Author's Accepted Manuscript

www.elsevier.com/locate/omega

Managing stochastic demand in an inventory routing problem with transportation procurement

Luca Bertazzi, Adamo Bosco, Demetrio Laganà

PII: S0305-0483(14)00121-2
DOI: http://dx.doi.org/10.1016/j.omega.2014.09.010
Reference: OME1437

To appear in: *Omega*

Received date: 30 July 2013
Revised date: 29 September 2014
Accepted date: 29 September 2014

Cite this article as: Luca Bertazzi, Adamo Bosco, Demetrio Laganà, Managing stochastic demand in an inventory routing problem with transportation procurement, *Omega,* http://dx.doi.org/10.1016/j.omega.2014.09.010

Managing Stochastic Demand in an Inventory Routing Problem with Transportation Procurement

Luca Bertazzi[*] Adamo Bosco[†] Demetrio Laganà[‡]

Abstract

We study an Inventory Routing Problem in which the supplier has a limited production capacity and the stochastic demand of the retailers is satisfied with procurement of transportation services. The aim is to minimize the total expected cost over a planning horizon, given by the sum of the inventory cost at the supplier, the inventory cost at the retailers, the penalty cost for stock–out at the retailers and the transportation cost. First, we show that a policy based just on the average demand can have a total expected cost infinitely worse than the one obtained by taking into account the overall probability distribution of the demand in the decision process. Therefore, we introduce a stochastic dynamic programming formulation of the problem that allows us to find an optimal policy in small size instances. Finally, we design and implement a matheuristic approach, integrating a rollout algorithm and an optimal solution of mixed–integer linear programming models, which is able to solve realistic size problem instances. Computational results allow us to provide managerial insights concerning the management of stochastic demand.

Keywords: Inventory Routing Problem, Stochastic Demand, Transportation Procurement, Dynamic Programming, Matheuristic.

[*]Department of Economics and Management, University of Brescia, Contrada Santa Chiara 50, 25122 Brescia, Italy, luca.bertazzi@unibs.it, tel: +39 030 2988585 (corresponding author)

[†]ITACA S.r.l., Ponte P. Bucci 41C, 87036 Arcavacata di Rende (CS), Italy, bosco@itacatech.it

[‡]Department of Mechanical, Energy and Management Engineering, University of Calabria, Ponte P. Bucci 41C, 87036 Arcavacata di Rende (CS), Italy, demetrio.lagana@unical.it

1 Introduction

In the last years, a large part of the logistic activities arising in the supply chain management were reorganized to take into account the impact of the Information Technology. The modern *Enterprise Resources Planning* (ERP) systems encouraged the integration of functions and data from various areas of the supply chain. The advantages of such integration stimulated the researchers to face new logistic challenges. In this context, a high integration level occurs, for example, in the *Vendor-Managed Inventory* (VMI) systems, where a single or multiple suppliers have to make simultaneous decisions to serve a set of retailers: 1) when to serve each retailer; 2) how much to deliver to each retailer, and 3) how to schedule the deliveries in each time period. The size of the deliveries can be defined according to different resupply policies. In general, the most useful policies are of two types: The *Order-up-to Level* (OU) policy, in which the quantity delivered to each retailer fulfills the storage capacity, and the *Maximum Level* (ML) policy, in which the supplier decides a delivery size that can not exceed the inventory capacity at each retailer. Solving the VMI implies to find solutions of a very hard combinatorial problem, named the *Inventory Routing Problem* (IRP), in which inventory and transportation decisions are integrated with the aim of minimizing the total cost resulting from using a specific resupply policy for a set of retailers. Operative constraints on the limited storage capacity at the retailers and the truckload capacity come into play when decisions have to be made.

In this paper, the focus is on the *Stochastic Inventory Routing Problem with Transportation Procurement* ($SIRP-TP$). In this problem, decisions on when and how much to deliver to a set of retailers are made over a finite and discrete planning horizon. The supplier has a limited production capacity at each time period, the demands of the retailers are modeled by discrete random variables and deliveries are performed using transportation procurement. We assume that the size of all deliveries is defined in accordance to the OU policy. This situation arises, for example, when inventory decisions related to a set of retailers occur in a VMI system where small package deliveries, fulfilling the capacity requirements, are frequently needed over a finite time horizon. Another application can be found in supermarket delivery, fuel delivery, refilling of vending machines. In these cases, transportation procurement can greatly reduce operating costs. An application is when transportation services are bought via Third Party Logistics (3PL) marketplaces that allow to pay the lowest transportation cost. Another application is when companies outsource last-mile deliveries of small quantities to carriers working as a freight contractors in a regional district. This allows a consolidation of the loads that significantly reduces the overall transportation cost. Additionally, administrative effort and legal problems of the vendor companies are reduced as the drivers are employees of a given transportation company or even self-employees. Outsourced deliveries are most likely to be efficient in large areas where a few clusters of retailers need to be frequently resupplied of small volumes over a given time horizon, and where the distances for traveling between the clusters are as long as unprofitable to be covered with a private fleet. The vendor company pays a fixed price to the carrier for using its (or a part of) truckload capacity. Consequently, the prices paid by the vendor company are independent from any routing decision of the freight contractor, but they may vary period by period of the time horizon according to the different size of the capacity leased by the carriers. A combinatorial auction is often used to obtain the lowest price. The aim of our problem is to minimize the total expected cost, represented by the sum of different cost components: 1) the total inventory cost at the supplier, 2) the total inventory cost at the retailers, 3) the total penalty cost arising whenever a stock-out occurs at the retailers, and 4) the total transportation cost to procure the transportation capacity needed to serve the retailers.

The $SIRP-TP$ can be defined as an Inventory Routing Problem, as it is a generalization of the Single Link Problem, i.e. the case with one supplier and one retailer, in which a fixed transportation cost is paid for each journey from the supplier to the retailer. The Single Link Problem is defined as an Inventory Routing Problem (see [1], **(author?)** [2] and **(author?)** [3]), as there is integration of inventory and transportation management over time.

The $SIRP-TP$ can be easily generalized in different ways. The extension to the case with several products is straightforward. The assumption about the transportation cost and capacity can be easily modified to obtain interesting variants of the problem, corresponding to different procurement modes of transportation services. The first is when the transportation capacity is infinite. This assumption corresponds to the case, very common in LTL, in which the sum of the maximum levels of inventory is very small with respect to the transportation capacity. The second is when several transportation capacities with different costs are made available at each time. In such a case there are several transportation companies and the supplier has to choose some of them at each time. The third is when the transportation cost function has a piecewise linear structure that alternates flat and increasing parts. This assumption corresponds to the classical case in which discount schemes are applied, i.e. the unit transportation cost decreases when the quantity increases (see for instance **(author?)** [4] and **(author?)** [5]).

The main scientific contributions of this paper can be summarized in the following points: a) providing a mathematical formulation of the $SIRP - TP$ with deterministic demand, b) proving that the $SIRP - TP$ is NP-hard even in the deterministic case, c) proving that the classical benchmark policy, based just on the average demand, can be infinitely worse than the optimal policy, d) providing a stochastic dynamic programming formulation of the problem, e) implementing an exact dynamic programming algorithm to determine an optimal policy in small instances, f) designing and implementing a rollout–based matheuristic algorithm to determine a near–optimal policy, g) providing managerial insights in the management of stochastic demands.

The remainder of the paper is organized as follows. A literature review is presented in Section 2. The problem is formally described in Section 3. The deterministic version of the problem and the corresponding complexity analysis is presented in Section 4. The benchmark policy is described and analyzed in Section 5. The dynamic programming formulation of the problem is presented in Section 6. The exact dynamic programming algorithm and the rollout–based matheuristic are described in Section 7. Computational results and the corresponding managerial insights are shown in Section 8.

2 Literature Review

Scientific research on Inventory Routing Problems (IRP) has sped up quickly in the last years. A large number of papers studying a few variants of the classical IRP have been produced, also inspired by real-cases. For a complete and well-organized analysis on the IRPs we refer to the recent survey of **(author?)** [6], **(author?)** [7], **(author?)** [1], **(author?)** [2] and **(author?)** [3]. Real cases are studied in **(author?)** [8], **(author?)** [9] and **(author?)** [10].

The most relevant contributions related to the one instant time horizon stochastic IRP are represented by the works proposed by **(author?)** [11], **(author?)** [12], **(author?)** [13], **(author?)** [14], **(author?)** [15], **(author?)** [16], **(author?)** [17] and **(author?)** [18]. In case of finite time horizon and stochastic demands, the major contributions are those proposed by **(author?)** [19], **(author?)** [20] and **(author?)** [21]. The mathematical model used to describe our problem is fundamentally rooted in modeling IRP with stochastic demand. In this context, we start the analysis from the pioneering paper by **(author?)** [22], who first proposed to model the coordination problem between different replenishment decisions as a Markov chain, in which the state in a given time period is defined by the inventory levels of all the retailers within such period, while the actions for moving from a state to another state are represented by routing decisions. The goal is to minimize the short and long-term cost of the action taken. This approach is applied to the setting of stochastic demands and with an unlimited number of vehicles by **(author?)** [22]. The same approach proposed by **(author?)** [22] is applied by **(author?)** [23] applied to an IRP with stochastic demands, in which a fleet with a limited number of vehicles is used and only direct deliveries are allowed. They developed a dynamic programming-based approximation method to compute the optimal value function of the IRP. Such IRP problem can be almost decomposed into as many subproblems as the retailers. Then, the solutions of these subproblems are used to estimate the expected value of long-term costs, while a greedy heuristic is developed to evaluate the expected total discounted value. In this approach stock-outs are allowed, but no backordering is permitted. In the next paper, **(author?)** [24] extended the approach proposed in their above work by allowing multiple deliveries per trip. In **(author?)** [25, 26], an IRP with stochastic demands is dealt by using the same approximated value function defined by **(author?)** [22] with the meaning of marginal prices for delivering. Such optimal dual prices are used into a linear program in order to approximate the future costs of current actions. Due to the infinite planning horizon, the main ideas on the basis of the works proposed by **(author?)** [27] and **(author?)** [28] is that the stocasticity can be modeled through a finite scenario tree. In **(author?)** [27] a GRASP algorithm is developed for the scenario tree problem which is constructed by assuming a progressive increase in the delivery volumes. In **(author?)** [28] a progressive hedging algorithm is applied to the problem, in which all identical sub-paths of the scenario tree are associated with the same decisions. Recently, **(author?)** [29] addressed a stochastic IRP with split delivery. The authors introduced different service levels at the suppliers and retailers with the aim of taking into account the stock-out. Such problem is modeled as an approximate stochastic IRP exploiting the transformation of stochastic components in deterministic ones, and near-optimally solved by using an hybrid approach based on a Lagrangian relaxation followed by local search techniques. A hybrid rollout algorithm for an IRP in which a supplier has to serve a set of retailers according to an OU level policy is proposed in **(author?)** [30]. For each retailer, a maximum inventory level is defined and a stochastic demand has to be satisfied over a given time horizon.

We now briefly review the papers with transportation procurement mostly related to our approach. In **(author?)** [31], a literature review on Third-Party Logistics (3PL) selection is presented. In **(author?)** [32], a single depot routing problem, named *Vehicle Routing Problem with Private fleet and Common carriers* (VRPPC), in which a set of retailers with known demand are served by a private fleet, and outsourced to a common carrier, is studied. The objective to be minimized consists in the classical transportation cost resulting from the variable travel costs, the fixed costs for vehicles, and fixed costs for the transportation procurement services. The heuristic algorithm proposed by **(author?)** [32] modifies the well-known **(author?)** [33] algorithm for the *Capacitated Vehicle Routing Problem* (CVRP) and performs a simple improvement phase. Very short computational experiments are presented for a dataset composed of only 5 instances. Several metaheuristic approaches for the VRPPC have been proposed in the last few years by **(author?)** [34], **(author?)** [35] and recently by **(author?)** [36] and **(author?)** [37]. The IRP with transshipments between customers is introduced by **(author?)** [38]. In **(author?)** [39], the problem of the selection of one or more companies in each echelon of a supply chain, and to manage production, inventory, and transportation is studied. In **(author?)** [40], a procurement setting in which a company needs to purchase a number of products from a set of suppliers to satisfy customer demand is studied. The suppliers offer total quantity discounts and transportation costs are based on truckload shipping rates. Combinatorial auctions are typically used in transportation procurement to determine the best offer of transportation services. We refer to **(author?)** [41] for an introduction of this tool, which could be used in our problem in a preliminary step, to determine the transportation cost and capacity available at each time period.

Referring to the main theme of this paper, the concept of controlling the stock-out at the retailers through a penalty cost due to the unsatisfied requests, combined with transportation procurement, has not yet been formally integrated within the context of IRP. If one looks at the recent survey on IRPs, one may observe that the most remarkable contributions in which the demands are stochastic relate to the cases with an infinite planning horizon and where a fleet composed of homogeneous vehicles is used for transportation. The problem dealt with in this paper is quite different. All the aforementioned considerations give the reasons for facing this problem and justify this research.

3 Problem Description

One product is shipped from a supplier 0 to a set $M = \{1, 2, \ldots, n\}$ of retailers over a given planning horizon H. The set of time periods is denoted as $T = \{1, 2, \ldots, H\}$. Each retailer $i \in M$ defines a maximum inventory level U_i and has a given starting inventory level $I_{i0} \leq U_i$, where U_i and I_{i0} are integer values. At each time $t \in T$, each retailer i has to satisfy the demand r_{it}, which is defined on the basis of a stationary random variable D_i. The probability distribution of D_i is discrete, denoted by $pr_i(k) = Pr(D_i = k)$, $k = 0, 1, \ldots, U_i$, and has mean q_i. The random variables D_i, $i \in M$, are independent. If i is visited at time t, then the quantity shipped to i at time t is such that the inventory level of i reaches its maximum value U_i (i.e. an OU policy is applied). More precisely, the shipped quantity is either equal to $U_i - I_{it}$ if a shipment to i is performed at time t, or equal to 0 otherwise. If we denote by z_{it} a binary variable equal to 1 if the retailer i is visited at time t and 0 otherwise, then the inventory level of the retailer i at time t is given by the maximum between 0 and the level at time $t-1$, plus the quantity $(U_i - I_{it-1})z_{it-1}$ shipped to i at time

$t-1$, minus the demand r_{it-1} at time $t-1$, that is

$$I_{it} = \max\{0, I_{it-1} + (U_i - I_{it-1})z_{it-1} - r_{it-1}\}, \tag{1}$$

where $z_{i0} = r_{i0} = 0$ and $\bar{I}_{i0}$ is given. Note that we are assuming that, when the level of the inventory is negative, the excess demand is not backlogged. Therefore, in this case the initial inventory level at the successive time period is set equal to zero. For each time $t \in T' = \{1, 2, \ldots, H+1\}$ and for each retailer $i \in M$, an inventory cost h_i is charged if the level of the inventory $I_{it-1} + (U_i - I_{it-1})z_{it-1} - r_{it-1}$ is positive, while a penalty cost $d_i > h_i$ is charged if it is negative.

The supplier has an initial integer inventory level $I_{00} = \bar{I}_{00}$. A deterministic integer quantity p_t is produced at each time $t \in T$. The inventory level of the supplier at time t is equal to the inventory level at time $t-1$, plus the quantity p_{t-1} produced at time $t-1$, minus the quantity shipped to the retailers at time $t-1$, that is

$$I_{0t} = I_{0t-1} + p_{t-1} - \sum_{i \in M} (U_i - I_{it-1})z_{it-1}. \tag{2}$$

For each time $t \in T'$, the inventory level I_{0t} cannot be negative and an inventory cost h_0 is charged if the inventory level is positive.

Deliveries are performed using transportation procurement. At each time $t \in T$, a transportation capacity C can be bought by paying a fixed cost f, to totally or partially use it.

The problem is to determine, for each time period $t \in T$, the subset of retailers to serve in order to minimize the sum of the expected inventory cost at the supplier, inventory cost at the retailers, penalty cost for stock–out at the retailers and transportation cost over the planning horizon.

Although the $SIRP-TP$ is the simplest stochastic inventory routing problem with transportation procurement, it is NP–hard even if the demand is deterministic and the transportation cost f is equal to zero (see Theorem 1).

4 The Deterministic Version

In this section, we study the particular case in which the demand is deterministic. We formulate a mixed–integer linear programming model and prove that this problem is

NP–hard. As a consequence, the $SIRP-TP$ is NP–hard. The optimal solution of the deterministic model is used in the next section to define a classical benchmark policy, based just on the average demand, instead on the probability distribution of the demand. We will show that this policy can be very suboptimal and therefore that taking into account the probability distribution of the demand is really important to be able to design good policies when the demand is stochastic. Finally, the optimal solution of this model will be embedded in the solution methods we propose to solve the $SIRP-TP$.

Let s_{it} be a non–negative variable representing the quantity sent to the retailer $i \in M$ at time $t \in T$; α_{it} be a non–negative variable representing the level of the inventory of the retailer i at time t, $i \in M$ and $t \in T'$; γ_{it} be a binary variable equal to 1 if $\alpha_{it} > 0$ and 0 otherwise; β_{it} be a non–negative variable representing the level of the stock–out of the retailer i at time t, $i \in M$ and $t \in T'$; δ_{it} be a binary variable equal to 1 if $\beta_{it} > 0$ and 0 otherwise; z_{it} be a binary variable equal to 1 if the retailer $i \in M$ is served at time $t \in T$ and 0 otherwise; y_t be a binary variable equal to 1 if at least one retailer is served at time $t \in T$ and 0 otherwise. The deterministic version of the problem can be formulated as follows.

Problem Det

$$\min \sum_{t \in T'} h_0 I_{0t} + \sum_{i \in M} \sum_{t \in T'} h_i \alpha_{it} + \sum_{i \in M} \sum_{t \in T'} d_i \beta_{it} + \sum_{t \in T} f y_t \tag{3}$$

$$I_{00} = \bar{I}_{00} \tag{4}$$

$$I_{0t} = I_{0t-1} + p_{t-1} - \sum_{i \in M} s_{it-1} \qquad t \in T' \tag{5}$$

$$\alpha_{i0} = \bar{I}_{i0} \qquad i \in M \tag{6}$$

$$\alpha_{it} - \beta_{it} = \alpha_{it-1} + s_{it-1} - r_{it-1} \qquad i \in M \quad t \in T' \tag{7}$$

$$\alpha_{it} \leq U_i \gamma_{it} \qquad i \in M \quad t \in T' \tag{8}$$

$$\beta_{it} \leq U_i \delta_{it} \qquad i \in M \quad t \in T' \tag{9}$$

$$\gamma_{it} + \delta_{it} \leq 1 \qquad i \in M \quad t \in T' \tag{10}$$

$$s_{it} \geq U_i z_{it} - \alpha_{it} \qquad i \in M \quad t \in T \tag{11}$$

$$s_{it} \leq U_i - \alpha_{it} \qquad i \in M \quad t \in T \tag{12}$$

$$s_{it} \leq U_i z_{it} \qquad i \in M \quad t \in T \tag{13}$$

$$\sum_{i \in M} s_{it} \leq C y_t \qquad t \in T \tag{14}$$

$$\sum_{i \in M} s_{it} \leq I_{0t} \qquad t \in T \tag{15}$$

$$I_{0t} \geq 0 \qquad t \in T' \tag{16}$$

$$\alpha_{it} \geq 0 \qquad i \in M \quad t \in T' \tag{17}$$

$$\beta_{it} \geq 0 \qquad i \in M \quad t \in T' \tag{18}$$

$$s_{it} \geq 0 \text{ integer} \qquad i \in M \quad t \in T \tag{19}$$

$$\gamma_{it} \in \{0,1\} \qquad i \in M \quad t \in T' \tag{20}$$

$$\delta_{it} \in \{0,1\} \qquad i \in M \quad t \in T' \tag{21}$$

$$z_{it} \in \{0,1\} \qquad i \in M \quad t \in T \tag{22}$$

$$y_t \in \{0,1\} \qquad t \in T \tag{23}$$

The objective function (3) expresses the minimization of the total cost, given by the sum of four terms: the inventory cost at the supplier, the inventory cost at the retailers, the penalty cost due to the stock–out at the retailers and the transportation cost. Constraints (4) and (5), where $s_{i0} = 0$, $i \in M$, and $p_0 = 0$, define the inventory level at the supplier. Constraints (6) and (7), where $s_{i0} = r_{i0} = 0$, $i \in M$, define the inventory and the stock-out level at the retailers. Constraints (8)-(10) ensure that, for each retailer $i \in M$ and for each time period $t \in T'$, either a positive inventory level α_{it} not greater than the maximum inventory level U_i is permitted or a stock-out quantity β_{it} not greater than U_i is permitted. The constraints (11)-(13) represent the order-up-to level constraints and guarantee that the quantity s_{it} shipped to each retailer i at each time $t \in T$ is either $U_i - \alpha_{it}$ if i is served at time t, and 0 otherwise. Note that, since the demand is not backlogged, the quantity $U_i - \alpha_{it}$ always allows to reach the order–up–to level when the retailer is served. Constraints (14) are the transportation capacity constraints. Constraints (15) guarantee that the total quantity sent to the retailers at each time $t \in T$ is not greater than the quantity available at the supplier. Finally, constraints (16)–(23) define the decision variable of the problem.

The following theorem shows that this problem is NP–hard, even in the case with $f = 0$. We use a reduction from the classical NP-hard 0–1 Knapsack Problem.

Theorem 1 *Problem Det is NP–hard.*

Proof The 0–1 Knapsack Problem is defined as follows: Given n objects of weight $w_i \in Z^+$ and value $v_i \in Z^+$, $i = 1, 2, \ldots, n$, and given a capacity $C < \sum_{i=1}^{n} w_i$, determine a subset A of $\{1, 2, \ldots, n\}$ with $\sum_{i \in A} w_i \leq C$ and of maximum value $\sum_{i \in A} v_i$ (see for instance **(author?)** [42]).

Given any instance of the 0-1 Knapsack Problem with n objects of weight w_i and size v_i, $i = 1, 2, \ldots, n$, and with capacity C, we construct an instance of the $SIRP-TP$ where $M = \{1, 2, \ldots, n\}$, $H = 1$, $\bar{I}_{i0} = 0$, $U_i = w_i$ $\forall i \in M$, $\bar{I}_{00} = \sum_{i \in M} U_i$, $h_0 = p_1 = 0$, $h_i < d_i = \frac{v_i}{U_i}$ $\forall i \in M$, $r_{i1} = U_i$ $\forall i \in M$ and $f = 0$.

Since $r_{i1} > 0$ and $\bar{I}_{i0} = 0$, all retailers require to be served at time 1. The quantity to deliver to each retailer $i \in M$ is U_i, as the OU policy is applied and $\bar{I}_{i0} = 0$. However, the transportation capacity $C < \sum_{i \in M} U_i$ forces some of the retailers to be not served. If retailer i is served at time 1, then $I_{i1} = I_{i2} = 0$ and therefore the corresponding total cost is 0, while if it is not served the cost is $d_i U_i$, as $r_{i1} = U_i$. Therefore, the objective function is $\min \sum_{i \in M} d_i U_i (1 - z_{i1})$, which is equivalent to $\max \sum_{i \in M} d_i U_i z_{i1}$. Therefore, the optimal solution of the problem can be obtained by optimally solving the following model:

$$\max \sum_{i \in M} d_i U_i z_{i1} \tag{24}$$

$$\sum_{i \in M} U_i z_{i1} \leq C \tag{25}$$

$$z_{i1} \in \{0, 1\} \tag{26}$$

Since we have defined $d_i U_i = v_i$ and $U_i = w_i$, the above problem is equivalent to the considered 0–1 Knapsack Problem. $\square$

Therefore, the following corollary holds.

Corollary 1 *$SIRP - TP$ is NP–hard.*

5 A Benchmark Policy

In this section, we aim at showing that policies obtained ignoring the probability distribution of the demands can be very suboptimal. In particular, we study the classical benchmark policy, referred to as *Policy EV*, based on the optimal solution of the deterministic model with demand equal to the average demand (see **(author?)** [43]). The average demand is a data that managers typically use to make decisions, while the probability distribution of the demand is often not known or not accurately known and, in any case, difficult to be embedded in the decision process.

The *Policy EV* can be formally described as follows:

Policy EV

1. *Demand Setting*: The demand r_{it} of each retailer i at each time t is set equal to the average demand q_i.

2. *Static Retailer Selection:* Problem *Det* is optimally solved by replacing the constraints (7) with the following constraints:

$$\alpha_{i1} = \alpha_{i0} \qquad i \in M \tag{27}$$

$$\alpha_{it} - \beta_{it} = \alpha_{it-1} + s_{it-1} - q_i \qquad i \in M \quad t = 2, 3, \ldots, H+1. \tag{28}$$

   We denote with z^*_{it} the optimal value of the variable z_{it}.

3. *Dynamic Service:* For each time period $t \in T$ at a time, starting from $t = 1$, first the quantity to deliver to the retailers is computed as $\sum_{i \in M}(U_i - \hat{I}_{it})z^*_{it}$, where $\hat{I}_{it}$ is the inventory level of retailer i at time t. If this quantity is greater than the transportation capacity C and/or greater than inventory level $\hat{I}_{0t}$ at the supplier at time t, the following Problem $\mathcal{Q}_t$ is optimally solved. Let θ_{it} be a binary variable equal to 1 if the quantity $(U_i - \hat{I}_{it})$ is delivered to retailer i and 0 otherwise.

Problem $\mathcal{Q}_t$

$$\max \sum_{i \in M} d_i (U_i - \hat{I}_{it})\theta_{it} \tag{29}$$

$$\sum_{i \in M} (U_i - \hat{I}_{it})\theta_{it} \leq C \tag{30}$$

$$\sum_{i \in M} (U_i - \hat{I}_{it})\theta_{it} \leq \hat{I}_{0t} \tag{31}$$

$$\theta_{it} \leq z^*_{it} \qquad i \in M \tag{32}$$

$$\theta_{it} \in \{0, 1\} \qquad i \in M \tag{33}$$

The rationale of Problem $\mathcal{Q}_t$ is to serve the retailers with high penalty cost selected among the ones served at time t in the optimal solution of Problem Det. Given the optimal solution of this problem, the demand at time t is revealed and the final inventory levels at time t are computed. For each positive level, the corresponding inventory cost is charged, while for each negative level the corresponding penalty cost is charged. Finally, the starting inventory levels at time $t+1$ are computed.

We now show that this policy can be infinitely worse than an optimal policy based on the probability distribution of the demand.

Theorem 2 *There exists an instance $\mathcal{I}$ such that $\frac{E[z^{EV}(\mathcal{I})]}{E[z^*(\mathcal{I})]} \to \infty$.*

Proof Consider the following instance $\mathcal{I}$: Time horizon $H = 2$, one retailer, $U_1 = 10$, $pr_1(0) = 0.5$, $pr_1(k) = 0$ for $k = 2, 3, \ldots, 9$, $pr_1(10) = 0.5$, $\bar{I}_{10} = 0$, $\bar{I}_{00} = 20$, $h_0 = p_1 = p_2 = 0$, $h_1 = 1$, $d_1 > 1$, $C = 10$ and $f = 1$.

Since the average demand is 5 and $\bar{I}_{1,0} = 0$, the optimal solution of the Problem Det with $\hat{r}_{1t} = q_1$, for each $t \in T$, is to serve the retailer just at time 1, delivering 10 units. Five units are used to satisfy the average demand at time 1 and five to satisfy the average demand at time 2. No stock–out occurs. Consider now the four scenarios of realization of the demand:

1. $r_{11} = r_{12} = 0$ with probability $1/4$: In this case, the application of the *Policy EV* gives a total cost of $h_1(\alpha_{11} + \alpha_{12} + \alpha_{13}) + d_1(\beta_{11} + \beta_{12} + \beta_{13}) + f(y_1 + y_2) = (0 + 10 + 10) + d_1(0 + 0 + 0) + 1(1 + 0) = 21$.

2. $r_{11} = 0$, $r_{12} = 10$ with probability $1/4$: In this case, the application of the *Policy EV* gives a total cost of $(0 + 10 + 0) + d_1(0 + 0 + 0) + 1(1 + 0) = 11$.

3. $r_{11} = 10$, $r_{12} = 0$ with probability 1/4: In this case, the application of the *Policy EV* gives a total cost of $(0+0+0) + d_1(0+0+0) + 1(1+0) = 1$.

4. $r_{11} = r_{12} = 10$ with probability 1/4: In this case, the application of the *Policy EV* gives a total cost of $(0+0+0) + d_1(0+0+10) + 1(1+0) = 1 + 10d_1$.

Therefore, $E[z^{EV}(\mathcal{I})] = \frac{1}{4}(21 + 11 + 1 + 10d_1) = \frac{33}{4} + \frac{5}{2}d_1$.

Given that the probability distribution of the demand is such that only a demand equal to 0 or 10 takes place, a better solution is obtained by a different policy in which 10 units are sent to the retailer whenever the inventory level is lower than 10. Doing that, the total cost is $(0+10+10)+d_1(0+0+0)+1 = 21$, $(0+10+0)+d_1(0+0+0)+1 = 11$, $(0+0+10) + d_1(0+0+0) + 2 = 12$ and $(0+0+0) + d_1(0+0+0) + 2 = 2$, respectively, in the four scenarios. Therefore, $E[z^*(I)] \leq \frac{1}{4}(21 + 11 + 12 + 2) = \frac{23}{2}$.

Hence

$$\frac{E[z^{EV}(\mathcal{I})]}{E[z^*(\mathcal{I})]} \geq \frac{\frac{33}{4} + \frac{5}{2}d_1}{\frac{23}{2}} \to \infty \qquad \text{for } d_1 \to \infty.$$

$\square$

The average performance of this policy will be shown in Section 8. These results will confirm that it is really important to embed the probability distribution of the demand in the decision process to be able to design good policies.

6 A Stochastic Dynamic Programming Formulation

In this section, we present a stochastic dynamic programming formulation (DP, for short) for the $SIRP-TP$, able to deal with the probability distribution of the demand in making decisions. We refer to **(author?)** [44] and **(author?)** [45] for comprehensive books on DP.

6.1 States

The state x_t at time $t = 0, 1, \ldots, H+1$ is given by the level of the inventory at the supplier and at each retailer $i \in M$:

$$x_t = (x_{0t}, x_{1t}, x_{2t}, \ldots, x_{nt}).$$

The initial state x_0 is given and equal to $(\bar{I}_{00}, \bar{I}_{10}, \ldots, \bar{I}_{n0})$. The state at time $t \in T$ is such that x_{0t} is an integer number $\bar{I}_{00} + \sum_{k=1}^{t-1} p_k - (t-1)C \leq x_{0t} \leq \bar{I}_{00} + \sum_{k=1}^{t-1} p_k$ and x_{it} is an integer number $0 \leq x_{it} \leq U_i$, $i \in M$. The terminal state x_{H+1} is such that x_{0H+1} is an integer number $\bar{I}_{00} + \sum_{k=1}^{H} p_k - HC \leq x_{0t} \leq x_{0H+1} \leq \bar{I}_{00} + \sum_{k=1}^{H} p_k$ and x_{iH+1} is an integer number $-U_i \leq x_{iH+1} \leq U_i$, $i \in M$.

6.2 Controls

The control $u_t(x_t)$ at time $t \in T$ corresponding to state x_t can be defined by the means of the binary variables z_{it}, $i \in M$:

$$u_t(x_t) = (z_{1t}, z_{2t}, \ldots, z_{nt}).$$

In fact, once the value of the variables z_{it}'s is defined, then the corresponding delivery quantities are given, as an OU policy is applied. The set $\mathcal{U}_t(x_t)$ of feasible controls at time t and state x_t is given by the controls $u_t(x_t)$ that satisfy the capacity constraint of the vehicle, that is

$$\sum_{i \in M} (U_i - x_{it}) z_{it} \leq C, \tag{34}$$

and do not exceed the quantity available at the supplier, that is

$$\sum_{i \in M} (U_i - x_{it}) z_{it} \leq x_{0t}. \tag{35}$$

6.3 Dynamic System

The discrete–time system can be described as follows. Suppose that x_t is the state of the system at time t, u_t is the control applied at time t and r_t is the vector of the demands r_{it}, $i \in M$, at time t. If we denote by $\hat{x}_{0t} = x_{0t} + p_t - \sum_{i \in M} (U_i - x_{it}) z_{it}$ and by $\hat{x}_{it} = x_{it} + (U_i - x_{it}) z_{it} - r_{it}$, then the state of the system at time $t+1$ is:

$$x_{t+1} = (\hat{x}_{0t}, \max\{0, \hat{x}_{1t}\}, \max\{0, \hat{x}_{2t}\} \ldots, \max\{0, \hat{x}_{nt}\}),$$

as the demand is not backlogged. The terminal state is

$$x_{H+1} = (\hat{x}_{0H}, \hat{x}_{1H}, \ldots, \hat{x}_{nH}).$$

6.4 Costs

The immediate cost $g(x_t, u_t, r_t)$ to be in the state x_t, to apply the control u_t and to have the demand r_t at time t is given by the sum of the inventory, penalty and transportation costs, that is:

$$g(x_t, u_t, r_t) = \sum_{i \in M} h_i \max\{0, \hat{x}_{it}\} + \sum_{i \in M} d_i \max\{0, -\hat{x}_{it}\} + f\Phi(\sum_{i \in M} z_{it}),$$

where $\Phi(\omega)$ is equal to 1 if $\omega > 0$ and 0 otherwise.

At time $H+1$, the terminal cost $g(x_{H+1})$ is 0 for each terminal state x_{H+1}.

6.5 Optimization Problem

Our aim is to find a policy that minimizes the expected total cost. We consider the set Π of feasible policies. Each of these policies consists of a sequence of functions $\pi = \{\mu_1, \mu_2, \ldots, \mu_H\}$, where μ_t maps each state x_t into a control $u_t = \mu_t(x_t)$ and is such that $\mu_t(x_t) \in \mathcal{U}_t(x_t)$ for all states x_t, $t \in T$. Starting from the given initial state x_0, the total expected cost of π is:

$$J_\pi(x_0) = E\left\{\sum_{t=1}^{H} g(x_t, \mu_t(x_t), r_t) + g_{H+1}(x_{H+1})\right\}.$$

Our aim is to find a policy that minimizes the total expected cost, that is a policy π^* such that:

$$J_{\pi^*}(x_0) = \min_{\pi \in \Pi} J_\pi(x_0).$$

7 Dynamic Programming Policies

The formulation provided in the previous section allows us to design exact and heuristic solution algorithms for the $SIRP-TP$ able to take into account the probability distribution of the demand in the decision process.

We first implement the classical *Exact DP* algorithm to determine the optimal policy of the $SIRP-TP$. Given the initial state $x_0 = (\bar{I}_{00}, \bar{I}_{10}, \ldots, \bar{I}_{n0})$, the optimal cost $J^*(x_0)$ is equal to $J_0(x_0)$, given by the last step of the following algorithm, which proceeds backward in time from period $H+1$ to period 1:

$$J_{H+1}(x_{H+1}) = g(x_{H+1}) \qquad \forall x_{H+1}$$

$$J_t(x_t) = \min_{u_t(x_t)\in\mathcal{U}_t(x_t)} E\{g(x_t, u_t, r_t) + J_{t+1}(x_{t+1})\} \qquad \forall x_t \quad t = 1, 2, \ldots, H.$$

The expectation is taken with respect to the probability distributions of the demand r_{it}, $i \in M$.

The availability of the optimal policy allows us to evaluate the performance of the heuristic policies in small instances, such as the *Policy EV* and the matheuristic described in the following. Instead, this algorithm cannot be used to solve realistic size instances of the problem. In fact, it suffers the three curses of dimensionality:

1. *State space*: The cardinality of the state space at each time $t \in T$ is related to $(\bar{U}+1)^n$, where $\bar{U} = \max_{i\in M} U_i$, as each retailer can have inventory level equal to $0, 1, \ldots, U_i$; moreover, note that this value has to be multiplied by the number of possible inventory level of the supplier;
2. *Control space*: The cardinality of the set of feasible controls is related to 2^n, as a binary variable is used for each retailer $i \in M$;
3. *Outcome space*: The cardinality of the set of possible outcomes at each time $t \in T$ is related to $(\bar{U}+1)^n$, as $U_i + 1$ values of demand are defined for each retailer $i \in M$.

Therefore, we propose a heuristic algorithm to be able to solve realistic size instances. In particular, we design and implement a matheuristic based on the rollout algorithm. A matheuristic is a new type of heuristic algorithm that makes use of mathematical programming models, typically Mixed Integer Linear Programming models (MILP), inside a heuristic/metaheuristic scheme (see, e.g., **(author?)** [46]). The computational effectiveness of commercial optimization software makes it interesting and promising the design of heuristic approaches that make use of the optimal solution of MILPs. We refer to **(author?)** [47] for an introduction to matheuristics for the solution of deterministic Inventory Routing Problems. The matheuristic we propose is rollout–based, as the heuristic scheme we use is a rollout algorithm. Rollout algorithms are a class of heuristic algorithms more and more frequently applied to solve deterministic and stochastic dynamic programming problems. The basic idea is to use, in a one-step lookahead policy, the cost of a well known heuristic, called base policy, to approximate the value of the optimal cost-to-go. These algorithms are very appealing from the practical point of view, as they are easy to be implemented and guarantee a not worse, and usually much better, performance than the corresponding base policies (see **(author?)** [44] and **(author?)** [48]). Rollout algorithms have been originally proposed in the context of Neuro-Dynamic Programming/Reinforcement Learning (see **(author?)** [49], **(author?)** [50] and **(author?)** [51]). They have been applied to stochastic scheduling by **(author?)** [52], to vehicle routing problems with stochastic demand by **(author?)** [53, 54, 55] and **(author?)** [56]. Computational papers related to rollout algorithms can be found in **(author?)** [57, 58]. They have been proposed for the solution of combinatorial problems by **(author?)** [49], **(author?)** [59] and **(author?)** [60]. They have been applied to multi–dimensional knapsack problems by **(author?)** [61]. A rollout algorithm for the solution of a stochastic Inventory Routing Problem with stock–out, with routing costs, has been proposed in [30].

In the matheuristic we propose, referred to as *Policy M*, the approximate rollout control at time $t = 1, 2, \ldots, H$ corresponding to state x_t is

$$\tilde{\mu}_t(x_t) = \arg \min_{u_t(x_t) \in \tilde{\mathcal{U}}_t(x_t)} \tilde{Q}_t(x_t, u_t(x_t))$$

where $\tilde{Q}_t(x_t, u_t(x_t)) = E\left\{g(x_t, u_t, r_t) + \tilde{J}_{t+1}(x_{t+1})\right\}$ is an approximate Q-factor, $\tilde{J}_{t+1}(x_{t+1})$ is an approximate cost–to–go and $\tilde{\mathcal{U}}_t(x_t)$ is an approximate set of controls. The approximate cost–to–go and the approximate set of controls are computed by solving MILP models.

Let us now describe how the approximate Q-factor, $\tilde{J}_{t+1}(x_{t+1})$ and $\tilde{\mathcal{U}}_t(x_t)$ are designed. Consider the state x_t and the control $u_t(x_t)$. To compute the approximate Q-factor $\tilde{Q}_t(x_t, u_t(x_t))$, we first generate a set $\mathcal{S}_t(x_t, u_t)$ of scenarios. Each scenario $s \in \mathcal{S}_t(x_t, u_t)$ is defined by a realization r_t of the demand of each retailer at time t. Three scenarios are obtained by setting r_{it} to 0 (minimum value), q_i (average value) and U_i (maximum value) for each retailer $i \in M$, respectively. These scenarios try to capture the two extreme cases (minimum and maximum demand for each retailer) and the most likely case (average demand for each retailer). The remaining scenarios are obtained by randomly generating r_{it} for each retailer i, according to the corresponding distribution probability. Let $S_t(x_t, u_t)$ be the total number of these additional scenarios. For each of these scenarios, it is possible to exactly compute the immediate cost $g(x_t, u_t, r_t)$ and to compute the corresponding state x_{t+1} at time $t+1$. The approximate cost–to–go $\tilde{J}_{t+1}(x_{t+1})$ is obtained as follows. We consider three scenarios at time $t+1$, obtained by setting the demand r_{it+1} of each retailer $i \in M$ to 0, q_i and U_i, respectively. For each of these scenarios, we solve Problem Det from $t+1$ to H, by setting the demand at time $t+1$ to the corresponding value, the demands from time $t+2$ to H to the average value q_i and the initial inventory level of each retailer $i \in M$ equal to $\bar{x}_{it+1}$. Then, $\tilde{J}_{t+1}(x_{t+1})$ is computed as the average value of the optimal costs obtained by solving Problem Det. Since each scenario can be generated several times, the approximate Q-factor $\tilde{Q}_t(x_t, u_t(x_t))$ is the average value of the costs of the different scenarios, where the weight of each scenario is the relative frequency of the scenario.

The approximate set of controls $\tilde{\mathcal{U}}_t(x_t)$ associated to each state x_t at each time t is composed of four types of controls: The control in which no retailer is served, the controls obtained by exactly solving Problem Det with demand of each retailer from t to H equal to the average value, the controls obtained by exactly solving four MILP models and the controls obtained by a heuristic algorithm. The four models are obtained by solving the following model with two different objective functions and two different scenarios, respectively. The two objective functions are

$$\max \quad \sum_{i \in M} z_{it} \tag{36}$$

$$\min \quad \sum_{i \in M} h_i \alpha_{it+1} + \sum_{i \in M} \beta_i d_{it+1} \tag{37}$$

while the two scenarios are: $r_{it} = 0$, $\forall i \in M$, and $r_{it} = U_i$, $\forall i \in M$. The constraints are the following, given the inventory levels at the supplier $\bar{I}_{0t}$ and at the retailers $\bar{I}_{it}$, $i \in I$:

$$I_{0t+1} = \bar{I}_{0t} + p_t - \sum_{i \in M} s_{it} \tag{38}$$

$$\alpha_{it+1} - \beta_{it+1} = \bar{I}_{it} + s_{it} - r_{it} \qquad i \in M \tag{39}$$

$$\alpha_{it+1} \leq U_i \gamma_{it+1} \qquad i \in M \tag{40}$$

$$\beta_{it+1} \leq U_i \delta_{it+1} \qquad i \in M \tag{41}$$

$$\gamma_{it+1} + \delta_{it+1} \leq 1 \qquad i \in M \tag{42}$$

$$s_{it} \geq U_i z_{it} - \bar{I}_{it} \qquad i \in M \tag{43}$$

$$s_{it} \leq U_i - \bar{I}_{it} \qquad i \in M \tag{44}$$

$$s_{it} \leq U_i z_{it} \qquad i \in M \tag{45}$$

$$z_{it} \leq s_{it} \qquad i \in M \tag{46}$$

$$\sum_{i \in M} s_{it} \leq C \tag{47}$$

$$\sum_{i \in M} s_{it} \leq I_{0t} \tag{48}$$

$$\alpha_{it+1} \geq 0 \qquad i \in M \tag{49}$$

$$\beta_{it+1} \geq 0 \qquad i \in M \tag{50}$$

$$s_{it} \geq 0 \text{ integer} \qquad i \in M \tag{51}$$

$$\gamma_{it+1} \in \{0,1\} \qquad i \in M \tag{52}$$

$$\delta_{it+1} \in \{0,1\} \qquad i \in M \tag{53}$$

$$z_{it} \in \{0,1\} \qquad i \in M \tag{54}$$

Constraint (38) computes the inventory level at the supplier at time $t+1$. Constraints (39) compute the inventory level at the retailers at time $t+1$. Constraints (40)–(42) guarantee that either an inventory level not greater than U_i or a stock–out not greater than $-U_i$ is reached for each retailer i. Constraints (43)–(45) are the OU policy constraints. Constraints (46) guarantee that $z_{it} = 0$ whenever $s_{i0} = 0$. Constraints (47)–(48) guarantee that the transportation capacity and the inventory level at the supplier are satisfied. Finally, constraints (49)–(54) define the decision variables of the problem.

Finally, the fourth type of controls is generated by applying the following heuristic algorithm, referred to as *Heuristic Ctrl*. An iteration of such an algorithm consists in selecting a retailer which state is modified from not served to served. For each iteration ℓ, the algorithm starts with a given solution and modifies it to obtain new (possibly feasible) solutions. Let $\bar{M} = \{z_1, z_2, \ldots, z_n\}$ be the current solution, where z_i is equal to 1 if the retailer i is served and 0 otherwise, with $i \in M$. The initial solution is $\bar{M} := \{0, 0, \ldots, 0\}$ and then is updated at each iteration. Let C_ℓ be the residual transportation capacity, I_ℓ be the residual inventory level at the supplier at the beginning of iteration ℓ.

1. Set $\ell := 0$, $\bar{M} := \{0, 0, \ldots, 0\}$, $C_\ell := C$, $I_\ell := \bar{I}_{i0}$.

2. Order $\bar{M}$ in a non-increasing way according to $d_i \max\{0, r_{it} - \bar{I}_{it}\}$.

3. While $\sum_{i \in M} z_i < n$ and $C_\ell > 0$ and $\bar{I}_\ell > 0$:

   (a) Select the first retailer, say j, in $\bar{M}$ with $z_j = 0$.

   (b) Generate a new control by setting $z_j := 1$.

   (c) Update the residual capacity, i.e. $C_\ell := C_\ell - \left(U_j - \bar{I}_{jt}\right)$, and the residual inventory level at the supplier: $\bar{I}_\ell := \bar{I}_\ell - (U_j - \bar{I}_{jt})$.

   (d) Check the feasibility of the new control: if $C_\ell \geq 0$ and $\bar{I}_\ell \geq 0$, then insert the control into the list of controls generated by the heuristic.

   (e) Set $\ell := \ell + 1$.

8 Computational Results

The *Policy EV*, the *Exact DP* algorithm and the *Policy M* have been implemented in Java and evaluated on randomly generated benchmark instances on a PC Intel(R) Xeon(R) CPU X5680 3.33GHz RAM 47 GB.

Two computational experiments have been carried out. The first aims at evaluating the performance of the heuristic algorithms with respect to the optimal policy or to a lower bound. The second aims at providing managerial insights.

The sets of instances are based on the instances proposed in **(author?)** [62] for a deterministic Inventory Routing Problem without stock–out and with routing costs. In particular, the following data are exactly the ones in **(author?)** [62]:

1. Inventory cost at the retailers h_i: In **(author?)** [62], they were randomly generated in the intervals $[0.01, 0.05]$ and $[0.1, 0.5]$;

2. Inventory cost at the supplier h_0: 0.03 when $h_i \in [0.01, 0.05]$, 0.3 when $h_i \in [0.1, 0.5]$.

Moreover, we fixed the following data, both in the small and realistic size instances:

1. Transportation capacity C: $= \frac{3}{2} \sum_{i \in M} q_i$ (as in **(author?)** [62]);

2. Transportation cost f: 10;

3. Distribution probability D_i of the demand of retailer $i \in M$: Binomial distribution $B(U_i, \frac{q_i}{U_i+1})$;

4. Initial inventory level at the retailers I_{i0}: $U_i - q_i$ (as in **(author?)** [62]);

5. Initial inventory level at the supplier I_{00}: $\sum_{i \in M} U_i$ (as in **(author?)** [62]);

6. Quantity p_t produced by the supplier at each time t: $\sum_{i \in M} q_i$ (as in **(author?)** [62]);

7. Penalty cost at the retailers d_i: $1.5f + h_i U_i$;

8. Number $S_t(x_t, u_t)$ of additional scenarios generated at time t: 100.

9. Time limit for the solution of each MILP: 7000 seconds.

8.1 Performance of the Heuristic Policies

The aim of the first computational experiment is to show the performance of the *Policy EV* and the *Policy M*. Given to the computational hardness of the *Exact DP* algorithm, a set of 9 instances is generated on the basis of the data described above and the following additional data, to compare the *Policy EV* and the *Policy M* with respect to the optimal policy:

1. Time horizon H: 3;

2. Average demand q_i of retailer $i \in M$: Randomly generated in $\{1, 2\}$ from a uniform distribution;

3. Number of retailers n: 2;

4. Maximum level U_i of retailer $i \in M$: If q_i is equal to 2, then $U_i = 2$; otherwise, randomly generated in $\{1, 2\}$ from a uniform distribution.

Table I shows the obtained results. The first column gives the number of the instance, while the last two columns show the percent increase of the expected cost of *Policy EV* and *Policy M* with respect to the optimal expected cost obtained by applying the *Exact DP* algorithm. Let $E[z^{EV}]$ and $E[z^*]$ be the expected cost obtained by applying *Policy EV* and the *Exact DP* algorithm, respectively. The percent increase of *Policy EV* is $\frac{E[z^{EV}]-E[z^*]}{E[z^*]} * 100$ and the one of *Policy M* is computed accordingly.

Instance	*Policy EV*	*Policy M*
1	5.00	0.15
2	18.03	0.10
3	0.80	0.00
4	5.96	0.11
5	12.49	0.92
6	14.98	0.68
7	15.17	0.35
8	3.57	0.08
9	14.20	0.00
Average percent increase	10.02	0.27

Table I: Heuristics vs Optimal: Percent increase

The results show that *Policy M* gives an expected cost very close to the optimal one. Moreover, note that *Policy M* significantly outperforms *Policy EV*.

To evaluate the performance of the *Policy EV* and *Policy M* on realistic size instances, we defined the following data:

1. Time horizon H: 3, 6;
2. Average demand q_i of retailer $i \in M$: Equal to the demand of each retailer i in **(author?)** [62], were it was randomly generated as an integer number in the interval $[10, 100]$;
3. Number of retailers n: $5k$, with $k = 1, 2, 4$;
4. Maximum level U_i of retailer $i \in M$: Equal to the same value in **(author?)** [62], where it was generated as $q_i g_i$, where g_i were randomly selected from the set $\{2, 3\}$.

Since the optimal expected cost cannot be computed on these instances and the complete enumeration of the trajectories cannot be performed, for each instance we generated 10 trajectories by randomly extracting a realization of the demands according to their probability distributions. For each trajectory, we applied the heuristic algorithms and optimally solved Problem *Det* to obtain a lower bound on

n	h_i	*Policy EV*	*Policy M*
5	[0.01,0.05]	43.78	2.78
5	[0.01,0.05]	146.96	10.75
5	[0.01,0.05]	293.02	7.17
5	[0.01,0.05]	164.35	3.98
5	[0.01,0.05]	265.09	11.88
10	[0.01,0.05]	289.54	2.21
10	[0.01,0.05]	304.66	3.03
10	[0.01,0.05]	190.72	3.98
10	[0.01,0.05]	150.87	1.53
10	[0.01,0.05]	320.48	4.80
20	[0.01,0.05]	266.58	1.98
20	[0.01,0.05]	241.60	2.13
20	[0.01,0.05]	202.42	2.84
20	[0.01,0.05]	232.57	1.43
20	[0.01,0.05]	222.09	3.35
Average percent increase		222.32	4.26
5	[0.1,0.5]	51.70	0.56
5	[0.1,0.5]	36.52	1.51
5	[0.1,0.5]	131.54	5.87
5	[0.1,0.5]	8.40	0.82
5	[0.1,0.5]	122.72	6.56
10	[0.1,0.5]	85.75	2.89
10	[0.1,0.5]	111.39	2.31
10	[0.1,0.5]	65.21	3.39
10	[0.1,0.5]	22.59	1.54
10	[0.1,0.5]	130.87	4.97
20	[0.1,0.5]	78.40	2.45
20	[0.1,0.5]	72.88	1.91
20	[0.1,0.5]	72.21	2.98
20	[0.1,0.5]	34.78	1.52
20	[0.1,0.5]	73.92	3.50
Average percent increase		73.26	2.85

Table II: Heuristics vs Lower Bound: Percent increase when $H = 3$

n	h_i	*Policy EV*	*Policy M*
5	[0.01,0.05]	201.74	1.68
5	[0.01,0.05]	112.27	6.45
5	[0.01,0.05]	225.43	5.54
5	[0.01,0.05]	93.18	1.67
5	[0.01,0.05]	295.52	8.91
10	[0.01,0.05]	215.40	2.63
10	[0.01,0.05]	229.28	3.26
10	[0.01,0.05]	176.04	4.06
10	[0.01,0.05]	191.74	1.78
10	[0.01,0.05]	243.47	4.66
20	[0.01,0.05]	242.24	1.69
20	[0.01,0.05]	201.51	1.79
20	[0.01,0.05]	246.77	2.17
20	[0.01,0.05]	168.45	1.19
20	[0.01,0.05]	214.04	3.07
Average percent increase		203.81	3.37
5	[0.1,0.5]	70.71	0.47
5	[0.1,0.5]	42.40	1.45
5	[0.1,0.5]	91.67	4.00
5	[0.1,0.5]	17.49	1.33
5	[0.1,0.5]	143.18	5.59
10	[0.1,0.5]	51.94	2.94
10	[0.1,0.5]	74.37	2.25
10	[0.1,0.5]	57.83	3.61
10	[0.1,0.5]	35.04	1.72
10	[0.1,0.5]	93.84	4.75
20	[0.1,0.5]	70.59	2.22
20	[0.1,0.5]	57.20	1.62
20	[0.1,0.5]	56.06	2.28
20	[0.1,0.5]	34.21	1.25
20	[0.1,0.5]	65.27	3.17
Average percent increase		64.12	2.58

Table III: Heuristics vs Lower Bound: Percent increase when $H = 6$

	Uniform	Binomial	Poisson
Total cost Policy EV	4998.79	880.58	1162.34
Total cost Policy M	1080.35	466.59	464.85
% Inventory Supplier	38.10%	51.42%	51.41%
% Inventory Retailers	21.49%	32.17%	32.06%
% Penalty	28.19%	0.00%	0.00%
% Transportation	12.21%	16.41%	16.52%
Number of visits	2.76	3.00	3.00
Delivered quantity	145.33	127.88	129.12

Table IV: Managerial Insights

the cost of the trajectory. Tables II and III show the obtained results. In particular, Table II concerns the case with time horizon $H = 3$, while Table III the case with $H = 6$. In each table, the first two columns give the number of retailers and the range of the inventory cost. The last two columns shows the ratio between the average cost of the heuristic policies and the average cost of Problem *Det*.

The results show that *Policy M* significantly outperforms the benchmark policy, confirming that it is really important to take into account the probability distribution of the demand in the decision process. Moreover, the average cost of *Policy M* is very close to the lower bound. Therefore, this policy is able to successfully solve the problem also in realistic size instances.

8.2 Managerial Insights

We now aim at giving managerial insights concerning the management of the stochastic demand on the basis of the solution obtained by *Policy M*. In order to do that, we run 100 instances with 5 retailers generated as described before, but with demand of each retailer i defined on the basis of three different probability distributions (Uniform, Binomial and Poisson), having mean q_i. In Table IV, we compare the expected cost of the *Policy M* with the expected cost of the *Policy EV*. Then, we show the composition of the expected cost of the *Policy M* in terms of inventory cost at the supplier, inventory cost at the retailers, penalty cost for stock–out and fixed transportation cost. Finally, we show the average number of visits to each retailer and the average delivery quantity.

The results show that the average number of visits and the average delivery quantity are quite independent of the probability distribution. In fact, in order to avoid stock–out, it is better to serve the retailers frequently. However, the total cost of the *Policy M* is much greater in the case of Uniform distribution with respect to the Binomial and Poisson distributions, due to the penalty cost for stock-out. We can also note that in this case the *Policy EV*, that completely ignores the probability distribution, is very poor. Therefore, we can conclude that, in order to reduce the total cost, it is better to serve the retailers frequently and that it is really important to embed the probability distribution of the demand in the decision process.

9 Conclusion

We studied the simplest *Stochastic Inventory Routing Problem with Transportation Procurement*. The theoretical and computational results provided in the paper allow us to conclude that this problem is really difficult to be solved to optimality. However, a matheuristic, based on the integration of a rollout algorithm with an optimal solution of mixed–integer linear programming models, is able to find near–optimal policies. We can also conclude that it is really important to embed the probability distribution of the demand in the decision process. In fact, policies based just on the average demand can be infinitely worse than the optimal policy. Moreover, since the expected cost significantly depends on the probability distribution of the demand, it is really important to have an accurate information about the probability distribution of the demand. Future research could be devoted to adapt the matheuristic approach to more complex transportation cost functions.

Acknowledgments

The authors wish to thank the Editor–in–Chief and an anonymous reviewer for useful suggestions that allowed us to improve the paper.

References

[1] L. Bertazzi and M.G. Speranza. Inventory routing problems: an introduction. *EURO Journal on Transportation and Logistics*, 1:307–326, 2012.

[2] L. Bertazzi and M.G. Speranza. Inventory routing problems with multiple customers. *EURO Journal on Transportation and Logistics*, 2:255–275, 2013.

[3] L.C. Coelho, J.F. Cordeau, and G. Laporte. Thirty years of inventory routing. *Transportation Science*, 2013.

[4] S. Nahmias. *Production and operations analysis.* McGraw-Hill, 2005.

[5] C. Archetti, L. Bertazzi, and M.G. Speranza. Polynomial cases of the economic lot sizing problem with cost discounts. *European Journal of Operational Research*, 2014.

[6] L. Bertazzi, M. Savelsbergh, and M.G. Speranza. *Inventory Routing, in The Vehicle Routing Problem, Latest Advances and New Challenges.* B. Golden and S. Raghavan and E. Wasil, editors, Springer, 2008.

[7] H. Andersson, A. Hoff, M. Christiansen, G. Hasle, and A. Løkketangen. Industrial aspects and literature survey: Combined inventory management and routing. *Computers & Operations Research*, 37:1515–1536, 2010.

[8] J. Kim and Y. Kim. A decomposition approach to a multi-period vehicle scheduling problem. *Omega*, 27(4):421–430, 1999.

[9] J.M. Day, P.D. Wright, T. Schoenherr, M. Venkataramanan, and K. Gaudette. Improving routing and scheduling decisions at a distributor of industrial gasses. *Omega*, 37(1):227–237, 2009.

[10] G.M. Kopanos, L. Puigjaner, and M.C. Georgiadis. Simultaneous production and logistics operations planning in semicontinuous food industries. *Omega*, 40(5):634–650, 2012.

[11] A. Federgruen and P.H. Zipkin. A combined vehicle routing and inventory allocation problem. *Operations Research*, 32:1019–1037, 1984.

[12] B.L. Golden, A. Assad, and R. Dahl. Analysis of a large scale vehicle routing problem with an inventory component. *Large Scale Systems*, 7:181–190, 1984.

[13] M. Dror, M. Ball, and B.L. Golden. A computational comparison of algorithms for the inventory routing problem. *Annals of Operations Research*, 4:3–23, 1985.

[14] A. Federgruen, G. Prastacos, and P.H. Zipkin. An allocation and distribution model for perishable products. *Annals of Operations Research*, 34:75–82, 1986.

[15] Y. Bassok and R. Ernst. Dynamic allocations for multi-product distribution. *Annals of Operations Research*, 29:256–266, 1995.

[16] Y.T. Herer and R. Levy. The metered inventory routing problem, an integrative heuristic algorithm. *International Journal of Production Economics*, 51:69–81, 1997.

[17] J.F. Bard, L. Huang, P. Jaillet, and M. Dror. A decomposition approach to the inventory routing problem with satellite facilities. *Transportation Science*, 32:189–203, 1998.

[18] O. Berman and R.C. Larson. Deliveries in an inventory/routing problem using stochastic dynamic programming. *Transportation Science*, 35:192–213, 2001.

[19] W.J. Bell, L.M. Dalberto, M.L. Fisher, A.J. Greenfield, R. Jaikumar, P. Kedia, R.G. Mack, and P.J. Pruzman. Improving the distribution of industrial gases with an on-line computerized routing and scheduling optimizer. *Interfaces*, 13:4–23, 1983.

[20] V. Gaur and M.L. Fisher. A periodic inventory routing problem at a supermarket chain. *Operations Research*, 52:813–822, 2004.

[21] S. Dauzère-Pérès, A. Nordli, A. Olstad, K. Haugen, U. Koester, P.O. Myrstad, and al. Omya Hustadmarmor optimizes its supply chain for delivering calcium carbonate slurry to european paper manufacturers. *Interfaces*, 37:39–51, 2007.

[22] A.S. Minkoff. A markov decision model and decomposition heuristic for a dynamic vehicle dispatching. *Operations Research*, 41:77–90, 1993.

[23] A. Kleywegt, V. Nori, and M. Savelsbergh. The stochastic inventory routing problem with direct deliveries. *Transportation Science*, 36:94–118, 2002.

[24] A. Kleywegt, V. Nori, and M. Savelsbergh. Dynamic programming approximations for a stochastic inventory routing problem. *Transportation Science*, 38:42–70, 2004.

[25] D. Adelman. Price-directed replenishment of subsets: methodology and its application to inventory routing. *Manufacturing & Service Operations Management*, 5(4):348–371, 2003.

[26] D. Adelman. A price-directed approach to stochastic inventory/routing. *Operations Research*, 52(4):499–514, 2004.

[27] L. M. Hvattum and A. Løkketangen. Using scenario trees and progressive hedging for stochastic inventory routing problems. *Journal of Heuristics*, 15:527–557, 2009.

[28] L. M. Hvattum, G. Laporte, and A. Løkketangen. Scenario tree-based heuristics for stochastic inventory-routing problems. *INFORMS Journal on Computing*, 21(2):268–285, 2009.

[29] Y. Yu, C. Chu, H. Chen, and F. Chu. Large scale stochastic inventory routing problems with split delivery and service level constraints. *Annals of Operational Research*, 197:135–158, 2012.

[30] L. Bertazzi, A. Bosco, F. Guerriero, and D. Laganà. A stochastic inventory routing problem with stock-out. *Transportation Research Part C: Emerging Technologies*, 27:89–107, 2013.

[31] A. Aguezzoul. Third-party logistics selection problem: A literature review on criteria and methods. *Omega*, 49:69–78, 2014.

[32] C.W. Chu. A heuristic algorithm for the truckload and less-than-truckload problem. *European Journal of Operational Research*, 165:657–667, 2005.

[33] G. Clarke and J.W. Wright. Scheduling of vehicles from a central depot to a number of delivery points. *Operations Research*, 12(4):568–581, 1964.

[34] M.C. Bolduc, J. Renaud, F. Boctor, and G. Laporte. A perturbation metaheuristic for the vehicle routing problem with private fleet and common carriers. *International Journal of the Operational Research Society*, 59(6):776–787, 2008.

[35] J.F. Côté and J.Y. Potvin. A tabu search heuristic for the vehicle routing problem with private fleet and common carrier. *European Journal of Operational Research*, 198(2):464–469, 2009.

[36] J.Y. Potvin and M.A. Naud. Tabu search with ejection chains for the vehicle routing problem with private fleet and common carrier. *International Journal of the Operational Research Society*, 62:326–336, 2011.

[37] A. Stenger, D. Vigo, S. Enz, and M. Schwind. An adaptive variable neighborhood search algorithm for a vehicle routing problem arising in small package shipping. *Transportation Science*, 47(1):64–80, 2013.

[38] L.C. Coelho, J.F. Cordeau, and G. Laporte. The inventory-routing problem with transshipment. *Computers & Operations Research*, 39(11):2537–2548, 2012.

[39] F. Pan and R. Nagi. Multi-echelon supply chain network design in agile manufacturing. *Omega*, 41(6):969–983, 2013.

[40] R. Mansini, M.W.P. Savelsbergh, and B. Tocchella. The supplier selection problem with quantity discounts and truckload shipping. *Omega*, 40(4):445–455, 2012.

[41] Y. Sheffi. Combinatorial auctions in the procurement of transportation services. *Interfaces*, 34(4):245–252, 2004.

[42] S. Martello and P. Toth. *Knapsack problems: algorithms and computer implementations*. John Wiley & Sons, Inc., 1990.

[43] J.R. Birge and F. Louveaux. *Introduction to stochastic programming*. Springer Verlag, 1997.

[44] D.P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, 2005.

[45] W.B. Powell. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. John Wiley & Sons, Inc., 2011.

[46] V. Maniezzo, T. Stèutzle, and S. Voss. *Matheuristics: hybridizing metaheuristics and mathematical programming*, volume 10. Springer, 2009.

[47] L. Bertazzi and M.G. Speranza. *Matheuristics for Inventory Routing Problems in Hybrid Algorithms for Service, Computing and Manufacturing Systems: Routing, Scheduling and Availability Solutions*. J.R. Montoya-Torres and A.A. Juan and L.H. Huatuco and J. Faulin and G.L. Rodriguez-Verjan, editors, IGI Global, 2011.

[48] L. Bertazzi. Minimum and worst-case performance ratios of rollout algorithms. *Journal of Optimization Theory and Applications*, 152:378–393, 2012.

[49] D.P. Bertsekas and J.N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, 1996.

[50] G. Tesauro and G.R. Galperin. On-line policy improvement using monte-carlo search. *Advances in Neural Information Processing Systems*, 9:1068–1074, 1997.

[51] R.S. Sutton and A.G. Barto. *Reinforcement learning: An introduction*, volume 1. Cambridge Univ Press, 1998.

[52] D.P. Bertsekas and D.A. Castanõn. Rollout algorithms for stochastic scheduling problems. *Journal of Heuristics*, 5(1):89–108, 1999.

[53] N. Secomandi. Comparing neuro-dynamic programming algorithms for the vehicle routing problem with stochastic demands. *Computers & Operations Research*, 27(11-12):1201–1225, 2000.

[54] N. Secomandi. A rollout policy for the vehicle routing problem with stochastic demands. *Operations Research*, 49:796–802, 2001.

[55] N. Secomandi. Analysis of a rollout approach to sequencing problems with stochastic routing applications. *Journal of Heuristics*, 9(4):321–352, 2003.

[56] C. Novoa and R. Storer. An approximate dynamic programming approach for the vehicle routing problem with stochastic demands. *European Journal of Operational Research*, 196(2):509–515, 2009.

[57] F. Guerriero and M. Mancini. A cooperative parallel rollout algorithm for the sequential ordering problem. *Parallel Computing*, 29(5):663–677, 2003.

[58] F. Guerriero and M. Mancini. Parallelization strategies for rollout algorithms. *Computational Optimization and Applications*, 31(2):221–244, 2005.

[59] D.P. Bertsekas, J.N. Tsitsiklis, and C. Wu. Rollout algorithms for combinatorial optimization. *Journal of Heuristics*, 3(3):245–262, 1997.

[60] F. Guerriero. Hybrid rollout approaches for the job shop scheduling problem. *Journal of Optimization Theory and Applications*, 139(2):419–438, 2008.

[61] D. Bertsimas and R. Demir. An approximate dynamic programming approach to multidimensional knapsack problems. *Management Science*, 48:550–565, 2002.

[62] C. Archetti, L. Bertazzi, G. Laporte, and M.G. Speranza. A branch-and-cut algorithm for a vendor-managed inventory-routing problem. *Transportation Science*, 41(3):382–391, 2007.