FINAL REPORT

To: The Tennessee Department of Transportation
Research Development and Technology Program

Project #:
Truck Congestion Mitigation through Freight Consolidation in Volatile Multi-item Supply Chains

Prepared by: [a]Dinçer Konur and [b]Mihalis M. Golias

[a]Department of Engineering Management and Systems Engineering
Missouri University of Science and Technology
206 Engineering Management, 600 W. 14th Street, Rolla, MO 65409
Office: 573.341.7256, Fax: 573.341.6990, Email: konurd@mst.edu

[b]Department of Civil Engineering, University of Memphis
104 Engineering Science Bldg, 3815 Central Avenue, Memphis, TN 38152
Office: 901.678.3048, Fax: 901.678.3026, Email: mgkolias@memphis.edu

March, 2014

Contents

List of Figures

List of Tables

INTRODUCTION

Transportation costs constitute the major part of the total costs in many retailing industries. Therefore, there has been tremendous amount of studies in the literature, which integrate inventory control and transportation decisions as inventory control policies determine how much and how often to ship. There is an obvious trade-off between inventory holding costs and order setup plus transportation costs in many practical scenarios: while replenishing inventory more frequently with smaller orders decreases inventory holding costs, it results in more order setup payments and frequent shipments from the suppliers; hence, increases order setup and transportation costs. Considering the non-linear nature of inventory related costs, jointly controlled inventories of multiple items, and demand uncertainties, integrated inventory control and transportation problems can be challenging.

This study focuses on a retailer's integrated inventory control and transportation problem for multiple items, each of which has its own stochastic demand. We explicitly model inbound transportation costs by taking into account that a retailer can use different freight trucks to ship an order. Furthermore, to utilize transportation capacity better, the retailer can possibly consolidate shipments of different items. To avail consolidation, it is assumed that a retailer adopts a time-based order-up-to-level inventory control policy, where the retailer replenishes each consolidated set of items in equal time intervals (this enables joint use of the transportation capacity by the consolidated items). The retailer's problem is to find the cost minimizing consolidation strategy, i.e., of which items' orders are replenished together, the time interval between two consecutive orders of a set of consolidated items, and the order-up-to-level for each item within a consolidation.

Due to stochastic demand environment, the retailer's objective is to minimize the expected costs. While the expected inventory holding costs, order setup costs, and penalty costs associated with shortages are well defined, the derivation of the expected inbound transportation costs is cumbersome due to the fact that freight truck choices for each order of a set of consolidated items are dynamic in nature. That is, the retailer can determine how many trucks of each truck type to be used for each order at order initiation depending on the order quantities of the individual items in the consolidation. This, in turn, makes the retailer's problem of expected cost minimization a bi-level optimization model with infinitely many lower level problems (each one is corresponding to a combination of the demands of the items within a given consolidation).

In this study, we first formulate the retailer's problem for a given consolidation of items. Here, a bi-level mixed-integer nonlinear optimization problem is modeled, where the retailer decides on the common replenishment cycle length for the consolidated items and the order-up-to-level for each item within the given consolidation. Then, a set partitioning problem is presented to find the best consolidation strategy. As a solution approach, we first provide an approximation formulation for a given consolidation and solve the approximated formulation with a neighborhood search heuristic. Then, an evolutionary heuristic method is discussed for the set partitioning problem of interest. A set of numerical studies are conducted to justify the approximation formulation and use of heuristic methods. Furthermore, through a set of numerical studies, we demonstrate the cost savings and environmental benefits of the proposed time-based order-up-to-level inventory control with shipment consolidation and explicit freight trucks modeling in multi-item stochastic inventory systems.

This study contributes to the inventory control literature and practice in the following fields: explicit transportation modeling, shipment consolidation, and stochastic joint replenishment problem. In the remainder of this section, we review the related literature on each of these fields and explain the similarities and differences along with contributions of our study in the respective field. In Section , the mathematical models for single item, a consolidation, and set partitioning problem are formulated. Section explains the details of the methods used to solve the resulting problem. Section documents the results of a set of numerical studies that illustrate the efficiency of the heuristic methods and demonstrate the potential cost and environmental benefits of the proposed consolidation strategy. Section summarizes the results and contributions of this study and suggests future research directions.

Explicit Transportation Modeling

It is well known that freight trucks is the most common transportation mode. Majority of the freight tonnage is shipped by trucks in the U.S. (FHWA, 2012). Two common practices of the trucking are less-than-truckload (LTL) and truckload (TL) transportation. Most of the integrated inventory control and transportation studies assume LTL transportation, where the shipment cost depends on the number (weight or volume) of items shipped. On the other hand, in TL transportation, the shipment cost depends on the number of trucks used. In this study, we assume TL transportation with the availability of heterogeneous freight trucks for inbound shipment.

Particularly, TL transportation with single truck type has been studied within inventory control models. Aucamp (1982), Lee (1986), Toptal et al. (2003), Toptal and Çetinkaya (2006), Toptal (2009), Toptal and Bingol (2011), and Konur and Toptal (2012) are some of the studies that account for TL transportation costs explicitly in single-item inventory control models. In particular, similar to these studies, TL transportation costs are modeled considering the per truck capacities and per truck costs in this study. In multi-item inventory settings, there is limited number of studies assuming TL transportation. Ben-Khedher and Yano (1994) analyze a multi-item deterministic joint replenishment problem with trucking costs as well as capacity constraints. They propose a heuristic method to solve the resulting NP-hard problem. In a similar setting, Kiesmuller (2009) analyzes a multi-item stochastic inventory system with periodic review and they account for TL transportation costs. Specifically, they propose a period order-up-to inventory policy where the trucks used for shipment have to be fully loaded; nevertheless, it is noted that full truckloads policy can be suboptimal for a retailer as it might lead to increased holding costs at such levels that decrease in shipping costs cannot counter balance it. A similar observation has been made by Toptal et al. (2003) in a single-item model; they note that it might be beneficial to have one of the trucks to be partially loaded.

In the aforementioned studies, only a single truck type is considered. We further generalize TL transportation modeling by taking different freight trucks into consideration. In case a retailer uses 2PL or 3PL for inbound transportation, there might be different TL carriers available, each of which has distinct truck fleets. Even in the case of a single TL carrier, it might be the case that the retailer can be forced to select among a set of different freight trucks for his/her inbound transportation. In such a case, the retailer needs to dynamically determine how many trucks of each truck type to use for the inbound shipment of each order. This study contributes to the multi-item inventory control models by providing generalized formulation for TL transportation with heterogeneous freight trucks. Specifically, we consider different per truck capacities and per truck costs for distinct truck types available for inbound shipment. Furthermore, the aforementioned studies define truck capacity in terms of the number of items that can be carried. We extend truck capacity definition by jointly regarding the weight and volume capacities for different truck types.

Shipment Consolidation

As mentioned previously, transportation costs constitute a significant part of total costs in many industries; therefore, utilization of transportation capacity can substantially save costs. The practice of shipment consolidation targets better utilization of the transportation capacity by merging shipments of small quantities to achieve a shipment with larger quantity that utilizes the transportation capacity better. This, in turn, reduces costs due to economies of scale in the transportation costs (Mutlu et al., 2010).

Three common shipment consolidation policies considered as quantity-based, time-based, and time-and-quantity-based consolidation (Cetinkaya et al., 2006). In the quantity-based shipment consolidation, the customer demands are accumulated until a specified quantity is achieved; and, then a shipment is released. On the other hand, in the time-based shipment consolidation, the customer demands are accumulated for a specified time period; and, then a shipment is released. In the time-and-quantity-based consolidation, the customer demands are accumulated until a specified quantity is achieved or a specified time period is ended; and, then a shipment is released. Çetinkaya (2005) provides a detailed review of coordinated inventory control models with shipment consolidation. In this study, we assume a time-based shipment consolidation policy, that is, an order is placed in equal time intervals. However, we note that we also formulate the decisions on which items to be consolidate.

Stochastic Joint Replenishment Problem

The joint replenishment problem (JRP) considers how to jointly replenish a set of different products in a multi-item inventory system. The main motivation for jointly replenishing the different products is the economies of scale of the order setup costs. Generally, order setup costs are defined by the transportation costs of a shipment. The reader is referred to a review of JRPs by Khouja and Goyal (2008) for different settings, models, and solution approaches studied in the literature for JRPs. In stochastic JRPs, each product has its own stochastic demand.

Balintfy (1964) proposes a can-order policy for a stochastic JRP, where each item has a must-order level s, a can-order level c, and an order-up-to-level S. In a can-order policy, denoted by (s, c, S), an item is ordered when its inventory level reaches the must-order level, and any other item, whose inventory level is below the can-order level, is then ordered with it such that the order quantities for the ordered items build their inventory levels to the specified order-up-to-levels. While Balintfy (1964) assumes continuous inventory review, Johansen and Melchiors (2003) analyze the can-order policy under periodic review noting that replenishment opportunities may only come once or twice a day and; therefore, a periodic review model can be superior for some customers.

Atkins and Iyogun (1988) analyze JRP strategies where the items are ordered up to an order-up-to-level R every time period of length T. These policies are referred to as (R, T) polices and Atkins and Iyogun (1988) investigate two (R, T) polices: a periodic policy, where all items are ordered with each replenishment and a modified periodic policy, where a base set of items is ordered with each replenishment and the remaining items are ordered at each specified consecutive replenishment. In this study, we adopt a (R, T) type of policy for a given set of consolidated items: the inventories of the items in the consolidation are replenished every T time units up to their individual order-up-to-levels. Atkins and Iyogun (1988) concludes that the periodic (R, T) type policies show more promise than the (s, c, S) type policies. However, Pantumsinchai (1992) notes that different policies can be superior to the others depending on the specific problem parameters.

Viswanathan (1997) introduces a new class of policies known as the $P(s, S)$ policy. The $P(s, S)$ policy is a periodic review policy where the amount of items on hand are reviewed at intervals of time

T. If the amount of items on hand is less than s then items are ordered to bring the inventory up to S. They test their algorithm against the same problems in Atkins and Iyogun (1988) and find that their proposed policy generally gives dominating solutions and that the extra computational requirement is nominal. Nielsen and Larsen (2005) use Markov decision theory and find an analytical solution to the $Q(s, S)$ policy, which was listed as a future research direction by Viswanathan (1997). In the $Q(s, S)$ policy, the total number of items are reviewed continuously but the items themselves are only reviewed once the total demand reaches Q. Nielsen and Larsen (2005) find the $Q(s, S)$ model to be superior to the periodic review $P(s, S)$ models. Ozkaya et al. (2006) propose a new hybrid (Q, S, T) policy. The policy is considered to be both continuous and periodic as orders are placed to the order-up-to level S whenever total demand level Q is reached or time T has elapsed since the last order. Using the same problem settings with Atkins and Iyogun (1988) and Viswanathan (1997) as a benchmark, Ozkaya et al. (2006) find their proposed method to be better 72% of the time.

All of the above models are unconstrained and Zhao et al. (2012) state that "Inventory systems with limited and sharable-common resource exist widely in the real logistics field, yet studies on such systems are limited." Minner and Silver (2005) develop a multi-product inventory replenishment problem where the inventory level at any time is constrained by a budget or space limitations. They assume a Poisson demand, zero lead time, and no backorders and formulate the problem as a semi-Markov decision process. Zhao et al. (2012) also study a constrained policy, specifically, the (r, Q) policy with a limited sharable common resource. In the (r, Q) policy, when an item's inventory drops below r then Q units of that item are ordered. Betts and Johnston (2005) study a similar model with a constraint on the investment capital available. In this study, the resource commonly shared is the transportation capacity, which is also a decision variable of the retailer at each replenishment.

PROBLEM FORMULATION

Consider a set of n items indexed by i, $i \in I$, where $I = \{1, 2, \dots, n\}$, such that each item has a stochastic demand. Let $f^i(D_i)$ and $F^i(D_i)$ denote the probability density function and cumulative distribution function of item i's demand, D_i, over unit time. We assume that the unit time demand for any item i is normally distributed with mean λ_i and standard deviation σ_i. Thus, item i's demand over a period of t time units is normally distributed with mean $\lambda_i t$ and standard deviation $\sigma_i \sqrt{t}$ (see, e.g., Nahmias, 2009). We denote $f_i(D_i^{(t)})$ as the probability density function of item i's demand over a period of t time units, where $D_i^{(t)}$ is the random variable defining item i's demand over t time units*.

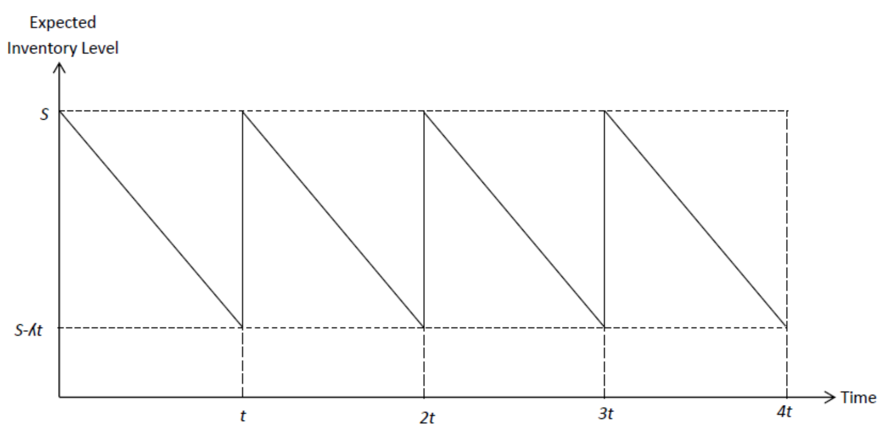
Under the current settings, the retailer is subject to inventory holding, order setup, and shortage

*The problem formulation and the solution methods presented can be easily modified for other demand distributions.

costs. In particular, let h_i denote the inventory holding cost per unit per unit time, a_i denote the order setup cost per each order, and p_i denote the penalty cost per unit shortage for item i. In addition to these costs, the retailer is subject to explicit transportation costs associated with each order. We assume that the retailer can use m different truck types for inbound shipment. Let different truck types be indexed by j, $j \in J$, where $J = \{1, 2, \ldots, m\}$ such that a single truck of type j has weight-capacity of W_j, volume-capacity of V_j, and cost of R_j. Furthermore, let each unit of item i have weight w_i and volume v_i.

The retailer is assumed to adopt a time-based order-up-to-level inventory control policy. That is, for a single item or a set of consolidated items, the retailer will place an order at identical time intervals such that each item's order quantity is determined to increase the inventory level of that item to a specific point. We assume that delivery lead time is negligible[†]. If the retailer plans to manage item i individually, his/her decision variables would be order-up-to-level for item i, denoted by s_i, and the replenishment cycle length t_i. Figure 1 illustrates the expected inventory level over time for a single item with replenishment cycle length t, order-up-to-level s, and λ demand per unit time.

Figure 1: Inventory Level in Time-based Order-Up-To-Level Control for Single Item



In the remainder of this section, we first formulate the case of a single-item being replenished individually. Then, a mathematical formulation is given to determine the common replenishment cycle length for a set of consolidated items along with each item's individual order-up-to-level decisions in the consolidated set. Following this, the model to decide on which items to consolidate is presented.

[†]We note that the problem formulation provided can be modified to handle constant lead times. Specifically, once the time interval for consecutive orders is determine, a retailer can initiate the order accordingly regarding the delivery lead time

Single-Item Time-based Order-up-to-level Inventory

Consider that item i is individually replenished. As noted previously, the retailer is subject to inventory holding, order setup, shortage, and inbound transportation costs. Due to stochastic demand, the retailer's objective is to minimize the total expected costs per unit time associated with item i. Expected inventory holding cost per unit time amounts to $h_i(s_i - \frac{\lambda_i t_i}{2})$. Order setup cost per unit time is a deterministic variable depending on t_i and it amounts to $\frac{a_i}{t_i}$. Now, let $n_i(s_i, t_i)$ be the expected number of shortages within one replenishment cycle as function of s_i and t_i. Then, expected shortage cost per unit time amounts to $\frac{p_i n_i(s_i,t_i)}{t_i}$. Note that the number of shortages within a replenishment cycle depends on both the replenishment cycle length t_i and the order-up-to-level s_i; hence, $n_i(s_i, t_i)$ is a function of s_i and t_i. One can show that $n_i(s_i, t_i) = \int_{s_i}^{\infty} \left(D_i^{(t_i)} - s_i\right) f_i(D_i^{(t_i)}) dD_i^{(t_i)}$. Therefore, expected shortage per unit time is $\frac{p_i}{t_i} \int_{s_i}^{\infty} \left(D_i^{(t_i)} - s_i\right) f_i(D_i^{(t_i)}) dD_i^{(t_i)}$.

The only remaining cost term is the expected inbound transportation costs. Recall that the retailer can use m different truck types for inbound transportation. At each order replenishment, the retailer needs to decide on how many of each truck type should be used. Let x_j be the integer number of type j trucks to be used for inbound transportation of an order and $\mathbf{x} = [x_1, x_2, \ldots, x_m]$. The order quantity to be shipped will be equal to the demand realized during the replenishment cycle, i.e., $D_i^{(t_i)}$. In this case, the retailer will determine the truck configuration $\mathbf{x}$ that will minimize inbound transportation costs to ship $D_i^{(t_i)}$ units. Therefore, the following problem needs to be solved at each replenishment:

$$\begin{array}{rcll} ITC_i(D_i^{(t_i)}) & = & \min_{\mathbf{x}} & \sum_{j \in J} x_j R_j \\ & & \text{s.t.} & \sum_{j \in J} x_j W_j \geq w_i D_i^{(t_i)} \\ & & & \sum_{j \in J} x_j V_j \geq v_i D_i^{(t_i)} \\ & & & x_j \in \{0, 1, 2, \ldots\} \ \forall j \in J. \end{array} \tag{1}$$

The objective function in the definition of $ITC_i(D_i^{(t_i)})$ given in Eq. (1) is the total trucking cost. The first and second constraints assure that the selected trucks cumulatively have the sufficient weight and volume capacity to ship $D_i^{(t_i)}$ units, respectively. The third set of constraints is the integer definition for the x_j values. (Note that if $D_i^{(t_i)} \leq 0$, $x_j = 0 \ \forall j = 1, 2, \ldots, m$; hence, $ITC_i(D_i^{(t_i)}) = 0$ for $D_i^{(t_i)} \leq 0$.) Then, expected inbound transportation cost per unit time amounts to $\frac{1}{t_i} \int_0^{\infty} ITC_i(D_i^{(t_i)}) f_i(D_i^{(t_i)}) dD_i^{(t_i)}$.

The retailer's total expected costs per unit time when item i is individually replenished, denoted by $g_i(s_i, t_i)$, amount to

$$\begin{array}{rcl} g_i(s_i, t_i) & = & h_i\left(s_i - \dfrac{\lambda_i t_i}{2}\right) + \dfrac{a_i}{t_i} + \dfrac{p_i}{t_i} \displaystyle\int_{s_i}^{\infty} \left(D_i^{(t_i)} - s_i\right) f_i(D_i^{(t_i)}) dD_i^{(t_i)} \\ & & + \dfrac{1}{t_i} \displaystyle\int_0^{\infty} ITC_i(D_i^{(t_i)}) f_i(D_i^{(t_i)}) dD_i^{(t_i)} \end{array} \tag{2}$$

where the first, second, third, and forth terms of Eq. (2) are the expected inventory holding, order setup, shortage, and inbound transportation costs per unit time. The retailer's optimization problem for individually replenished item i then reads as

$$
\begin{aligned}
(\mathbf{P}^i) \quad & \min_{(s_i,t_i)} \quad g_i(s_i,t_i) \\
& \text{s.t.} \quad t_i \geq 0 \\
& \qquad\quad s_i \geq 0 \\
& \qquad\quad ITC_i(D_i^{(t_i)}) = \min_{\mathbf{x}} \quad \sum_{j\in J} x_j R_j \\
& \qquad\qquad\qquad\qquad\quad \text{s.t.} \quad \sum_{j\in J} x_j W_j \geq w_i D_i^{(t_i)} \\
& \qquad\qquad\qquad\qquad\qquad\quad\; \sum_{j\in J} x_j V_j \geq v_i D_i^{(t_i)} \\
& \qquad\qquad\qquad\qquad\qquad\quad\; x_j \in \{0,1,2,\ldots\} \qquad \forall j \in J.
\end{aligned}
$$

Consolidated Time-based Order-up-to-level Inventory

Now suppose that a set of items are ordered together, that is, their shipments are consolidated. The retailer's objective is to determine the order-up-to-level for each item in the consolidation and the replenishment cycle length for the consolidation so that the total expected costs per unit time for the items in the consolidation are minimized. Any subset of the set of items I is a possible consolidation; thus, there are $2^n - 1$ subsets of items that can be consolidated. Let each possible subset of items be indexed by k, $k \in K$ where $K = \{1, 2, \ldots, 2^n - 1\}$ and Ω_k denote a subset. Furthermore, let T_k denote the common replenishment cycle when Ω_k is selected as a consolidation, i.e., $t_i = T_k \; \forall i \in \Omega_k$.

Similar to single-item case, a consolidated set of items has inventory holding, order setup, shortage, and inbound transportation costs. Note that inventory holding, order setup, and shortage costs of the items in a consolidation are individual cost terms; therefore, total expected holding, order setup, and shortage costs per unit time for the consolidation will be equal to the sum of the expected holding, order setup, and shortage cost per unit time of each item in the consolidation. That is, the total expected holding cost per unit time of consolidation Ω_k $k \in K$ is equal to the sum of the expected holding costs per unit time of the consolidated items. The total expected holding cost per unit time of the consolidation is, therefore, equal to $\sum_{i\in\Omega_k} h_i s_i - \frac{T_k}{2}\sum_{i\in\Omega_k} h_i\lambda_i$. Similarly, it follows that the total order setup cost per unit time for Ω_k amounts to $\frac{1}{T_k}\sum_{i\in\Omega_k} a_i$, and the total shortage cost per unit time for Ω_k is equal to $\frac{1}{T_k}\sum_{i\in\Omega_k} p_i \left(\int_{s_i}^{\infty}\left(D_i^{(T_k)} - s_i\right) f_i(D_i^{(T_k)}) dD_i^{(T_k)}\right)$.

Unlike the inventory holding, order setup, and shortage costs for Ω_k, the inbound transportation costs will not be equal to the sum of the individual items' transportation costs as different items can share truck capacities due to being replenished simultaneously. In particular, at each replenishment, the retailer needs to decide on the number of trucks of each type to ship the realized demands of the items in the consolidation. Let $\mathbf{D}_{\Omega_k}^{(T_k)}$ be the $|\Omega_k|$-vector of $D_i^{(T_k)}$ values for $i \in \Omega_k$. The following problem then should be solved at each replenishment to determine the inbound transportation cost of consolidation Ω_k:

$$\begin{array}{rcll} ITC_{\Omega_k}(\mathbf{D}_{\Omega_k}^{(T_k)}) & = & \min_{\mathbf{x}} & \sum_{j\in J} x_j R_j \\ & & \text{s.t.} & \sum_{j\in J} x_j W_j \geq \sum_{i\in\Omega_k} w_i D_i^{(T_k)} \\ & & & \sum_{j\in J} x_j V_j \geq \sum_{i\in\Omega_k} v_i D_i^{(T_k)} \\ & & & x_j \in \{0,1,2,\ldots\}\ \forall j \in J. \end{array} \quad (3)$$

Similar to Eq. (1), the objective function in the definition of $ITC_{\Omega_k}(\mathbf{D}_{\Omega_k}^{(T_k)})$ given in Eq. (3) is the total trucking cost. The first and second constraints guarantee that the selected trucks cumulatively have the sufficient weight and volume capacity to ship $D_i^{(t_i)}$ $\forall i \in \Omega_k$, respectively. The third set of constraints is the integer definition for the x_j values. Now, let us assume that $\Omega_k = \{1, 2, \ldots, \ell\}$ such that $\ell \leq n$. Then, expected inbound transportation cost per unit time amounts to $\frac{1}{T_k}\int_0^\infty ITC_{\Omega_k}(\mathbf{D}_{\Omega_k}^{(T_k)}) f(\mathbf{D}_{\Omega_k}^{(T_k)}) d\mathbf{D}_{\Omega_k}^{(T_k)} = \frac{1}{T_k}\int_0^\infty \int_0^\infty \ldots \int_0^\infty ITC_{\Omega_k}(\mathbf{D}_{\Omega_k}^{(T_k)}) f_1(D_1^{(T_k)}) f_2(D_2^{(T_k)}) \ldots f_\ell(D_\ell^{(T_k)}) dD_1^{(T_k)} dD_2^{(T_k)} \ldots dD_\ell^{(T_k)}$.

The retailer's total expected costs per unit time when items in Ω_k are consolidated, denoted by $G_k(\mathbf{S}_k, T_k)$, amount to

$$\begin{aligned} G_k(\mathbf{S}_k, T_k) = & \sum_{i\in\Omega_k} h_i s_i - \frac{T_k}{2}\sum_{i\in\Omega_k} h_i\lambda_i + \frac{1}{T_k}\sum_{i\in\Omega_k} a_i + \frac{1}{T_k}\sum_{i\in\Omega_k} p_i \left(\int_{s_i}^\infty \left(D_i^{(T_k)} - s_i\right) f_i(D_i^{(T_k)}) dD_i^{(T_k)}\right) \\ & + \frac{1}{T_k}\int_0^\infty ITC_{\Omega_k}(\mathbf{D}_{\Omega_k}^{(T_k)}) f(\mathbf{D}_{\Omega_k}^{(T_k)}) d\mathbf{D}_{\Omega_k}^{(T_k)} \end{aligned} \quad (4)$$

where $\mathbf{S}_k$ is a $|\Omega_k|$-vector of s_i values for $\forall i \in \Omega_k$. The first, second, third, and forth terms of Eq. (4) are the expected inventory holding, order setup, shortage, and inbound transportation costs per unit time for the consolidation Ω_k. The retailer's optimization problem for consolidation Ω_k then reads as

$$\begin{array}{lll} (\mathbf{P}^{\Omega_k}) & \min\limits_{(\mathbf{S}_k, T_k)} & G_k(\mathbf{S}_k, T_k) \\ & \text{s.t.} & T_k \geq 0 \\ & & s_i \geq 0 \\ & & ITC_{\Omega_k}(\mathbf{D}_{\Omega_k}^{(T_k)}) = \min\limits_{\mathbf{x}} \sum\limits_{j\in J} x_j R_j \\ & & \qquad \text{s.t.} \sum\limits_{j\in J} x_j W_j \geq \sum\limits_{i\in\Omega_k} w_i D_i^{(T_k)} \\ & & \qquad \sum\limits_{j\in J} x_j V_j \geq \sum\limits_{i\in\Omega_k} v_i D_i^{(T_k)} \\ & & \qquad x_j \in \{0,1,2,\ldots\} \quad \forall j \in J. \end{array}$$

Let $\mathbf{S}_k^*$ and T_k^* denote an optimum solution of $(\mathbf{P}^{\Omega_k})$.

Consolidation Decisions

Ultimately, the retailer's goal is to determine which items will be consolidated and what will be the common replenishment cycle length for each consolidation and order-up-to-level for each set of items in the consolidations. Therefore, the retailer needs to select which subsets of items will be consolidation such that each item will be replenished within a single consolidation. A given consolidation Ω_k can be defined by c_{ik} values such that

$$c_{ik} = \begin{cases} 1 & \text{if item } i \text{ is in consolidation } \Omega_k, \\ 0 & \text{otherwise.} \end{cases}$$

Let

$$y_k = \begin{cases} 1 & \text{if consolidation } \Omega_k \text{ is selected,} \\ 0 & \text{otherwise.} \end{cases}$$

Assuming that the retailer will adopt the optimum common replenishment cycle length and order-up-to-levels for any consolidation Ω_k, i.e., $\mathbf{S}_k^*$ and T_k^*, the retailer's consolidation problem reads as

$$\begin{aligned} (\mathbf{P}) \quad \min_{(\mathbf{y})} \quad & C(\mathbf{y}) = \sum_{k \in K} y_k G_k(\mathbf{S}_k^*, T_k^*) \\ \text{s.t.} \quad & \sum_{k \in K} c_{ik} y_k = 1 && \forall i \in I \\ & y_k \in \{0, 1\} && \forall k \in K. \end{aligned}$$

where $\mathbf{y}$ is the binary $(2^n - 1)$-vector of y_k values. The objective function of $(\mathbf{P})$ minimizes the total expected costs per unit time. The first set of constraints ensures that each item is included within one of the selected consolidations. The second set of constraints are the binary definitions for the decision variables. We note that $(\mathbf{P})$ is a set partitioning problem, which is known to be NP-hard (see, e.g., Garey and Johnson, 1979). Furthermore, definitions of $\mathbf{S}_k^*$ and T_k^* require to solve bi-level mixed-integer non-linear optimization problems. Therefore, in the next section, we focus on an evolutionary heuristic method to solve $(\mathbf{P})$ and discuss an approximation formulation that reduces $(\mathbf{P}^{\Omega_k})$ to single-level mixed-integer non-linear optimization problems, for which we discuss another heuristic method.

Solution Analysis

In this section, we propose a genetic algorithm based meta-heuristic approach for solving problem $(\mathbf{P})$, denoted by GA-P. GA-P has the following four main steps: (i) chromosome representation and initialization, (ii) fitness evaluation, (iii) mutation, and (iv) termination. In what follows, we discuss the details of each step.

Chromosome Representation and Initialization

Note that the retailer can select at most n consolidations (when each item is individually replenished), that is, $\sum_{k \in K} y_k \leq n$. Therefore, a solution to $(\mathbf{P})$ can be presented by an integer n-vector $chrom = [c_1, c_2, \ldots, c_n]$, where c_i denotes the consolidation number that item i belongs to. Note that one should have $1 \leq c_i \leq n \; \forall i \in I$. The important point about defining a solution for $(\mathbf{P})$ as a *chrom* vector is that the corresponding consolidation decisions are feasible for $(\mathbf{P})$ as each item is guaranteed to be within one consolidation. For instance, for a problem instance with $n = 5$ items, let $chrom = [3, 1, 2, 3, 2]$; then, items 1 and 4 form one consolidation, items 3 and 5 form one consolidation, and item 2 forms one consolidation. That is, $\{1, 4\}, \{3, 5\}, \{2\}$ are the three consolidations selected. Furthermore, *chrom* representation enables mutation operations to be simply executed. As an initialization, we randomly generate nm number of *chrom* vectors by randomly generating c_i values such that $1 \leq c_i \leq n \ \forall i \in I$.

Fitness Evaluation

Now suppose that a set of chromosomes are given. For each chromosome, one can determine the number of consolidations and the items in each consolidation as explained above. The fitness value for a chromosome is the total expected costs of the consolidations in the chromosome. Therefore, one needs to find the total expected costs per unit time for each consolidation of a given chromosome and get their summation to find the fitness value of the chromosome. To do so, problem $(\mathbf{P}^{\Omega_k})$ should be solved for each consolidation associated with the chromosome. Note that $(\mathbf{P}^{\Omega_k})$ is a bi-level mixed-integer non-linear optimization problem due to the calculation of expected inbound transportation costs present in the objective function, i.e., Eq. (3). Even the simplest bi-level optimization problems, when optimization problems at both levels are linear, are shown to be NP-hard (see, e.g., Hansen et al., 1992). Furthermore, one needs to solve $(\mathbf{P}^{\Omega_k})$ at least once and at most n times for each chromosome to be evaluated. Therefore, an efficient method to solve $(\mathbf{P}^{\Omega_k})$ is required. In what follows, we first discuss an approximated reformulation for $(\mathbf{P}^{\Omega_k})$, which gives a single-level mixed-integer nonlinear optimization problem; then, we discuss a local search algorithm to solve the resulting single-level mixed-integer nonlinear optimization problem.

Approximated Reformulation for A Consolidation

In determining $\mathbf{S}_k^*$ and T_k^* for a given consolidation Ω_k, the retailer should consider how much inbound transportation costs on average will be paid. However, inbound transportation decisions, i.e., $\mathbf{x}$ are dynamic in the sense that the retailer will find his/her optimal truck choices with every replenishment. Nevertheless, since $\mathbf{S}_k$ and T_k heavily affect the replenishment quantities, problem $(\mathbf{P}^{\Omega_k})$, therefore, explicitly includes the expected inbound transportation costs in finding $\mathbf{S}_k^*$ and T_k^*. This, in turn, results in the bi-level optimization problem given by $(\mathbf{P}^{\Omega_k})$. Specifically, the lower level of $(\mathbf{P}^{\Omega_k})$ is required in order to find exact expected inbound transportation costs per unit time. As aforementioned, bi-level optimization problems are complex, we, therefore, approximate $(\mathbf{P}^{\Omega_k})$ with a single-level optimization problem as follows.

Note that expected order quantity for each item in Ω_k will be equal to the expected demand during one replenishment cycle, i.e., $\lambda_i T_k\ \forall i \in \Omega_k$. Then, we approximate Eq. (3) by defining expected number of trucks of type j used for consolidation Ω_k, denoted by $\widetilde{x}_{jk}$. That is, we define Eq. (3) by assuming that, on average, the retailer decides to use $\widetilde{x}_{jk}$ number of type j trucks in each replenishment of the items in Ω_k. Let $\widetilde{\mathbf{x}}^k$ be the m-vector of $\widetilde{x}_{jk}$ values. Using this approximation, average shipment cost per replenishment of Ω_k amounts to $ITC_{\Omega_k}(\widetilde{\mathbf{x}}^k) = \sum_{j\in J} \widetilde{x}_{kj} R_j$. Then, the retailer's approximated total expected costs per unit time when items in Ω_k are consolidated, denoted by $\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k)$, are equal to

$$\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k) = \sum_{i\in\Omega_k} h_i s_i - \frac{T_k}{2}\sum_{i\in\Omega_k} h_i\lambda_i + \frac{1}{T_k}\sum_{i\in\Omega_k} a_i + \frac{1}{T_k}\sum_{i\in\Omega_k} p_i n(s_i, T_k) + \frac{1}{T_k}\sum_{j\in J}\widetilde{x}_{kj}R_j. \tag{5}$$

The only difference between Eq. (5) and Eq. (4) is that Eq. (5) uses $ITC_{\Omega_k}(\widetilde{\mathbf{x}}^k)$ while Eq. (4) requires the solution of Eq. (3) for any combinations of demand realizations of the items in Ω_k. Using Eq. (5), the retailer's optimization problem for consolidation with approximated total expected costs per unit time reads as

$$
\begin{array}{lll}
(\widetilde{\mathbf{P}}^{\Omega_k}) \quad \min\limits_{(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k)} & \widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k) & \\
\text{s.t.} & T_k \geq 0 & \\
& s_i \geq 0 & \forall i \in \Omega_k \\
& \sum_{j\in J} \widetilde{x}_{jk} W_j \geq \sum_{i\in\Omega_k} w_i \lambda_i T_k & \\
& \sum_{j\in J} \widetilde{x}_{jk} V_j \geq \sum_{i\in\Omega_k} v_i \lambda_i T_k & \\
& x_j \in \{0, 1, 2, \ldots\} & \forall j \in J.
\end{array}
$$

$(\widetilde{\mathbf{P}}^{\Omega_k})$ is a single-level mixed-integer nonlinear optimization problem. We note that $(\widetilde{\mathbf{P}}^{\Omega_k})$ is NP-hard as a special case of $(\widetilde{\mathbf{P}}^{\Omega_k})$ when $w_i = 0\ \forall i \in I$ (or $W_j \to \infty$) is an integer knapsack problem for given $\mathbf{S}_k$ and T_k. Therefore, we next develop an heuristic method to solve $(\widetilde{\mathbf{P}}^{\Omega_k})$.

Local Search Heuristic for Consolidation Approximation

We propose a local search heuristic for solving $(\widetilde{\mathbf{P}}^{\Omega_k})$, denoted by LSH-k. Particularly, LSH-k works as follows. Given $\widetilde{\mathbf{x}}^k$, we first determine $\mathbf{S}_k$ and T_k by solving $(\widetilde{\mathbf{P}}^{\Omega_k})$ with the given $\widetilde{\mathbf{x}}^k$. Given $\widetilde{\mathbf{x}}^k$, $(\widetilde{\mathbf{P}}^{\Omega_k})$ reduces to the following optimization problem:

$$
\begin{array}{ll}
(\widetilde{\mathbf{P}}^{\widetilde{\mathbf{x}}^k}) \quad \min\limits_{(\mathbf{S}_k, T_k)} & \widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k | \widetilde{\mathbf{x}}^k) \\
\text{s.t.} & T_k \leq \min\left\{ \frac{\sum_{j\in J}\widetilde{x}_{jk}W_j}{\sum_{i\in\Omega_k} w_i\lambda_i}, \frac{\sum_{j\in J}\widetilde{x}_{jk}V_j}{\sum_{i\in\Omega_k} v_i\lambda_i} \right\} \\
& T_k \geq 0 \\
& s_i \geq 0\ \forall i \in \Omega_k
\end{array}
$$

$(\widetilde{\mathbf{P}}^{\widetilde{\mathbf{x}}^k})$ is a nonlinear optimization problem. A common method to solve such nonlinear models is the Interior-Point (IP) method. Since $(\widetilde{\mathbf{P}}^{\widetilde{\mathbf{x}}^k})$ needs to be solved many times within LSH-k (which is also needed to be executed many times within GAP-P), we focus on developing an efficient method to find solutions for $(\widetilde{\mathbf{P}}^{\widetilde{\mathbf{x}}^k})$ in less computational times. In particular, given $\mathbf{S}_k$, if we overestimate the number of expected shortages for any item within one replenishment cycle and assume it is equal to the expected demand for that item within one replenishment cycle, i.e., $n_i(s_i, T_k) \cong \lambda_i T_k$; then, one can easily show that $T_k = \min\left\{\frac{\sum_{j\in J}\widetilde{x}_{jk}W_j}{\sum_{i\in\Omega_k} w_i\lambda_i}, \frac{\sum_{j\in J}\widetilde{x}_{jk}V_j}{\sum_{i\in\Omega_k} v_i\lambda_i}\right\}$ minimizes $\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k|\widetilde{\mathbf{x}}^k, \mathbf{S}_k)$ over the feasible T_k values of $(\widetilde{\mathbf{P}}^{\widetilde{\mathbf{x}}^k})$. Furthermore, given T_k, $\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k|\widetilde{\mathbf{x}}^k, T_k)$ is separable in and convex with respect to each s_i $i \in \Omega_k$; thus, it follows from the first order condition that s_i that minimizes $\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k|\widetilde{\mathbf{x}}^k, T_k)$ will be the solution of $F_i^{(T_k)}(s_i) = 1 - \frac{h_i T_k}{p_i}$, where $F_i^{(T_k)}(\cdot)$ is the cumulative distribution function of item i's demand over T_k time units (i.e., cumulative distribution of the normal random variable, $D_i^{(T_k)}$, with mean $\lambda_i T_k$ and standard deviation $\sigma_i\sqrt{T_k}$). Therefore, we accept the solution of $(\widetilde{\mathbf{P}}^{\widetilde{\mathbf{x}}^k})$, denoted by $\widetilde{\mathbf{S}}_k$ and $\widetilde{T}_k$, as given in the following equations:

$$\widetilde{T}_k = \min\left\{\frac{\sum_{j\in J}\widetilde{x}_{jk}W_j}{\sum_{i\in\Omega_k} w_i\lambda_i}, \frac{\sum_{j\in J}\widetilde{x}_{jk}V_j}{\sum_{i\in\Omega_k} v_i\lambda_i}\right\}, \tag{6}$$

$$F_i^{(T_k)}(\widetilde{s}_i) = 1 - \frac{h_i\widetilde{T}_k}{p_i}. \tag{7}$$

In Section , we compare Eqs. (6) and (7) to IP and it can be seen from Table 2 that Eqs. (6) and (7) are computationally very efficient compared to IP. Furthermore, the solution qualities are very close over the problem instances solved. Therefore, we use Eqs. (6) and (7) to solve $(\widetilde{\mathbf{P}}^{\widetilde{\mathbf{x}}^k})$.

Once $(\widetilde{\mathbf{P}}^{\widetilde{\mathbf{x}}^k})$ is solved, we calculate $\widetilde{G}_k(\widetilde{\mathbf{S}}_k, \widetilde{T}_k, \widetilde{\mathbf{x}}^k)$ as the cost value of $\widetilde{\mathbf{x}}^k$. After that, we check all neighbors of $\widetilde{\mathbf{x}}^k$. To do so, we increase and decrease (if possible) the number of trucks of each type by 1. That is, we increase $\widetilde{x}_{jk}$ by 1 and decrease $\widetilde{x}_{jk}$ by 1 (if $\widetilde{x}_{jk} \geq 1$) for each j. This generates all neighbors of $\widetilde{\mathbf{x}}^k$. If there is a neighbor with lower cost value, we take the neighbor with the lowest cost as the new solution and repeat the neighbor search with this solution. This process is repeated until no neighbor with lower cost value is determined. At termination, we are guaranteed with a local minimum.

To avoid getting a high-cost local minimum, we start the LSH-k with multiple $\widetilde{\mathbf{x}}^k$. Initially, we randomly generate m $\widetilde{\mathbf{x}}^k$ vectors such that $0 \leq \widetilde{x}_{jk} \leq u_k$ where $u^k = \max_{j\in J}\left\{\left\lceil\frac{\sum_{i\in\Omega_k} w_i\lambda_i t^{max}}{W_j}\right\rceil, \left\lceil\frac{\sum_{i\in\Omega_k} v_i\lambda_i t^{max}}{V_j}\right\rceil\right\}$ and $t^{max} = \max_{i\in\Omega_k}\{\sqrt{\frac{2a_i}{h_i\lambda_i}}\}$ (note that $\sqrt{\frac{2a_i}{h_i\lambda_i}}$ is the replenishment cycle length of item i assuming that $\sigma_i = 0$, i.e., the economic order quantity model); thus, u_k is the maximum number of trucks needed to ship total order quantity of the items in the consolidation assuming that each item's order quantity is given by the economic order quantity and a single truck type is used. The details of LSH-k for a given starting solution are explained below.

Local Search Heuristic for ($\widetilde{\mathbf{P}}^{\widetilde{\mathbf{x}}^k}$) (LSH-k)	
Step 0:	Let $\widetilde{\mathbf{x}}^k$ be given for a consolidation Ω_k.
Step 1:	Calculate $\mathbf{S}_k$ and T_k using Eqs. (6) and (7) and determine $\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k \vert \widetilde{\mathbf{x}}^k)$
Step 2:	For $j = 1 : m$
Step 3:	Let $\widetilde{\mathbf{x}}^{k[-j]} = \widetilde{x}_{jk}^{[+j]} = \widetilde{\mathbf{x}}^k$. If $\widetilde{x}_{jk}^{[-j]} > 0$, let $\widetilde{x}_{jk}^{[-j]} = \widetilde{x}_{jk}^{[-j]} - 1$; and, let $\widetilde{x}_{jk}^{[+j]} = \widetilde{x}_{jk}^{[+j]} + 1$
Step 4:	Calculate $\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^{k[-j]} \vert \widetilde{\mathbf{x}}^{k[-j]})$ and $\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^{k[+j]} \vert \widetilde{\mathbf{x}}^{k[+j]})$ using Eqs. (6) and (7)
Step 5:	End
Step 6:	If $\min_{j \in J}\{\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^{k[-j]} \vert \widetilde{\mathbf{x}}^{k[-j]}), \widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^{k[+j]} \vert \widetilde{\mathbf{x}}^{k[+j]})\} < \widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k \vert \widetilde{\mathbf{x}}^k)$
Step 7:	Set $\widetilde{\mathbf{x}}^k = \arg\min_{j \in J}\{\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^{k[-j]} \vert \widetilde{\mathbf{x}}^{k[-j]}), \widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^{k[+j]} \vert \widetilde{\mathbf{x}}^{k[+j]})\}$, go to Step 2
Step 7:	Else, terminate and return $\widetilde{\mathbf{x}}^k$

Mutation

Now suppose that we have a population of evaluated chromosomes, that is, the total approximated expected cost per unit time for each chromosome is known. Let $chrom^{dl}$ be the d^{th} $d \in \{1, 2, \ldots, pop^l\}$ chromosome in the l^{th} population, where pop^l is the number of chromosomes in the l^{th} population. Furthermore, let $\widetilde{C}(chrom^{dl})$ the total approximated expected cost per unit time of $chrom^{dl}$. Without loss of generality, let $\widetilde{C}(chrom^{1l}) < \widetilde{C}(chrom^{2l}) < \ldots < \widetilde{C}(chrom^{pop^l l})$. To generate the $(l+1)^{st}$ population, we execute the following three mutation operations:

(i) Local Mutation: A local mutation is applied to the chromosomes that are randomly selected from the first 45% of the pop^l chromosomes within the l^{th} population, i.e., the best 45% of the population. Local search mutation randomly picks an item i from a selected chromosome and randomly increases or decreases c_i of the chromosome by 1. For a given population of evaluated chromosomes, we generate $\lceil 0.45 pop^l \rceil$ new chromosomes at the end of local mutation operations.

(ii) Cross-Over: Cross-over mutation is applied to the chromosomes in the best 50% of the population. We randomly create pairs of two chromosomes from the best 50% of the population and perform a random single-point cross-over. Each pair of chromosomes crossed-over generates two new chromosomes, one from each chromosome within the pair. For a given population of evaluated chromosomes, we generate $\lceil 0.5 pop^l \rceil$ new chromosomes at the end of cross-over operations.

(iii) Random Mutation: Random mutation is applied to create a number of chromosomes so that the new population has the same population size with the current population. First, the number of chromosomes needed after local mutation and cross-over operations is determined. Then, chromosomes are randomly selected from the best 50% of the population and random mutation is applied. A random mutation on a selected chromosome randomly generates a c_i value such that $1 \leq c_i \leq n$ for a randomly selected item i.

At the end of mutation operations, the newly generated population has the same number of chromosomes with the previous population.

Termination

If there is no improvement in $\widetilde{C}(chrom^{1l})$ for L consecutive populations or O populations are evaluated, the GAP-P terminates.

NUMERICAL ANALYSIS

In this section, we focus on two sets of numerical analysis. In the first set of numerical analysis, the subroutine defined by Eqs. (6) and (7) is compared to IP and the approximated reformulation of a consolidation is tested with a simulation study. In the second set of numerical analysis, the cost and environmental benefits of consolidating items and using multiple truck types for shipment are illustrated. In both of the numerical analyses, the demand per unit time for any item i is assumed to be normally distributed with mean λ_i and standard deviation σ_i. The problem instances are randomly generated using uniform distributions with the given ranges in Table 1. Similar numerical values are assumed for these parameters in the literature on integrated inventory control and transportation (see, e.g., Toptal et al., 2003, Toptal and Çetinkaya, 2006, Toptal, 2009, Konur and Toptal, 2012). In all of the

Table 1: Problem Parameters

λ	$\sim$	$U[1750, 2250]$	w_i	$\sim$	$U[1, 4]$
σ	$\sim$	$U[150, 250]$	v_i	$\sim$	$U[0.5, 2]$
h_i	$\sim$	$U[1, 5]$	W_j	$\sim$	$U[200, 600]$
a_i	$\sim$	$U[50, 250]$	V_j	$\sim$	$U[100, 300]$
p_i	$\sim$	$U[2, 10]$	R_j	$\sim$	$U[150, 450]$

following analysis, we consider 15 different problem classes, each of which corresponds to a combination of $n = \{5, 10, 15, 20, 25\}$ and $m = \{5, 10, 15\}$. For each problem class, 10 problem instances are generated. The values shown in the tables of this section for a given problem class are the average values over all 10 problem instances solved within that problem class.

We first compare Eqs. (6) and (7) to IP. Here, we assume that all of the items are consolidated in one single group and the approximated truck choices for the consolidation is given as we are comparing two alternative solution methods for problem $(\widetilde{\mathbf{P}}^{\widetilde{\mathbf{x}}^k})$. That is, $\widetilde{\mathbf{x}}^k$ is given for Ω_k such that $\Omega_k = I$. Given the number of truck types, $\widetilde{\mathbf{x}}^k$ is randomly generated such that $\widetilde{x}_{jk} \in [0, 5]$. For each problem class, Table 2 shows average values, over the 10 randomly generated problem instances, for T_k and corresponding

$\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k|\widetilde{\mathbf{x}}^k)$ values along with the cpu times (in seconds) for Eqs. (6) and (7) and IP. Furthermore, the cost difference column gives the average difference in $\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k|\widetilde{\mathbf{x}}^k)$ values between Eqs. (6) and (7) and IP.

Table 2: Comparing Solution Methods for ($\widetilde{\mathbf{P}}^{\widetilde{\mathbf{x}}^k}$)

		Eqs. (6) and (7)			Interior-Point (IP)			Cost
n	m	$\widetilde{T}_k$	$\widetilde{G}_k$	cpu	$\widetilde{T}_k$	$\widetilde{G}_k$	cpu	Difference
5	5	0.93	36,307	0.001	0.78	35,927	0.272	0.90%
	10	1.91	48,520	0.001	1.64	47,238	0.275	2.63%
	15	2.70	57,123	0.001	2.02	53,281	0.260	6.74%
10	5	0.88	74,563	0.001	0.81	74,044	0.702	0.68%
	10	1.93	91,088	0.001	1.50	87,296	0.757	4.35%
	15	2.65	112,154	0.001	2.00	104,929	0.746	6.89%
15	5	0.80	106,073	0.001	0.76	105,705	1.270	0.35%
	10	1.65	135,966	0.001	1.44	132,389	1.390	2.66%
	15	2.62	164,629	0.001	1.90	154,024	1.290	6.70%
20	5	0.90	146,863	0.001	0.82	146,040	1.367	0.58%
	10	1.72	183,057	0.001	1.47	179,075	1.365	2.25%
	15	2.75	217,362	0.001	1.93	203,171	1.370	6.78%
25	5	0.87	183,909	0.001	0.79	183,602	1.362	0.21%
	10	1.69	229,604	0.001	1.37	223,512	1.356	2.70%
	15	2.62	274,293	0.001	1.87	256,578	1.358	6.79%
Average		1.78	137,434	0.001	1.41	132,454	1.009	3.41%

As it can be seen from Table 2, the average computational time with Eqs. (6) and (7) is significantly lower than the average computational time with IP. Moreover, while the IP method results in lower approximated costs, i.e., $\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k|\widetilde{\mathbf{x}}^k)$ values, Eqs. (6) and (7) were able to find good quality solutions; the increase in costs is less than 4% on average. Finally, T_k values returned by each alternative method are very close on average. Therefore, we can conclude that Eqs. (6) and (7) are efficient for solving ($\widetilde{\mathbf{P}}^{\widetilde{\mathbf{x}}^k}$) and we use them in GAP-P.

Next, we evaluate the approximated reformulation of a given consolidation. Recall that truck choice decisions are dynamic as the retailer can select the number of trucks of each type to ship each order. However, calculation of expected transportation costs resulted in bi-level optimization problem ($\mathbf{P}^{\Omega_k}$), which has been approximated by problem ($\widetilde{\mathbf{P}}^{\Omega_k}$). Particularly, in ($\widetilde{\mathbf{P}}^{\Omega_k}$), $\widetilde{\mathbf{x}}^k$ defines approximated number of trucks of each truck type to be used by the retailer for a given consolidation. To see how close $G_k(\mathbf{S}_k, T_k)$ and $\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k)$, we simulate the truck choice decisions as well as order quantity decisions for a given consolidation. Particularly, given a problem instance, we assume that all of the items are consolidated in one single group. Then, $\mathbf{S}_k$, T_k, and $\widetilde{\mathbf{x}}^k$ values are determined using LSH-k. After that, with the determined $\mathbf{S}_k$ and T_k values, we simulate 1,000 replenishment cycles for the problem instance (to do so, for each item $i \in I$, 1,000 demand realizations, i.e., $D_i^{(T_k)}$ values, are generated using normal distribution with mean $\lambda_i T_k$ and standard deviation $\sigma_i\sqrt{T_k}$). At each replenishment of the simulation, the best truck choices for the order are determined by solving Eq. (3) with CPLEX (as the number of decision variables are 15 maximum, it was not very time consuming to solve Eq. (3) at each of the 1,000 replenishment). As a result of simulation, we find the mean value of the cost per cycle and then determine the mean value of the cost per unit time, denoted by $\overline{G}_k(\mathbf{S}_k, T_k)$. Furthermore, we find the mean number of trucks of each type used, denoted by $\overline{x}_{jk}$. Note that, in the approximated formulation, $\overline{x}_{jk}$ is assumed to be given by $\widetilde{x}_{jk}$ values.

Table 3: Comparing Approximated and Simulated Results for A Given Consolidation

			Approximation		Simulation	
n	m	T_k	$\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k)$	$\sum_{j\in J}\widetilde{x}_{jk}$	$\overline{G}_k(\mathbf{S}_k, T_k)$	$\sum_{j\in J}\overline{x}_{jk}$
5	5	0.090	26,198	5.2	26,390	5.5
	10	0.121	25,574	7.1	23,793	7.0
	15	0.126	24,504	7.1	22,996	7.1
10	5	0.044	77,365	6.0	70,302	6.2
	10	0.077	60,447	10.0	50,552	9.5
	15	0.101	58,872	13.2	47,650	12.2
15	5	0.030	142,823	5.9	129,527	6.0
	10	0.053	116,794	10.7	92,551	9.5
	15	0.073	96,696	14.6	74,720	13.3
20	5	0.022	221,195	5.9	205,388	6.1
	10	0.041	164,768	11.0	135,303	10.7
	15	0.056	146,727	15.2	111,474	13.5
25	5	0.018	327,658	6.0	300,110	6.1
	10	0.035	212,767	10.5	176,580	9.6
	15	0.049	183,637	15.9	145,397	13.8
Average		0.062	125,735	9.6	107,516	9.1

In Table 3, the average values over the 10 randomly generated problem instances for T_k, $\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k)$, $\sum_{j\in J}\widetilde{x}_{jk}$, $\overline{G}_k(\mathbf{S}_k, T_k)$, and $\sum_{j\in J}\overline{x}_{jk}$ are documented for each problem class. It can be observed that from Table 3 that $\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k)$ over estimates $\overline{G}_k(\mathbf{S}_k, T_k)$ for most of the problem classes (and this was the case in most of the problem instances solved). This result was expected since the approximation reformulation does not define the minimum transportation costs in each replenishment. Specifically, $\widetilde{G}_k(\mathbf{S}_k, T_k, \widetilde{\mathbf{x}}^k)$ over estimated $\overline{G}_k(\mathbf{S}_k, T_k)$ by approximately 17% on average. Nevertheless, $\widetilde{x}_{jk}$ can over or under estimate $\overline{x}_{jk}$ values and the same observation holds for the total number of trucks used for inbound shipment; however, the difference between $\sum_{j\in J}\widetilde{x}_{jk}$ and $\sum_{j\in J}\overline{x}_{jk}$ is within ±15% and 6% on average. Based on these observations, we believe that approximation reformulation of a consolidation is sufficiently well reflecting the actual costs; hence, can be naively used to evaluate the cost performance of a given consolidation and find good $\mathbf{S}_k$ and T_k values for a given consolidation Ω_k.

The following numerical analyses document the cost and environmental benefits of consolidation. Specifically, we compare three consolidation policies: (i) consolidation policy, the consolidations adopted as the solution of problem (**P**) via GAP-P, (ii) no-consolidation policy, when all of the items are individually replenished, and (iii) single-consolidation policy, when all of the items are consolidated in a single group. For each policy, we determine approximated expected costs (denoted by $\widetilde{C}$) and truck density (denoted by ϕ). Truck density is defined as the average number of trucks used per unit time. Particularly, for a given consolidation of items, Ω_k, truck density, ϕ_k, is defined as follows:

$$\phi_k = \frac{\sum_{j \in J} \widetilde{x}_{jk}}{T_k}.$$

Then, truck density of a consolidation policy, is equal to the sum of the truck densities of the consolidations suggested by the policy. Table 4 gives the average values over 10 problem instances solved within each problem class for $\widetilde{C}$ and ϕ for each consolidation policy. Furthermore, we give the average values of the percent increases in $\widetilde{C}$ and ϕ (denoted by $\Delta\widetilde{C}$ and $\Delta\phi$, respectively) due to adopting no-consolidation and single-consolidation policies over the consolidation policy.

Table 4: Comparing Consolidation Strategies for (**P**)

		Consolidation		No Consolidation				Single Consolidation			
n	m	$\widetilde{C}$	ϕ	$\widetilde{C}$	ϕ	$\Delta\widetilde{C}$	$\Delta\phi$	$\widetilde{C}$	ϕ	$\Delta\widetilde{C}$	$\Delta\phi$
5	5	21,650	65.6	32,266	79.2	51.1%	34.0%	25,322	64.5	14.7%	9.1%
	10	20,896	58.1	28,099	63.3	35.9%	23.6%	25,422	57.2	17.1%	6.3%
	15	19,814	55.8	30,439	68.3	53.7%	19.5%	25,707	57.9	27.3%	4.3%
10	5	45,067	135.0	61,120	138.1	37.5%	8.4%	69,559	128.2	53.6%	3.4%
	10	42,020	118.5	62,761	150.3	50.5%	38.2%	63,063	131.6	53.1%	14.0%
	15	39,230	111.3	63,512	144.6	61.8%	25.5%	56,851	126.6	46.6%	14.1%
15	5	76,210	200.1	109,856	241.0	44.0%	23.4%	164,045	222.2	119.9%	23.4%
	10	60,394	179.3	99,554	230.3	64.8%	22.9%	104,445	194.0	77.1%	4.3%
	15	56,535	168.4	96,132	206.4	70.0%	12.5%	93,314	186.9	68.5%	10.3%
20	5	84,902	244.3	127,190	297.1	50.2%	20.4%	205,520	260.1	138.5%	11.3%
	10	78,609	233.1	131,619	278.5	68.2%	40.9%	156,509	260.6	103.4%	29.8%
	15	74,479	213.5	126,769	276.1	70.0%	33.3%	134,975	257.2	77.1%	23.5%
25	5	106,471	291.5	163,163	345.5	54.0%	26.0%	287,827	312.3	161.0%	7.1%
	10	106,574	288.7	182,291	395.7	72.3%	49.0%	235,357	347.1	126.2%	28.1%
	15	102,087	288.9	165,087	398.0	62.4%	24.6%	184,080	323.1	89.2%	9.1%
Average		62,329	176.8	98,657	220.8	56.4%	26.8%	122,133	195.3	78.2%	13.2%

As it can be seen from Table 4, consolidation policies heavily affect the costs and truck density. Specifically, a retailer can save in costs by efficiently determining which items will be consolidated. We note that both single-consolidation and no-consolidation policies are suboptimal for problem (**P**);

therefore, as expected, consolidation results in lower costs than no-consolidation and single-consolidation policies. Compared to no-consolidation policy, consolidation can save costs over 50% on average; and, compared to single-consolidation policy, consolidation can save costs over 75% on average over the problem instances solved. Furthermore, efficient consolidation can reduce truck density. As expected, truck density is the highest on average for no-consolidation policy as utilization of truck capacities is minimum in no-consolidation policy. Compared to no-consolidation policy, consolidation can decrease truck density over 25% on average; and, compared to single-consolidation policy, consolidation can decrease truck density over 10% on average. These observations suggest that efficient consolidation in multi-item inventory systems can save costs and result in environmental benefits significantly.

Finally, we compare the case when the retailer uses single truck type instead of multiple truck types for inbound transportation. Specifically, we assume that the retailer will select the truck type which minimizes the total of the approximated expected costs of the consolidations selected, i.e., sum of the costs defined in Eq. (5) over the consolidations. To find the single truck type to be used, we find the consolidation policy assuming single truck type via GAP-P for each truck type, and select the one which gives lower approximated expected costs. Table 5 gives the average values over 10 problem instances solved within each problem class for $\widetilde{C}$ and ϕ for inbound transportation with consideration of multiple truck types and single truck type. Furthermore, we give the average values of the percent increases in $\widetilde{C}$ and ϕ (denoted by $\Delta\widetilde{C}$ and $\Delta\phi$, respectively) due to adopting restricting single truck type for inbound shipment.

As expected and can be observed in Table 5, restricting single truck type for inbound shipment increases costs. On average, single truck type inbound shipment increases costs by 2.3% compared to allowing use of different truck types for inbound shipment. Furthermore, single truck type restriction increases the truck density by 4.9% on average over the problem instances solved. Therefore, one can conclude that consideration of different truck types simultaneously for inbound shipment can have cost savings as well as environmental benefits.

CONCLUSIONS

This report studies a multi-item inventory system with shipment consolidation and explicit TL transportation in a stochastic demand environment. A time-based order-up-to-level inventory policy is proposed for a set of consolidated items. Furthermore, a retailer's consolidation decisions are formulated as a set partitioning problem. Due to the complexity of the problem, heuristic methods are developed. First, for a given consolidation, an approximated reformulation of the time-based order-up-to-level

Table 5: Comparing Consolidation with Multiple Truck Types to Single Truck Type

		Multiple-Truck		Single-Truck			
n	m	$\widetilde{C}$	ϕ	$\widetilde{C}$	ϕ	$\Delta\widetilde{C}$	$\Delta\phi$
5	5	21,650	65.6	21,991	69	1.6%	4.8%
	10	20,896	58.1	21,108	59	0.8%	0.8%
	15	19,814	55.8	20,490	62	3.3%	9.7%
10	5	45,067	135.0	45,713	139	1.3%	3.5%
	10	42,020	118.5	42,785	123	1.8%	3.1%
	15	39,230	111.3	39,714	113	1.2%	1.7%
15	5	76,210	200.1	78,276	216	2.6%	6.3%
	10	60,394	179.3	62,054	189	2.8%	5.4%
	15	56,535	168.4	58,435	177	3.2%	4.7%
20	5	84,902	244.3	90,581	281	5.8%	15.3%
	10	78,609	233.1	79,519	240	1.1%	3.2%
	15	74,479	213.5	75,105	214	0.9%	0.3%
25	5	106,471	291.5	107,900	298	1.4%	2.0%
	10	106,574	288.7	110,478	311	3.7%	6.7%
	15	102,087	288.9	104,624	308	2.4%	6.2%
Average		62,329	176.8	63,918	186.6	2.3%	4.9%

inventory policy with heterogeneous freight trucks is provided. A local search heuristic is proposed for the approximated reformulation. This search heuristic is utilized in a genetic algorithm to find good-quality consolidation strategies for the retailer's consolidation problem.

This study contributes to the literature on multi-item inventory systems by explicitly accounting for transportation costs when heterogeneous freight trucks can be used for inbound shipment, proposing a practical inventory control policy for a set of consolidated items with distinct characteristics, and developing a solution method for determining consolidation strategies.

With a set of numerical studies, the accuracy of the approximated reformulation of a consolidation is presented. Furthermore, a set of numerical studies is conducted to illustrate the economical as well as environmental benefits of shipment consolidation with heterogeneous freight trucks. Specifically, it is observed that shipment consolidation not only saves costs but also reduces truck density. Reduced truck density implies less transportation emissions and less truck congestion.

A future research direction would be to analyze different inventory control policies for a given set of consolidated items. For instance, quantity-based order-up-to-level policy can be studied and compared to the time-based order-up-to-level policy examined in this report. Furthermore, joint replenishment problem with explicit transportation costs considering the availability of different truck types is a remaining problem to be investigated.

Bibliography

Atkins, D. R., Iyogun, P. O., 1988. Periodic versus 'can-order' policies for coordinated multi-item inventory systems. Management Science 34 (6), pp. 791–796.

Aucamp, D., 1982. Nonlinear freight costs in the EOQ problem. European Journal of Operational Research 9, 61–62.

Balintfy, J. L., 1964. On a basic class of multi-item inventory problems. Management Science 10 (2), 287–297.

Ben-Khedher, N., Yano, C. A., 1994. The multi-item replenishment problem with transportation and container effects. Transportation Science 28, 37–54.

Betts, J. M., Johnston, R. B., 2005. Determining the optimal constrained multi-item (q, r) inventory policy by maximising risk-adjusted profit. IMA Journal of Management Mathematics 16 (4), 317–338.

Çetinkaya, S., 2005. Coordination of inventory and shipment consolidation decisions: a review of premises, models, and justification. In: Geunes, J., Akçali, E., Pardalos, P. M., Romeijn, E. H. (Eds.), Applications of Supply Chain Management and e-commerce research in industry. Kluwer, New York, Ch. 1.

Cetinkaya, S., Mutlu, F., Lee, C.-Y., 2006. A comparison of outbound dispatch policies for integrated inventory and transportation decisions. European Journal of Operational Research 171 (3), 1094–1112.

FHWA, 2012. Freight facts and figures 2012. Tech. rep., U.S. Department of Transportation, Federal Highway Administration.

Garey, M. R., Johnson, D. S., 1979. Computers and Intractability: A Guide to the Theory of NP-Completeness. W. H. Freeman & Co., New York, NY.

Hansen, P., Jaumard, B., Savard, G., 1992. New branch-and-bound rules for linear bilevel programming. SIAM Journal on Scientific and Statistical Computing 13, 1194–1217.

Johansen, S. G., Melchiors, P., 2003. Can-order policy for the periodic-review joint replenishment problem. The Journal of the Operational Research Society 54 (3), pp. 283–290.

Khouja, M., Goyal, S., 2008. A review of the joint replenishment problem literature: 1989–2005. European Journal of Operational Research 186 (1), 1 – 16.

Kiesmuller, G., 2009. A multi-item periodic replenishment policy with full truckloads. International Journal of Production Economics 118 (1), 275 – 281.

Konur, D., Toptal, A., 2012. Analysis and applications of replenishment problems under stepwise transportation costs and generalized wholesale prices. International Journal of Production Economics 140 (1), 521–529.

Lee, C.-Y., 1986. The economic order quantity for freight discount costs. IIE Transactions 18 (3), 318–320.

Minner, S., Silver, E. A., 2005. Multi-product batch replenishment strategies under stochastic demand and a joint capacity constraint. IIE Transactions 37 (5), 469–479.

Mutlu, F., Cetinkaya, S., Bookbinder, J., 2010. An analytical model for computing the optimal time–and–quantity–based policy for consolidated shipments. IIE Transactions 42 (5), 367–377.

Nahmias, S., 2009. Production and Operations Analysis, 6th Edition. McGraw-Hill Irwin, New York, US.

Nielsen, C., Larsen, C., 2005. An analytical study of the q(s,s) policy applied to the joint replenishment problem. European Journal of Operational Research 163 (3), 721 – 732.

Ozkaya, B. Y., Gurler, U., Berk, E., 2006. The stochastic joint replenishment problem: A new policy, analysis, and insights. Naval Research Logistics (NRL) 53 (6), 525–546.

Pantumsinchai, P., 1992. A comparison of three joint ordering inventory policies. Decision Sciences 23, 111–127.

Toptal, A., 2009. Replenishment decisions under an all-units discount schedule and stepwise freight costs. European Journal of Operational Research 198, 504–510.

Toptal, A., Bingol, S. O., 2011. Transportation pricing of a truckload carrier. European Journal of Operational Research 214, 559–567.

Toptal, A., Çetinkaya, S., 2006. Contractual agreements for coordination and vendor-managed delivery under explicit transportation considerations. Naval Research Logistics 53 (5), 397–417.

Toptal, A., Çetinkaya, S., Lee, C.-Y., 2003. The buyer-vendor coordination problem: Modeling inbound and outbound cargo capacity and costs. IIE Transactions 35 (11), 987 – 1002.

Viswanathan, S., 1997. Periodic review (s, s) policies for joint replenishment inventory systems. Management Science 43 (10), pp. 1447–1454.

Zhao, X., Qiu, M., Xie, J., He, Q., 2012. Computing (r, q) policy for an inventory system with limited sharable resource. Computers & Operations Research 39 (10), 2368 – 2379.