A CONSTRUCTIVE HEURISTIC FOR THE INTEGRATED INVENTORY-DISTRIBUTION PROBLEM

Tamer F. Abdelmaguid*, Maged M. Dessouky** and Fernando Ordonez***

*Assistant Professor, Mechanical Design and Production Department
Faculty of Engineering, Cairo University, Giza 12613 - Egypt.
**Professor, ***Assistant Professor, Daniel J. Epstein Dept. of Industrial and Systems
Engineering, University of Southern California, California, USA.
tabdelmaguid@eng.cu.edu.eg

ABSTRACT

We study the integrated inventory distribution problem which is concerned with multiperiod inventory holding, backlogging, and vehicle routing decisions for a set of customers who receive units of a single item from a depot with infinite supply. We consider an environment in which the demand at each customer is deterministic and relatively small compared to the vehicle capacity, and the customers are located closely such that a consolidated shipping strategy is appropriate. We develop a constructive heuristic to obtain an approximate solution for this NP-hard problem and demonstrate its effectiveness through computational experiments.

KEYWORDS

Vehicle routing, inventory management, heuristics.

1. INTRODUCTION

Recent decades have seen fierce competition in local and global markets, forcing manufacturing enterprises to streamline their logistic systems, as they constitute over 30% of the cost of goods sold for many products [1]. The major components of logistic costs are transportation costs, representing approximately one third, and inventory costs, representing one fifth [2]. The transportation and inventory cost reduction problems have been thoroughly studied separately; while, the integrated problem has recently attracted more interest in the research community as new ideas of centralized supply chain management systems, such as vendor managed inventory (VMI), have gained acceptance in many supply chain environments.

The integration of transportation and inventory decisions is represented in the literature by a general class of problems referred to as dynamic routing and inventory (DRAI) problems. As defined by Baita et al. [3], this class of problems is “characterized by the simultaneous vehicle routing and inventory decisions that are present in a dynamic framework such that earlier decisions influence later decisions.” Baita et al. [3] classify the approaches used for DRAI problems into two categories. The first category operates in the frequency domain where the decision variables are replenishment frequencies, or headways between shipments. Examples in the literature include [4-8].
The second category, referred to as the time domain approach, determines the schedule of shipments. With discrete time models, quantities and routes are decided at fixed time intervals. Within this category the most famous problem is the inventory routing problem (IRP), which arises in the application of the distribution of industrial gases. The main concern for this kind of application is to maintain an adequate level of inventory for all customers and to avoid any stockout. In the IRP, it is assumed that each customer has a fixed demand rate and the focus is on minimizing the total transportation cost; while inventory costs are mostly not of concern. Examples of this application in the literature include [9-14].

In this paper, we consider a DRAI problem that addresses the integrated inventory and vehicle routing decisions in the time domain at the operational planning level. This problem, referred to as the Integrated Inventory Distribution Problem (IIDP), considers multiple planning periods, both inventory and transportation costs, and a situation in which backorders are permitted. The kind of application that permits backorders is, of course, different from the distribution of industrial gases, where no shortage is allowed. The proposed model is suitable to industrial applications in which a manufacturer distributes its product to geographically disbursed factories/retailers which are located in cities close to its warehouse. At the operational planning level, backorder decisions are generally justified in two cases. The first is when there is a transportation cost saving that is higher than the incurred shortage cost by a customer. The second case is when there is insufficient vehicle capacity to deliver to a customer given that renting additional vehicles is not an option due to technological or economic constraints.

In the literature, the integration of vehicle routing and inventory decisions with the consideration of inventory costs in the time domain approaches of the DRAI problems has taken different forms. In a few cases a single period planning problem has been addressed as found in [15] and [16]. In the multi-period problem, the decisions are conducted for a specific number of planning periods, or the problem is reduced to a single period problem by considering the effect of the long term decisions on the short term ones. Examples include [17-19].

Other researchers take into consideration various forms such as distributing perishable products [20], and the consideration of the time value of money for long-term planning [21]. Some work focused on different structures of the distribution network such as the case of satellite facilities [22], the case where warehouses act as transshipment points in a 3-level distribution network [23], and the case of a multi-depot problem [24-25].

Solution heuristics that have been proposed in the literature for the different variations of the integrated inventory-distribution problem, particularly the inventory routing problem, are either based on subgradient optimization of a Lagrangian relaxation as in [9] and [16] or constructive and improvement heuristics. The constructive heuristics are broadly classified into heuristics that allocate customers to service days and then solve a VRP to generate vehicle routes for each day [12]; and heuristics that allocate customers to days and vehicles and then solve a traveling salesman problem for every assignment [11]. Improvement heuristics found in the literature [15 and 26] are generally considered as extensions to the arc-exchange and node-exchange heuristics as found in the vehicle routing literature.

In the literature of the time domain approaches of the DRAI problems, some models in the case of multi-period planning may include shortage or stockout costs; however, backorder
decisions are generally not explicitly considered. Instead, the shortage or stock-out cost is treated as the penalty cost that is incurred due to making direct deliveries to customers whose demand is not fulfilled in the regular delivery route in a given period. Examples of such models in the literature include [19] and [27]. In this paper, we consider a situation in which backorder decisions are either unavoidable or more economical, and they have to be coordinated with other inventory holding and vehicle routing decisions over a specific planning horizon. We introduce a constructive heuristic for solving this NP-hard problem, and benchmark it against lower and upper bounds obtained by a commercial software package, CPLEX. The constructive heuristic introduced in this paper is an enhanced version of a previously developed one in [28].

The rest of this paper is organized as follows. In section 2 we formulate the problem as a mixed integer linear program. The motivating ideas and search plan for the developed heuristic is presented in section 3. Section 4 provides description of the developed constructive heuristic. In section 5, the experimental results are presented followed by the conclusion and directions for future research in section 6.

2. PROBLEM DESCRIPTION AND MIXED INTEGER PROGRAMMING MODEL

In the IIDP, we study a distribution system consisting of a depot, denoted 0, and geographically dispersed customers, indexed 1,...,N. Each customer i faces a different demand \( d_{it} \) for a single item per time period \( t \) (day/week). As traditionally considered, a single item does not restrict the problem to the case of a single product distribution, as the word ‘item’ can refer to a unit weight or volume of the distributed products and each customer can be viewed as a consumption center for packages of unit weight or volume [08]. Accordingly, the proposed model can be applied to the case of multiple products given that the values of the inventory holding and shortage costs per unit volume/weight have small variance among the different products. We consider the case in which the demand of each customer is relatively small compared to the vehicle capacity, and the customers are located closely such that a consolidated shipping strategy is appropriate. Deliveries to customers 1,...,N are to be made by a capacitated heterogeneous fleet of \( V \) vehicles, each with capacity \( q_v \) starting from the depot at the beginning of each period. Vehicles must return to the depot at the end of the period, and no further delivery assignments should be made in the same period. In this model, we consider the case in which renting additional vehicles is not an option due to technological or economic constraints, and it is assumed that the fleet of vehicles remains unchanged throughout the planning horizon.

Each customer \( i \) maintains its own inventory up to capacity \( C_i \) and incurs inventory carrying cost of \( h_i \) per period per unit and a backorder penalty (shortage cost) of \( \pi_i \) per period per unit on the end of period inventory position. We assume that the depot has sufficient supply of items that can cover all customers’ demands throughout the planning horizon. The planning horizon considers \( T \) periods. Transportation costs include \( f_t \) a fixed usage cost per vehicle, which depends on the period \( t \), and \( c_{ij} \) a variable transportation cost between \( i \) and \( j \), which satisfies the triangular inequality. The objective is to minimize the overall transportation, inventory carrying and backlogging costs incurred over a specific planning horizon. We consider an integer variable \( x_{vij} \), which equals 1 if vehicle \( v \) travels from \( i \) to \( j \) in period \( t \), and 0 if it does not. The amount transported on that trip is represented by \( y_{vij} \).
inventory and backorder at the end of time \( t \) is \( I_{it} \) and \( B_{it} \) respectively. The following is a mixed integer programming formulation for the problem.

\[
\text{Min} \sum_{t=1}^{T} \left[ \sum_{j=1}^{N} \sum_{v=1}^{V} f_{ij} \cdot x_{0jv}^v + \sum_{i=0}^{N} \sum_{j=0}^{N} \sum_{l=0}^{N} \sum_{v=1}^{V} c_{ij} \cdot x_{ijv}^v + \sum_{i=1}^{N} \left( h_i \cdot I_{it} + \pi_i \cdot B_{it} \right) \right]
\]

subject to:

1. \( \sum_{j=0}^{N} x_{ijv}^v \leq 1 \quad i = 0, \ldots, N, \ t = 1, \ldots, T \ \text{and} \ v = 1, \ldots, V \) (1)

2. \( \sum_{k=0}^{N} x_{ilk}^v - \sum_{l=0}^{N} x_{lij}^v = 0 \quad i = 0, \ldots, N, \ t = 1, \ldots, T \ \text{and} \ v = 1, \ldots, V \) (2)

3. \( y_{ijv}^v - q_i \cdot x_{ijv}^v \leq 0 \quad i = 0, \ldots, N, \ j = 0, \ldots, N, \ i \neq j, \ t = 1, \ldots, T \ \text{and} \ v = 1, \ldots, V \) (3)

4. \( \sum_{k=0}^{N} y_{ilk}^v - \sum_{l=0}^{N} y_{lij}^v \leq 0 \quad i = 1, \ldots, N, \ t = 1, \ldots, T \ \text{and} \ v = 1, \ldots, V \) (4)

5. \( I_{it-1} - B_{it-1} - I_{it} + B_{it} + \sum_{v=1}^{V} \left( \sum_{l=0}^{N} y_{ilk}^v - \sum_{l=0}^{N} y_{lij}^v \right) = d_{it} \quad i = 1, \ldots, N \ \text{and} \ t = 1, \ldots, T \) (5)

6. \( I_{it} \leq C_i \quad i = 1, \ldots, N \ \text{and} \ t = 1, \ldots, T \) (6)

7. \( I_{it} \geq 0 \quad i = 1, \ldots, N \ \text{and} \ t = 1, \ldots, T \) (7)

8. \( B_{it} \geq 0 \quad i = 1, \ldots, N \ \text{and} \ t = 1, \ldots, T \) (8)

9. \( y_{ijv}^v \geq 0 \quad i = 0, \ldots, N, \ j = 0, \ldots, N, \ i \neq j, \ t = 1, \ldots, T \ \text{and} \ v = 1, \ldots, V \) (9)

10. \( x_{ijv}^v = 0 \ \text{or} \ 1 \quad i = 0, \ldots, N, \ j = 0, \ldots, N, \ i \neq j, \ t = 1, \ldots, T \ \text{and} \ v = 1, \ldots, V \) (10)

The objective function (0) includes transportation costs and inventory carrying and shortage costs on the end of period inventory position. Constraints (1) make sure that a vehicle will visit a location no more than once in a time period, and constraints (2) ensure route continuity. Constraints (3) serve for two purposes. The first one is to ensure that the amount transported between two locations will always be zero whenever there is no vehicle moving between these locations, and the second is to ensure that the amount transported is less than or equal to the vehicle’s capacity. Constraints (4) are necessary to eliminate sub-tours. Constraints (5) are the inventory balance equations for the customers. Constraints (6) limit the inventory level of the customers to the corresponding storage capacity. It is assumed that the amount consumed
by each customer in a given period is not kept in the customer’s storage location; accordingly, it is not accounted for in constraints (6). Constraints (7) to (10) are the domain constraints.

3. MOTIVATING IDEAS AND HEURISTIC DESIGN

The IIDP is NP-hard since it includes the capacitated vehicle routing problem (VRP) as a subproblem. In this section we present the key ideas behind the proposed constructive heuristic for this problem.

A key decision in solving the IIDP is the amount delivered to customer \( i \) in period \( t \), as this quantity, let us define it by

\[
w_{it} = \sum_{v=1}^{V} \left( \sum_{j=0}^{N} y_{ijt}^v - \sum_{k=0}^{N} y_{ikt}^v \right),
\]

effectively separates the routing and inventory problems. In fact, given delivery values \( w_{it} \) for all customers and periods, the inventory and backorder problem is decided by minimizing the last term in the objective function of the MIP formulation subject to constraints (5) to (8). At the same time, the best routing solution for these \( w_{it} \) is obtained by solving \( T \) separate capacitated vehicle routing problems. Each VRP computes the optimal transportation costs to deliver \( W_{it} = (w_{it}, i = 1, \ldots, N) \) in period \( t \) by solving the following feasible problem if the delivery amounts satisfy \( \sum_{i=1}^{N} w_{it} \leq \sum_{v=1}^{V} q_v \):

\[
TC(A(W_i)) = \min \sum_{j=0}^{N} \sum_{v=1}^{V} f_{ijv} x_{ijt}^v + \sum_{j=0}^{N} \sum_{v=1}^{V} \sum_{i=0}^{N} c_{ijv} x_{ijt}^v
\]

Subject to:

\[
\sum_{j=0}^{N} x_{ijt}^v \leq 1 \quad i = 0, \ldots, N \text{ and } v = 1, \ldots, V \quad (1')
\]

\[
\sum_{k=0}^{N} x_{ikt}^v - \sum_{l=0}^{N} x_{lit}^v = 0 \quad i = 0, \ldots, N \text{ and } v = 1, \ldots, V \quad (2')
\]

\[
y_{ijt}^v - q_j x_{ijt}^v \leq 0 \quad i, j = 0, \ldots, N, i \neq j \text{ and } v = 1, \ldots, V \quad (3')
\]

\[
\sum_{k=0}^{N} y_{ikt}^v - \sum_{l=0}^{N} y_{lit}^v \leq 0 \quad i = 1, \ldots, N, \text{ and } v = 1, \ldots, V \quad (4')
\]

\[
\sum_{v=1}^{V} \left( \sum_{l=0}^{N} y_{ilt}^v - \sum_{k=0}^{N} y_{ikt}^v \right) = w_{it} \quad i = 1, \ldots, N \quad (10)
\]

\[
y_{ijt}^v \geq 0 \text{ and } x_{ijt}^v = 0 \text{ or } 1 \quad i, j = 0, \ldots, N, i \neq j \text{ and } v = 1, \ldots, V \quad (11)
\]

Therefore, the key in solving the IIDP is to be able to identify the optimal delivery amounts \( w_{it} \) since for a given values of the \( w_{it} \) variables, the IIDP can be separated into inventory and
routing problems for which there exist several efficient algorithms. Our proposed heuristic builds on this observation by focusing on how to determine the \( w_{it} \) variables efficiently. The procedure used to determine the \( w_{it} \) values must take into consideration the tradeoff existing between inventory and transportation costs.

In section 4 we propose a constructive heuristic that sets the delivery amounts by balancing this tradeoff. The idea of the heuristic is to estimate a transportation cost value for each customer in each period from an approximate routing solution. Actual delivery amounts, \( w_{it} \), are then decided by comparing these transportation cost estimates with the corresponding inventory costs. This process is done sequentially from the first period onward and in each period the comparison of transportation and inventory costs is done in two phases. The first phase looks into backorder decisions that are either imposed by insufficient vehicle capacity or preferred due to savings in transportation costs that are higher than backordering costs. The second phase investigates inventory decisions that would cover demand requirements in future periods in the case that excess vehicle capacity is available at the current period. The heuristic looks into inventory decisions that provide savings in future transportation costs that are higher than inventory carrying costs.

A key step in this heuristic is to be able to effectively estimate the transportation cost of each customer. Below we present a result that provides insight into the structure of the total transportation cost in period \( t \) as a function of the delivery amount \( W_t \).

**Proposition 1.** \( TC(W_i) \) is a multi-dimensional monotone increasing step function.

**Proof.** Given that the definition of \( TC(W_i) \) is based on an MIP model for the capacitated vehicle routing problem (VRP) in which triangular inequality holds. Starting from an optimal solution of a specific VRP at an initial \( W^0_t = (w^0_i : i = 1, \ldots, N) \), and by adding \( \Delta W^+_t = (\partial w^0_i : \partial w^0_i \geq 0, i = 1, \ldots, N) \) to \( W^0_t \) (i.e. increasing the demand values for a subset of the customers) such that \( \sum_{i=1}^N (w^0_i + \partial w^0_i) \leq \sum_{v=1}^V q_v \), one of two possible consequences will occur: 1) new arc or arcs will be added to the current solution to satisfy the vehicle capacity constraints \((3')\), which will increase \( TC(W^0_t) \) by the corresponding \( c_{ij} \) and/or \( f_i \) amounts as needed, or 2) the current VRP solution remains optimal. Thus \( TC(W_t + \Delta W^+_t) \geq TC(W_t) \) when \( W_t + \Delta W^+_t > W_t \). Since the changes of \( TC(W_t) \) occur at discrete points according to the vehicle capacities, \( TC(W_t) \) takes the form of a multidimensional step function. □

As a result of proposition 1, the solution scheme can focus only on those values of the continuous variables, \( w_{its} \), at which changes to the transportation cost occur. We can look at this result from another perspective. Given planned delivery amounts to customers in a period, by reducing the delivery quantity of a specific customer, the transportation costs will be reduced at discrete points and the maximum possible reduction will occur when the delivery to that customer is dropped to zero. Although proposition 1 is proven for optimal solutions to the VRP, this result can still be used for solutions generated by efficient heuristics as an approximation, such as the savings algorithm [29].
4. THE CONSTRUCTIVE HEURISTIC

As mentioned earlier, the constructive heuristic is based on the idea of estimating a transportation cost value for each customer in each period, which is necessary to facilitate the comparison between transportation and inventory carrying and shortage costs. We therefore refer to the constructive heuristic as the Estimated Transportation Costs Heuristic (ETCH). In subsection 4.1, we describe how the transportation cost estimates are evaluated and continuously updated throughout the course of the heuristic. Using these estimates, we show in subsection 4.2 how the inventory problem in IIDP can be decomposed into two subproblems that are solved by the heuristic in two phases. The solution techniques for these subproblems are illustrated in subsection 4.3.

4.1. Estimating transportation costs

Let \( w_{it}^{PL} \) be the planned delivery amount for customer \( i \) in period \( t \). For period \( \tau \) in which
\[
\sum_{j=1}^{N} w_{jt}^{PL} \leq \sum_{v=1}^{V} q_v,
\]
let \( W_\tau = (w_{jt} : w_{jt} = w_{jt}^{PL}, j = 1, \ldots, N) \). For customer \( i \) whose \( w_{it}^{PL} > 0 \), let
\[
W_\tau^{(i)} = (w_{jt} : w_{jt} = 0, j = 1, \ldots, N, j \neq i).
\]
Then, the transportation cost reduction that would result from reducing customer \( i \)'s delivery in period \( \tau \) to zero can be calculated as
\[
TC_\tau(W_\tau) - TC_\tau(W_\tau^{(i)}).
\]
Since the transportation cost function involves the solution of a VRP, which is known to be NP-hard, it may not be possible to calculate its exact value, especially for large problem sizes; instead, an efficient heuristic can be used to approximate it. In our implementation, the savings algorithm is used for this purpose.

Let \( ATC_\tau(W_\tau) \) be an approximation for \( TC_\tau(W_\tau) \) when the savings algorithm is used to solve the associated VRP. The transportation cost estimate for customer \( i \) in period \( \tau \) is calculated as
\[
ETC_\tau(W_\tau) = ATC_\tau(W_\tau) - ATC_\tau(W_\tau^{(i)}).
\]
However, resolving a VRP every time the transportation cost estimate for each customer is calculated may be computationally inefficient. Instead, a faster approximation scheme can be constructed by evaluating the transportation cost saving that will result when a customer is removed from its delivery tour assigned to it in a given VRP solution. This means that for given delivery amounts, \( W_\tau \), the associated VRP will be solved only once and the resulting vehicle tours will be used for generating transportation cost estimates.

\( ATC_\tau(W_\tau) \) and \( ETC_\tau(W_\tau) \) are functions of the planned delivery amounts \( w_{it}^{PL} \) which are determined based on the customers’ net demand requirements in period \( t \). However, the values of \( w_{it}^{PL} \) must be defined such that the vehicle capacity constraint,
\[
\sum_{j=1}^{N} w_{jt}^{PL} \leq \sum_{v=1}^{V} q_v,
\]
is satisfied.

Given the inventory position at the beginning of period \( t \), \( I_{i,t-1} - B_{i,t-1} \), and the demand requirements \( d_{it} \) for all periods \( \tau \geq t \), ETCH evaluates the net demand requirement for each customer, and based on that it estimates \( w_{it}^{PL} \). If the vehicle capacity constraint is not satisfied in a given period, the \( w_{it}^{PL} \) values are adjusted such that customers with the lowest unit
shortage costs, $\pi_i$, will have part of their demand requirements postponed to future periods. The following list describes the steps of this approach.

Procedure PLNDLV($t$)

1. Let $OC = \text{ordered set of all customers in which customers are sorted in a non-increasing order of their } \pi_i$ values;
2. For every customer $i \in OC$, let $inv_i = I_{i,t-1} - B_{i,t-1}$;
3. For period $\tau = t$ to $T$ do
   3.1. Let $Q_{\text{max}} = \sum_{v=1}^{V} q_v$;
   3.2. For every customer $i \in OC$ using the order in set $OC$ do
      3.2.1. $w^{PL}_i = \min(Q_{\text{max}}, \max(d_i - inv_i, 0))$;
      3.2.2. $Q_{\text{max}} = \max(Q_{\text{max}} - w^{PL}_i, 0)$;
      3.2.3. $inv_i = inv_i + w^{PL}_i - d_i$;
   End-Loop;
End-Loop;

The resultant $w^{PL}_i$ values can be safely used in evaluating both functions $ATC_t(W_t)$ and $ETC_t(W_t)$. During the course of the algorithm, if a change in the planned delivery amounts occurs, a VRP for the period in which the change occurred is instantiated and solved to update the values of the transportation cost estimates.

4.2. Problem decomposition and solution scheme

In the ETCH procedure, the comparison between the transportation cost estimates and inventory carrying and shortage costs is separated into two subproblems that are solved sequentially. This comparison is conducted for every period $t$ starting from the first period onward. The first subproblem is concerned with deciding whether to have backorders on period $t$ and the second subproblem is concerned with deciding whether to use remaining vehicle capacity in period $t$, if any, to cover future customer demand.

Backorders can be profitable for two reasons; it is either cheaper to pay the backorder cost than the transportation cost, or there is insufficient capacity in the vehicles to satisfy demand. Let $\delta_{i,t} = \max(d_{i,t} - I_{i,t-1} + B_{i,t-1}, 0)$ be the outstanding demand at customer $i$ at the beginning of period $t$, and $CD$ be the set of customers that have $\delta_{i,t} > 0$. The following subproblem decides whether to deliver to customer $i$ in period $t$ or not ($z_i = 1$ or $0$ respectively) and the quantity $r_i$ to deliver such that the sum of backorder cost and estimated transportation cost is minimized and vehicle capacity constraints are satisfied.

\[
[\text{SUB1}] - \text{Backorder decisions subproblem}
\]
\[
\begin{align*}
\text{Min} & \quad ATC_t(\Omega_t) + \sum_{i \in CD} \pi_i (\delta_{i,t} - r_i) \\
\text{Subject to:} & 
\end{align*}
\]
In SUB1, the objective function is composed of two parts, an approximation of the transportation costs in period $t$ and backorder penalty costs. Both parts are functions of the decision variables $r_i$. Constraint (12) ensures that we do not exceed the total vehicle capacity, and constraints (13) enforces that we deliver the exact amount of the outstanding demand only to customers included in the delivery in period $t$. Constraint (14) defines the vector of delivery amounts used in approximating the transportation cost function.

The main outcome from solving SUB1 is the backorder decisions evaluated as $B_{it} = \delta_{ij}$ for every customer $i \in CD$ that has $z_i = 0$, and accordingly $w_{it} = 0$, in the solution of SUB1. The delivery amounts, $w_{it}$, for customers in set $CD$ that have $z_i = 1$ in the solution of SUB1 are not decided yet as future demand requirements may be added. These decisions are investigated through subproblem SUB2. For every other customer $j \notin CD$, $w_{jt} = 0$, $B_{jt} = 0$ and $I_{jt} = I_{jt-1} - d_{jt}$.

Let $FD$ be the set of customers that have $z_i = 1$ in the solution of SUB1. Consider the integer variable $u_{it}$ to decide whether to deliver customer $i$’s demand for period $t$, where $\tau > t$. Let $Q'$ denote the total remaining vehicle capacity, i.e. $Q' = \sum_{v=1}^{V} q_v - \sum_{i \in CD} r_i$, and let $T_{i,\max}^\tau$ be the latest period where customer $i$’s demand can be considered without violating its storage capacity constraint, i.e. $T_{i,\max}^\tau = \max \left\{ \arg\max_{L} \left( \sum_{\tau=t+1}^{L} d_{it} \leq C_i \right), T \right\}$. We also define $T_{\max} = \max_{i} (T_{i,\max}^\tau)$.

Let $w_{it}^{pl}$ be the planned delivery amount for customer $i$ in a future period $\tau > t$. The values of $w_{it}^{pl}$ are initially calculated using the PLNDLV($t+1$) procedure as described in subsection 4.1 with a small modification to make sure that for every customer $j \in FD$, initial values of $w_{jt}^{pl} = d_{jt}$. If it is not possible to achieve this condition in a future period $\tau$ for customer $j \in FD$, $T_{j,\max}^\tau$ is set to $t-1$. The $w_{it}^{pl}$ values for customers that do not belong to set $FD$ are fixed; however, the values of $w_{it}^{pl}$ for customers in set $FD$ change with the change of the $u_{it}$ decision variables. The following problem decides whether to include future demand for any customer in the current delivery by minimizing the total transportation and inventory costs and satisfying capacity limits. This part is formulated as follows:

\[
\text{[SUB2] - Inventory decisions subproblem}
\]

\[
\begin{align*}
\min & \sum_{\tau=t+1}^{T_{\max}} \sum_{i \in FD} \sum_{\tau=t+1}^{T_{\max}} A_{IT}(\Omega_i) + \sum_{i \in FD} \sum_{\tau=t+1}^{T_{\max}} [(\tau-t)h_i d_{it}] u_{it} \\
\end{align*}
\]
Subject to:

\[ \sum_{i=FD}^{T_{\text{max}}} d_{i\tau} u_{i\tau} \leq Q \]  
(16)

\[ u_{i\tau+1} \geq u_{i\tau} \]  
\( \tau = t+1, \ldots, T_{i\text{max}}, \forall i \in FD \)  
(17)

\[ w_{i\tau}^{pl} = d_{i\tau} (1-u_{i\tau}) \]  
\( \tau = t+1, \ldots, T_{i\text{max}}, \forall i \in FD \)  
(18)

\[ \Omega_{\tau} = (\omega_{i\tau} : \omega_{i\tau} = w_{i\tau}^{pl}, i = 1, \ldots, N) \]  
\( \tau = t+1, \ldots, T_{i\text{max}} \)  
(19)

\[ u_{i\tau} = 0 \text{ or } 1. \]  
\( \tau = t+1, \ldots, T_{i\text{max}}, \forall i \in FD \)  
(20)

Constraint (16) represents the available vehicle capacity limit. For simplification, the customers’ storage limits are represented by the time index \( T_{i\text{max}} \), which is computed in advance as described earlier. The precedence constraints (17) are added to represent the fact that future demand in a certain period is to be considered only if the customer’s preceding period demand is fulfilled. Constraints (18) define the relationship between the future planned delivery amounts for customers in set \( FD \) and the decision variables \( u_{i\tau} \). SUB2 neglects the effect of changes in transportation cost in period \( t \) that may result from changing delivery amounts in that period.

\[ \text{Fig. 1. An outline of ETCH} \]
By solving SUB2, the delivery amounts for customers in set FD can be calculated as \( w_{it} = r_i + \sum_{t=1}^{T_{\text{max}}} d_{it}u_{it} \). Accordingly, the inventory and backorder decision variables in period \( t \) can be easily calculated. Finally, delivery routes in period \( t \) are decided by solving a VRP using the resulting delivery amounts. The flow chart in Fig. 1 summarizes the major steps of the proposed heuristic. The following subsection provides the algorithmic solutions for both subproblems and their related analyses.

4.3. Solving subproblems

The two subproblems are resource allocation problems in which the scarce resource is the associated available vehicle capacity and the main decision variables, \( z_i \) and \( u_{it} \), are binary variables. Accordingly, both of them can be solved optimally using dynamic programming (DP) as described in [30]. However, with the increase of the problem size, mainly due to the number of customers and the planning horizon, the DP implementations suffer from the curse of dimensionality. In this section, we present efficient heuristics that can be used instead. First, we present the following result that characterizes optimal solutions to subproblem SUB1.

**Proposition 2.** There is an optimal solution to SUB1 that only makes deliveries to customer \( i \) if the quantity delivered satisfies \( r_i > \text{ETC}_i(\Omega_i) / \pi_i \). Also, every optimal solution to SUB1 only makes deliveries if \( r_i \geq \text{ETC}_i(\Omega_i) / \pi_i \).

**Proof.** Assume that in the optimal solution to SUB1, some customer \( i \) is delivered \( r_i \) that satisfies \( r_i \leq \text{ETC}_i(\Omega_i) / \pi_i \), or equivalently \( \pi_i(\delta_i - r_i) + \text{ATR}_i(\Omega_i) \geq \pi_i\delta_i + \text{ATR}_i(\Omega_i(i)) \). If we consider the modified solution obtained by setting \( z_i = r_i = 0 \), then the previous inequality shows that the modified solution, which is feasible, is at least as good as the optimal solution. In the case when \( r_i < \text{ETC}_i(\Omega_i) / \pi_i \), then the modified solution is strictly better. Thus, the original solution cannot be optimal. \( \square \)

Proposition 2 gives a necessary condition for the optimality of the delivery decision made for a specific customer; however, satisfying this condition for all customers that have planned deliveries does not guarantee optimality for the solution of SUB1. Yet, since backorder decisions are generally not preferable, we will consider solutions that have this characteristic sufficiently good. We design the following algorithm that utilizes this rule.

Let \( DL_k = \{ dl: dl \subseteq CD \text{ and } |dl| = |CD| - k \} \), where \( |.| \) denotes the size of a set. We define \( f^{\text{SUB1}}(dl) \) as the objective function value of subproblem SUB1 when \( z_i = 1 \) for every customer \( i \in dl \) and \( z_j = 0 \) for every customer \( j \in CD - dl \), where \( dl \in DL_k \) for some \( k \). If the vehicle capacity constraint of SUB1 associated with setting \( z_i = 1 \) for all customers in a set \( dl \) is not satisfied, we define \( f^{\text{SUB1}}(dl) = \infty \). The following list describes the steps of a breadth-first-based heuristic approach that searches for efficient solutions to SUB1.

**Procedure SUBALG1**

1. Let \( k = 1 \) and \( dl^{\text{min}} = CD \);
2. If \( f^{\text{SUB1}}(dl^{\text{min}}) \neq \infty \) and \( r_i \geq \text{ETC}_i(\Omega_i) / \pi_i \ \forall \ i \in dl^{\text{min}} \) then go to 9;
3. For every \( dl \in DL_k \) evaluate \( f^{\text{SUB1}}(dl) \);
4. Find \( dl' \) from set \( DL_k \) that has the minimum \( f^{\text{SUB1}}(dl) \) selected from the members of \( DL_k \) that satisfy the following conditions:
   \[ f^{\text{SUB1}}(dl) \neq \infty \ \text{and} \ r_i \geq \text{ETC}_i(\Omega_i) / \pi_i \ \forall \ i \in dl; \]
5. If \( dl' \neq \emptyset \) then let \( dl^{\text{min}} = dl' \) go to 9;
6. Find \( dl' \) from set \( DL_k \) that has the minimum \( f^{\text{SUB1}}(dl) \);
7. If \( f^{\text{SUB1}}(dl') < f^{\text{SUB1}}(dl^{\text{min}}) \) then let \( dl^{\text{min}} = dl' \);
8. If \( k < |CD| \) then let \( k = k + 1 \), go to 3;
9. Generate a solution for \( \text{SUB1} \) in which deliveries are only made to customers in set \( dl^{\text{min}} \).

\text{SUBALG1} \text{ evaluates the } f^{\text{SUB1}}(dl) \text{ value for every set } dl \in DL_k \text{ at values of } k=0,\ldots,|CD|. \text{ If at some level of } k, \text{ the condition that } r_i \geq \text{ETC}_i(\Omega_i) / \pi_i \text{ is satisfied for all } i \in dl, \text{ we find an approximate solution and the algorithm terminates. However, if steps 2, 4 and 5 are removed, the algorithm guarantees that an optimal solution for \( \text{SUB1} \) has been identified.}

\begin{tabular}{|c|c|c|c|}
\hline
\multicolumn{2}{|c|}{Customers} & \multicolumn{3}{c|}{period} \\
\hline
\multicolumn{2}{|c|}{in set \( FD \)} & \( t+1 \) & \( t+2 \) & \( t+3 \) & \( t+4 \) \\
\hline
1 & & & & & \\
\hline
2 & & & & & \\
\hline
3 & & & & & \\
\hline
4 & & & & & \\
\hline
\end{tabular}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig2}
\caption{Graphical illustration of subproblem \( \text{SUB2} \) for a sample case}
\end{figure}

\text{Subproblem \( \text{SUB2} \) can be illustrated graphically. Consider the sample case for \( \text{SUB2} \) illustrated in Fig. 2. The decision variables } u_{it} \text{ are represented by directed arcs, where the cost saving associated with each arc } S_{it} = \text{ETC}_i(\Omega_i) - (\tau - t)h_i d_{it}. \text{ A solid vertical line is drawn to represent the time limit } T_{max}^i \text{ for customer } i. \text{ Starting from node 0, arcs are to be selected using the order given by their directions, such that the total cost saving is maximized and the vehicle capacity constraint is satisfied. We note here that if one or more arcs in a given period are selected, the saving values } S_{it} \text{ of the unselected arcs in the same period will be changed due to changes in the transportation cost estimates and therefore have to be recalculated.}

\text{1598}
Inspired by this graphical representation, subproblem SUB2 can be dealt with as a precedence constrained knapsack problem (PCKP) in which the coefficients of the objective function, \( S_{i \rightarrow t} \), are dependent on the decision variables. The PCKP is known to be NP-hard [31]; however, Johnson and Niemi [32] provide a dynamic programming algorithm for the PCKP that can solve the problem in a pseudo-polynomial time, given that the underlying precedence graph is a tree, which is fortunately a property of SUB2 as can be seen in Fig. 2.

We present here a simpler algorithm based on a greedy search that selects the next possible arc (see Fig. 2) that has the maximum positive saving. This algorithm does not guarantee optimality to the solution of SUB2; however, it can produce relatively good solutions in polynomial time. The following steps describe the algorithm.

Procedure SUBALG2

1. Let \( D^{\text{max}} = Q \) and \( TD = FD \);
2. For every customer \( i \) in set \( TD \), Let \( \Delta t_i = 1 \);
3. Find customer \( j \) in set \( TD \) that has the largest positive value of 
   \[ ETC_j (\Omega_{t+\Delta t_j}) - \Delta t_j h_j d_{j,t+\Delta t_j} \]; 
   If none found then terminate;
4. If \( D^{\text{max}} \geq d_{j,t+\Delta t_j} \) then 
   Let \( D^{\text{max}} = D^{\text{max}} - d_{j,t+\Delta t_j} \);
   Add \( d_{j,t+\Delta t_j} \) to customer \( j \)'s delivery amount and update 
   transportation cost estimates in period \( t+\Delta t_j \);
   Let \( \Delta t_j = \Delta t_j + 1 \);
   If \( \Delta t_j > T_j^{\text{max}} \) then remove customer \( j \) from set \( TD \);
   End-If
   Else remove customer \( j \) from set \( TD \);
5. If \( TD = \emptyset \) then terminate; Else go to step 3.

5. EXPERIMENTATION AND RESULTS

Two versions of ETCH have been implemented. In the first one, optimal solutions for the two subproblems are generated using a complete breadth-first search for SUB1 and a dynamic programming algorithm for SUB2. We refer to this implementation as ETCH-O. The second version uses the provided breadth-first heuristic for SUB1 and the greedy-search algorithm for SUB2, and is referred to as ETCH-H. These heuristics are programmed and compiled using Borland C++ Builder version 3 and benchmarked against the lower and upper bounds obtained by AMPL-CPLEX 8.1 with a specified execution time limit under an Intel Pentium 4 processor running with a clock speed of 2.40 GHz with 1GB RAM.

5.1. Experimental design

We consider two different scenarios to examine the effectiveness of the developed heuristics under different circumstances. These scenarios simulate the integrated inventory-distribution decisions faced by manufacturing companies that deal with small number of customers, each
located in a different major city. An example for similar cases in the literature can be found in [33].

The first scenario is designed to test the quality of the inventory holding decisions of ETCH; while, in the second one, some parameters are tuned to provide conditions in which backorder decisions are economical, so that the backorder decisions of ETCH are assessed. The main factors that are controlled to produce such cases are the ratio of the available vehicle capacity to the average daily demand by customers, the average unit shortage cost and the transportation cost per unit distance.

In both scenarios, customers are allocated in a square of 20×20 distance units and their coordinates are generated using a uniform distribution within these limits. The depot is located in the middle of the square. Customers’ unit holding costs are generated using a normal distribution with a mean of 0.1 and a standard deviation of 0.02, and each customer has a storage capacity of 120 items.

In the first scenario, the transportation cost per unit distance is set to 1, the customers’ unit shortage costs are generated using a normal distribution with a mean of 5 and a standard deviation of 0.5, and the customers’ demands are generated using a uniform distribution from 25 to 50 items per day. In the second scenario, we set the parameter values so it is optimal to carry backorders. In this scenario, the transportation cost per unit distance is set to 2, the customers’ unit shortage costs are generated using a normal distribution with a mean of 3 and a standard deviation of 0.5, and the customers’ demands are generated using a uniform distribution from 5 to 50 items per day.

For each scenario, sixty problems have been generated by varying the number of customers (N), the number of planning periods (T) and the number of homogenous vehicles (V). We generate three levels of N (5, 10 and 15), two levels of T (5 and 7), and two levels of V (1 and 2). For each problem setting defined by a combination of N, T, and V, we randomly generate five problems. The total vehicle capacity in the first scenario is selected to be fixed at 500, 1000, and 1500 for each level of N, respectively. In the second scenario, the selected total vehicle capacities are 150, 300, and 450.

The naming convention used for the test problems starts with a number that refers to the scenario. After a hyphen, two digits are assigned for the number of customers, followed by a digit representing the length of the planning horizon. The next digit represents the number of vehicles. Finally, the replicate number is given at the last digit after a hyphen. Thus, the problem 1-0551-1 represents the first run of the first scenario with 5 customers, a planning horizon of 5 periods and 1 vehicle.

5.2. Results and discussion

The detailed experimental results are not included in this paper for the sake of brevity. The percentage differences between the total cost obtained by each heuristic and the lower bound are used as performance indicators. The percentage difference, also referred to as optimality gap, is calculated by taking the ratio of the difference between the heuristic’s total cost and the lower bound to the lower bound. A comparison against the lower bound provides a measure of deviation from optimality. The CPLEX upper bound in a maximum of one-hour
running time is used as an alternate heuristic and its percentage difference against the lower bound is similarly calculated.

The results of the first scenario problems, as shown in Fig. 3, indicate that the total cost obtained by ETCH-O is less than that of ETCH-H in most of the cases. ETCH-O generates the lower cost by generally reducing the transportation costs with a slight increase in inventory holding costs as compared to ETCH-H. This discrepancy in the cost elements of both versions of the constructive heuristic is attributed to the search technique used for solving the second subproblem. The dynamic programming part of ETCH-O allows for investigating larger number of combinations for allocating the available vehicle capacity to future demand coverage. Consequently, ETCH-O investigates additional solution alternatives that may have lower total cost for the second subproblem. However, solving the second subproblem optimally does not guarantee generating lower total cost to the whole problem due to the myopic nature of the decisions made for the second subproblem. This explains why ETCH-H can generate solutions in a few instances with lower total cost. With a similar argument, the results obtained for the second scenario problems show that solving the first subproblem optimally does not guarantee generating lower cost for the whole problem. However, on average, ETCH-O is capable of generating solutions with lower total cost. We note that these general findings do not change with varying problem parameters of $N$, $T$, or $V$ for both scenarios. An ANOVA analysis showed no effect of $N$, $T$, or $V$ on the optimality gaps for the two heuristics (results of the ANOVA analysis not included in the paper for brevity).

We next investigate the impact of the solution quality as a function of the problem size in terms of the number of binary variables. That is, for each problem setting, the average of the percentage differences of the 5 replicates is calculated and plotted against the number of binary variables of that setting as shown in Figs. 3 and 4. Table 1 summarizes the average computational time of the developed heuristics for each problem set in both scenarios.

As shown in Figs. 3 and 4, the constructive heuristic outperforms the CPLEX upper bound for instances with more than 15 customers in both scenarios. While the growth of the CPLEX optimality gap is steady with the increase of the number of binary variables, the optimality gap for the developed heuristics is below 30% on average and remains almost level with the increase of the number of binary variables. The ETCH-O version of the constructive heuristic is on average 2% closer to the lower bound than ETCH-H.

The computational time for ETCH-H is found to be less than one second in all the tested cases. For ETCH-O, due to the dynamic programming part of the algorithm, the computational time increases with the increase of the problem size; however on average, it has not reached the 90 seconds limit in all problem sets. The increase of computational time of ETCH-O is mostly attributed to the increase in both $N$ and $T$; while, the number of vehicles, $V$, does not seem to have a significant effect on computational time.
A CONSTRUCTIVE HEURISTIC FOR THE INTEGRATED INVENTORY-DISTRIBUTION PROBLEM

Fig. 3. Average percentage differences against lower bounds for the first scenario problems

Fig. 4. Average percentage differences against lower bounds for the second scenario problems
Table 1. Average computational times (in minutes) for the developed heuristics

<table>
<thead>
<tr>
<th>N</th>
<th>T</th>
<th>V</th>
<th># binary variables</th>
<th>First scenario</th>
<th>Second scenario</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ETCH-O</td>
<td>ETCH-H</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ETCH-O</td>
<td>ETCH-H</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>1</td>
<td>150</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>2</td>
<td>150</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>2</td>
<td>210</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>2</td>
<td>210</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>1</td>
<td>300</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>15</td>
<td>5</td>
<td>2</td>
<td>300</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>15</td>
<td>7</td>
<td>2</td>
<td>300</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

From the previous results we conclude the following. The ETCH-O version of the constructive heuristic is capable of generating slightly better solutions compared to ETCH-H with up to 2% difference on average in the optimality gap. However, with the increase of the problem size, especially the number of customers and the number of planning periods, the computational time of ETCH-O will be significantly higher than the computational time of ETCH-H.

To investigate the performance of the developed heuristics with larger problem sizes, we construct an additional experimental set based on a third scenario. In this scenario, medium vehicle capacity to average daily demand ratio is used such that a situation in the middle of the first two extreme scenarios is addressed. This scenario considers similar parameters as in the second one with some modifications to reduce the frequency in which backorder decisions are needed. The main difference between the parameters used in the third scenario as compared to the second one is that the travel cost per unit distance is set to 1 and the customers daily demand is generated using a uniform distribution between 0 and 25. We consider three different levels of the number of customers, \( N \): 20, 25 and 30, and a total vehicle capacity of 300, 350 and 400 at each level of \( N \) respectively. We only consider one level for both \( T \) and \( V \) at 7 and 2 respectively. Five random replicates are generated at each level of \( N \). We use the previously defined naming convention for the third scenario problems.

For the third scenario problems, CPLEX lower and upper bounds are obtained after a running time of three hours. Due to the inability of the optimization routines to find solutions for the two subproblems for these large problem instances, we only ran the heuristic ETCH-H. The average cost and time results for the ETCH-H version are shown in table 2.

We can see that the rate of increase of the heuristic's optimality gap is almost constant with the increase of the number of customers. When we compare this with the rapid rate of increase for the CPLEX upper bound percentage difference, we can see the potential benefit of the developed constructive heuristic for larger problem sizes. In terms of computational time, the ETCH-H version of the constructive heuristic remains below one second for larger problems with up to 30 customers.
A CONSTRUCTIVE HEURISTIC FOR THE INTEGRATED INVENTORY-DISTRIBUTION PROBLEM

Table 2. Average results for the third scenario problems

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th># binary variables</th>
<th>CPLEX UB LB diff %</th>
<th>ETCH-H LB diff %</th>
<th>Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>7</td>
<td>2</td>
<td>5880</td>
<td>75.91</td>
<td>34.16</td>
<td>0.00</td>
</tr>
<tr>
<td>25</td>
<td>7</td>
<td>2</td>
<td>9100</td>
<td>126.39</td>
<td>37.10</td>
<td>0.00</td>
</tr>
<tr>
<td>30</td>
<td>7</td>
<td>2</td>
<td>13020</td>
<td>200.84</td>
<td>39.83</td>
<td>0.01</td>
</tr>
</tbody>
</table>

It is important to note that the above experimental results are representing the specific algorithmic implementations that are based on the Savings algorithm for providing solutions to the VRP subproblem. More efficient algorithms for the VRP may provide better solutions as a result of generating lower transportation costs. Investigating this issue is beyond the scope of this paper.

Regarding the CPLEX runs, it was found that the "presolve" heuristics that are applied at the preparation stage to generate efficient cuts before applying the branch and bound algorithm play an important role in generating efficient lower and upper bounds. However, these bounds improve with a very slow rate as the branch and bound algorithm proceeds. It was found that runs longer than the used time limits do not improve the bounds significantly.

6. CONCLUSION AND FUTURE WORK

This article addressed the integrated inventory distribution problem in which multiperiod vehicle routing and inventory holding and backlogging decisions for a set of customers are to be made. We considered an environment in which the demand at each customer is relatively small compared to the vehicle capacity, and the customers are closely located such that a consolidated shipping strategy is appropriate. We presented a constructive heuristic based on the idea of allocating single transportation cost estimates for each customer. Two subproblems, comparing inventory holding and backlogging decisions with these transportation cost estimates, are formulated and their solution methods are incorporated in the developed heuristic. The main idea behind the constructive heuristic as seen in the formulation of the two subproblems is to consider only delivery plans in which fulfillment of part of the current or the future demand requirements in a currently studied period is not allowed. A mixed integer programming formulation is provided and used to obtain lower and upper bounds using AMPL-CPLEX to assess the performance of the developed heuristics.

For small sized problems with up to 15 customers, the experimental results show that the developed constructive heuristic can achieve solutions that are on average not farther than 30% from the optimal in a few minutes. With the increase of problem size, the optimality gap of the developed heuristics increases with almost a constant rate and results can be obtained in a few minutes. This shows the potential benefit of the developed heuristics for larger problem sizes.

Future research may consider building upon the algorithmic ideas developed in this paper to solve other problems in the manufacturing industry that have wider scope in the supply chain. The integration of production and inventory-distribution decisions at the operational planning level is a good example of one such problem.
REFERENCES