A two-echelon inventory model for a deteriorating item with stock-dependent demand, partial backlogging and capacity constraints

Yousef Ghiami\textsuperscript{a,}\textsuperscript{*}, Terry Williams\textsuperscript{b}, Yue Wu\textsuperscript{a}

\textsuperscript{a} CORMSIS, School of Management, University of Southampton, Southampton SO17 1BJ, United Kingdom
\textsuperscript{b} Business School, Hull University, Hull HU6 7RX, United Kingdom

\textbf{Abstract}

This study investigates a two-echelon supply chain model for deteriorating inventory in which the retailer’s warehouse has a limited capacity. The system includes one wholesaler and one retailer and aims to minimise the total cost. The demand rate in retailer is stock-dependent and in case of any shortages, the demand is partially backlogged. The warehouse capacity in the retailer (OW) is limited; therefore the retailer can rent a warehouse (RW) if needed with a higher cost compared to OW. The optimisation is done from both the wholesaler’s and retailer’s perspectives simultaneously. In order to solve the problem a genetic algorithm is devised. After developing a heuristic a numerical example together with sensitivity analysis are presented. Finally, some recommendations for future research are presented.

\section{Introduction}

In the classical inventory model for deteriorating products it is usually assumed that the warehouse has no limits in the capacity. However, in the real-life problem the situation is different. There are a number of factors which influence the optimum solution in different ways. Sometimes these factors may suggest retailers to buy more than their own warehouse (OW) capacity. In these situations, the retailers can benefit from a rented warehouse (RW).

Another assumption that can greatly influence the optimal policies is to take a supply chain perspective when analysing inventory models. In multi-echelon inventory models, actors try to integrate their businesses in order to improve the overall performance of the system (e.g. by increasing service level, profit of decreasing the cost). Implementing such integrated models however remains challenging especially when the actors are independent businesses and should collaborate very closely (see Fawcett & Magnan, 2002; Power, 2005). In such cases apart from close collaboration between players in the supply chain, there should be a fair mechanism to distribute the incentives between the actors to encourage the integration. Prajogo and Olhager (2012) argue that establishing any mechanisms for supply chain integration is only possible if there is a long-term relationship between the supply chain partners.

To date, very few studies on deteriorating inventory in two-echelon systems have been carried out (see Bakker, Riezebos, & Teunter, 2012; Goyal & Giri, 2001; Li, Lan, & Mawhinney, 2010; Nahmias, 1982; Raafat, 1991). Considering the general product and deteriorating items, it can be said that deteriorating product literature is still in its infancy and has a long way to go compared to the maturity acquired by general product supply chain literature.

As Cohen (1976) notes in his research, Zyl (1964) is one of the first researchers who addresses deteriorating inventory. Zyl (1964) considers a perishable product with fixed lifetime. Cohen (1976) develops a model for a deteriorating item with m-period lifetime. In this single echelon model, the demand is stochastic and any shortages are completely backlogged. Mak (1982) considers the same set of assumptions with constant rate of demand. The researcher chooses continuous variables and analyses the system using differential equations, which is the most popular analytical approach in analysing deteriorating product’s models.

When working on models for deteriorating items, researchers consider specific factors on which they make assumptions. The main factors that are considered by researchers in developing new models are namely, demand pattern, lead-time, deterioration rate, shortages, supply chain structure, reviewing policy, system type (inventory vs. production/inventory) and warehouse capacity.

The definition of the deteriorating item includes quite a wide spectrum of products such as food, fruit, blood, flower, medicine and clothes. The demand for these deteriorating items therefore varies depending on the product characteristics and the consumption pattern followed by customers. A large group of researchers consider a constant rate for demand. In these studies “D” units of the product is demanded per unit of time. Mak (1982), Wee (1993), Wee (1998), Chung, Liu, and Tsai (1997), Abad (2000), Yang and Wee (2000), Rau, Wu, and Wee (2003), Yang (2004a), Dye,

Bhunia and Maiti (1998), Chung and Tsai (2001), Moon, Giri, and Ko (2005), Yang (2005) and Lee and Hsu (2009) assume the demand to be a time-dependent function. Examples for time-dependent demand applications can be seen in grocery retailing industry that some products’ demand rate varies in different weekdays. Another example is clothing industry in which seasonality changes the demand level in different seasons. Wu and Ouyang (2000) and Manna and Chaudhuri (2006) develop models with ramp-type demand which is a combination of constant rate and time-dependent demand. One example for this demand pattern is when a new consumer good is introduced to the market. The demand for this new product (in case of success) goes up as the time passes and finally it converges to specific level and becomes constant (Wu & Ouyang, 2000).

Some products are price elastic which means by changing the price of the product the demand will change, however the elasticity may vary from one product group to another. Yang (2004b) and Dye (2007) explore systems with price-dependent demand. Changes in demand in different periods mean that retailers should follow the changes in demand by storing inventory or changing inventory policies frequently which results in large amount of cost in their supply chain if the demand fluctuations are large. Hence, they strive to benefit from the elasticity of demand and influence the demand pattern to minimise their supply chain cost.

Hou (2006) discusses that for some products such as consumer goods the demand may be influenced by the amount of the product displayed on shelves. With this regard a group of researchers have considered stock-dependent demand in their research such as Mandal and Phayujar (1989), Giri, Pal, Goswami, and Chaudhuri (1996) and Hou (2006). Similar to the case of price-dependent demand, retailers use this characteristic to influence the demand rate for a product by displaying a large amount of that item on shelves. Therefore they can decrease their inventory cost and increase their sale on the product.

Researchers make different assumptions regarding policies on shortages when considering the product and market characteristics. Liao (2007) and Yang, Teng, and Chern (2010) develop models where shortages are not allowed. This is a critical assumption when developing a blood bank model or optimising a distribution system for a group of pharmaceuticals. In this type of models, service level is the objective function rather than cost and benefit as these products are connected to health issues.

For some products when a retailer is out of stock, the demand is lost which means the customer finds the item or a similar one in another store. Hsu (2000) and Lu, Song, and Zhu (2008) study models in which any shortages are lost. This case may happen when there are similar products in the market and differences are negligible such as milk or bread that can be found in every grocery shop.

Sometimes customers tend to wait or have to wait for inventory replenishment in case of shortages. The reason for this can be the specific characteristic or the outstanding quality of the product for instance an especial type of cheese which only can be found in a specific shop. Some researchers such as Yang (2004a) and Olson and Tydesjö (2010) study systems with shortages and consider that any unmet demand is backlogged. Yang (2005), Law and Wee (2006) and Dye (2007) challenge the last two groups of research and discuss that most of the time the situation lies somewhere between where the unsatisfied demand is partially backlogged and partially lost.

When analysing inventory models for deteriorating items, researchers have mainly considered single-echelon systems and have strived to optimise models from a single business point of view. In recent years, however, the supply chain perspective for deteriorating items has gained more attention. Even though developing logistics models with a supply chain perspective makes the models more realistic, implementing these models is complicated and requires collaboration between the supply chain members. Yang and Wee (2000) and Yang et al. (2010) develop an inventory model for a two-echelon system with constant rate for demand and deterioration in which shortages are not allowed. Yang et al. (2010) develop a similar model with a difference in input pattern as the model is a production/inventory system. Yang (2004a) studies a two-echelon inventory model, similar to the above mentioned models, in which the demand is a function of price and the lead-time is constant. Another important feature of this study is the influence of the time value of money which is taken into account. Law and Wee (2006) and Lo, Wei, and Huang (2007) investigate a two-echelon system with partial backlogging, two-parameter deterioration rate and constant rate of demand. The former research is conducted allowing for permissible delays in payment while the latter considers inflation rate in its calculations. Zanoni and Zavanella (2007) study a two-echelon system for determining optimal inventory policies with constant rate of demand and no shortages. In this model the perishable product has a constant lifetime.

In most of the studies, researchers consider no limit for the capacity of a warehouse. This however, can be one of the most important issues in a real-world problem. For the first time, Sarma (1987) developed a model while assuming a limited warehouse capacity. In this model extra capacity can be obtained by renting warehouse (RW) as the owned warehouse (OW) has limited capacity. Yang (2004a) studies an inventory system with limited capacity. The demand rate is constant and any shortages are completely backlogged. The item’s lifetime is based on an exponential distribution (constant rate of in-hand inventory is deteriorated). In another study, Yang (2006) develops a similar model with partial backlogging. Wee, Yu, and Law (2005) investigate a model with partial backlogging in which the product lifetime is based on the two-parameter Weibull distribution. Dal, Das, Panda, and Bhunia (2005) explore a deteriorating item with a two-warehouse system in which the demand is time-dependent and the demand is partially backlogged in case of shortages. Lee (2006) develops a production-inventory model in which the inventory level is increased by a finite production rate. The other assumptions in this model are similar to the research done by Yang (2004a). Chung and Huang (2007) study an inventory model with no shortages, and permissible delays in payments are assumed. Hsieh, Dye, and Ouyang (2008) study an inventory system similar to Yang (2004a) and optimise the model taking net present value approach. Rong, Mahapatra, and Maiti (2008) consider an inventory model with fuzzy lead-time and complete backlogging in which the demand is connected to price. Singh, Kumar, and Kumari (2009) explore a two-warehouse model with complete backlogging and time-dependent demand. This model is analysed and optimised considering permissible delays in payments. Lee and Hsu (2009) develop a production-inventory model with time-dependent demand and no shortages. Gayen and Pal (2009) analyse an inventory model in which shortages are not allowed and the demand is stock-dependent. Liao and Huang (2010) investigate a similar model to Chung and Huang (2007) by adopting a different approach. One can find many examples where a retailer needs to rent a warehouse. A new business on high street where space is very expensive can be one example. In such case the retailer uses the space to display the items and moves the warehousing processes to less costly areas. Only after securing a high demand level, the retailer would think about “owning a warehouse” as a strategic decision if feasible.

Almost in all the deterministic studies done on deteriorating items, the objective function is to either maximise the profit or minimise the cost. In stochastic models however there are models...
with an objective function on service level. For a more thorough survey of deteriorating item models see Nahmias (1982), Raafat (1991), Goyal and Giri (2001), Li et al. (2010) and Bakker et al. (2012).

All the two-warehouse inventory models are based on a single company’s point of view and they fall short of taking an overall supply chain perspective. In this study we consider a two-echelon system consisting of a wholesaler and a retailer in which there is limit in the retailer’s warehouse capacity. The demand is considered to be stock-dependent and the product is deteriorated with a constant rate. An analytical model is developed based on the above mentioned assumptions and the costs incurred by retailer and wholesaler are analysed. Using this optimisation model, the optimal inventory policies for both actors in this two-echelon system are obtained. In order to solve the problem a heuristic method is developed.

2. Assumptions and notation

2.1. Assumptions

In this paper we consider a two-echelon system which delivers a perishable item to the end customer. This system consists of one wholesaler and one retailer. The perishable product is supplied to the retailer only by this wholesaler and in the same way this retailer is the only downstream actor to which the perishable product can be sent. The product has exponential life time which means there is a constant rate for deterioration. The lead time for both retailer and wholesaler is zero. Shortages are allowed at the retailing level and are partially backlogged while this is not allowed at the wholesaler level. The retailer has a capacity of \( W \) in its own warehouse (OW) which is limited; therefore, if the order quantity exceeds this capacity, it is suggested that the retailer rents a temporary warehouse (RW) which has higher carrying costs and unlimited capacity. In the case that RW is used at the retailer level, the consumption of goods from OW starts only after the inventory in RW is depleted at time \( t_r \). The inventory level in OW reaches zero at time \( t_o \), and from this time to the end of the inventory period \( (t_o) \) shortages occur and are partially backlogged until the next replenishment. Warehouse capacity is not an issue for the wholesaler. The deteriorated items cannot be repaired or replaced. This model is an inventory system with fixed inventory period and order quantity and the optimal solution will specify how much and how often the members of this supply chain should order. The total cost of the system consists of purchasing, holding and deterioration cost for both members and shortage cost for the retailer. Purchasing cost for both members is a linear function of their order quantity including the replenishment cost. The unit deterioration cost and the unit holding cost per unit of time are constant. The unit lost sale cost and the unit shortage cost per unit of time for backlogged demand are also constant.

2.2. Notations

In order to develop a mathematical model based on the above mentioned assumptions, the following notations are adopted. The time unit can be either a day or a week, but it is important to consider all the time-based costs properly.

\[
\begin{align*}
D(t) & \quad \text{the demand rate at time } t, \text{ the demand is assumed to be deterministic and stock-dependent:} \\
D(t) &= c_d(t) + d \\
W & \quad \text{the capacity of the OW} \\
L(t) & \quad \text{the inventory level at the OW at time } t \\
D_t(t) & \quad \text{the inventory level at the OW at time } t \\
S(t) & \quad \text{the shortage level at the retailer} \\
L_w(t) & \quad \text{the inventory level at the wholesaler} \\
D_w(t) & \quad \text{the inventory level at the wholesaler during } \text{ith} \text{ interval of its inventory period} \\
\alpha & \quad \text{the deterioration rate at the OW} \\
\beta & \quad \text{the deterioration rate at the RW (} \beta > \alpha \text{)} \\
\gamma & \quad \text{the deterioration rate at the wholesaler} \\
\delta & \quad \text{the percentage of demand which is backlogged during shortage time} \\
A_R & \quad \text{the replenishment fixed cost per order for the retailer} \\
A_W & \quad \text{the replenishment fixed cost per order for the wholesaler} \\
P_R & \quad \text{the purchasing price for the retailer per unit} \\
P_W & \quad \text{the purchasing price for the wholesaler per unit} \\
h_o & \quad \text{the unit holding cost per unit of time at the OW} \\
h_r & \quad \text{the unit holding cost per unit of time at the RW} \\
h_w & \quad \text{the unit holding cost per unit of time at the wholesaler} \\
c_{sf} & \quad \text{the unit lost sale cost} \\
c_{sw} & \quad \text{the unit lost sale cost per unit of time for backlogged demand} \\
cas & \quad \text{the time at which the inventory level at the OW reaches zero} \\
t_r & \quad \text{the time at which the inventory level at the RW reaches zero} \\
t_o & \quad \text{the time during which the retailer is out of stock and demand is partially backlogged} \\
T_R & \quad \text{the inventory period at the retailer} \\
Q_R & \quad \text{the order quantity at the retailer} \\
T_W & \quad \text{the inventory period at the wholesaler} \\
Q_W & \quad \text{the order quantity at the wholesaler}
\end{align*}
\]

3. The model

3.1. Inventory level at the retailer (OW and RW)

Considering the model, in case of having inventory at both the OW and the RW, the retailer is willing to use the inventory at the RW first. Let \( l_d(t) \) and \( l_r(t) \) be the inventory level at the OW and the RW respectively. At the RW the inventory is depleted by a demand which is connected to the inventory level at the OW. Therefore the changes of inventory level at the RW between the start of the inventory period and \( t_r \) can be presented by the following differential equations:

\[
\frac{dl_r(t)}{dt} = -c_d(t) - \beta l_r(t), \quad 0 \leq t \leq t_r
\]  

(1)

While the retailer is using the inventory at the RW to meet the demand, the inventory level at the OW goes down by a constant rate of the inventory level due to deterioration as shown in (2).

\[
\frac{dl_o(t)}{dt} = -\alpha l_o(t), \quad 0 \leq t \leq t_r
\]  

(2)

At time \( t_r \), the inventory at the RW is depleted completely and the inventory at the OW is used. The inventory level at the OW decreases due to the demand and deterioration until it reaches zero at \( t_o \). This change in the inventory level is presented by the differential equations in (3).

\[
\frac{dl_o(t)}{dt} = -c_d(t) - \alpha l_o(t), \quad t_r \leq t \leq t_o
\]  

(3)
From $t_0$ to $T_k$ the system is out of stock and unmet demand is partially backlogged (4). All the changes of inventory level at the RW and the OW are depicted graphically in Fig. 1.

$$\frac{dl_k(t)}{dt} = -\delta d, \quad t_0 \leq t \leq T_k$$  \hfill (4)

In order to solve the presented differential equations, the following boundary conditions should be considered:

$I_1(0) = W, \quad I_1(t_f) = 0, \quad I_1(t_j) = 0$

By solving the differential equations in (1)–(4), the inventory levels at the OW and the RW are obtained:

$I_1(t) = \frac{cW e^{\frac{-tr}{\beta + \alpha}} - d e^{\frac{tr}{\beta + \alpha}} - 1}{\beta} + d e^{\frac{tr}{\beta + \alpha}} - 1, \quad 0 \leq t \leq t_r$  \hfill (5)

$I_1(t) = We^{-\frac{tr}{\beta + \alpha}}, \quad 0 \leq t \leq t_r$  \hfill (6)

$I_2(t) = \frac{1}{c + \alpha} (1 - e^{(c + \alpha)(t_0 - t)}), \quad t_r \leq t \leq t_0$  \hfill (7)

$S(t) = -\delta(t - t_0), \quad t_0 \leq t \leq T_k$  \hfill (8)

The inventory level at the OW at $t = t_r$ can be calculated from both (6) and (7). This equality results in (9) which shows $t_0$ as a function of $t_r$:

$t_0 = t_r + \frac{1}{c + \alpha} \ln \left(1 + \frac{c + \alpha}{d} We^{-\frac{tr}{\beta + \alpha}}\right)$  \hfill (9)

The order quantity for the retailer is the sum of the initial inventory level at the RW and the OW and the total backlogged demand during one inventory period. As can be seen in (10), $Q_{rk}$ is a function of $t_0$ and $t_r$.

$Q_{rk} = I_1(0) + I_1(t_0) + S(T_k)$

$$= \frac{cW}{\beta + \alpha} \left(e^{(\beta - \alpha)t_r} - 1\right) + \frac{d}{\beta} (e^{\beta t_r} - 1) + W + \delta t_s$$  \hfill (10)

The length of the inventory period at the retailer is the sum of $t_0$ and $t_r$. Based on (9), $t_0$ is a function of $t_r$. Therefore, $T_k$ is obtained as a function of $t_r$ and $t_0$:

$T_k = t_0 + t_r + \frac{1}{c + \alpha} \ln \left(1 + \frac{c + \alpha}{d} We^{-\frac{tr}{\beta + \alpha}}\right) + t_r$  \hfill (11)

As can be seen in (10) and (11), inventory policies for the retailer ($Q_{rk}, T_k$) are functions of $t_r$ and $t_0$.

3.2. Inventory level at the wholesaler

The inventory policies for the wholesaler are $Q_{W}$ and $T_{W}$. If we assume $T_{W}$ to be a multiplication of $T_k (T_{W} = kT_k)$, it can be proven that $k$ is an integer. If $k$ is not an integer, it means that the wholesaler receives a new batch in their warehouse while there is no demand for new replenishment from the retailer’s side, as the retailer has not passed the inventory period yet. This means for that fraction of $T_k$, the wholesaler is carrying inventory and bearing deterioration costs which are not needed until the end of the retailer’s inventory period. Thus, $k$ should be an integer.

The order quantity for the wholesaler is equal to the inventory needed for $k$ periods at the retailer, plus the amount of deterioration during the wholesaler’s inventory cycle. It should be noted that during the $k$th interval, there is no inventory in the wholesaler and, after receiving $Q_k$ of product at the end of $T_{W}$, $Q_k$ is sent to the retailer. Therefore there is no deterioration during this interval in the wholesaler. The order quantity for the wholesaler can be calculated as follows:

$Q_{W} = kQ_{rk} + $ Deterioration during the wholesaler’s inventory cycle

$$= k \left(Q_{rk} + \delta t_s\right)$$  \hfill (12)

As discussed before, one inventory period in the wholesaler consists of $k$ retailer’s inventory periods. Fig. 2 illustrates the inventory level in the wholesaler. At the time $(k-2)T_k$ and $(k-1)T_k$, the inventory level at the wholesaler drops by $Q_k$ and a constant rate of the inventory is deteriorated during the interval $[(k-2)T_k, (k-1)T_k]$. The change in inventory level at the wholesaler during this interval can be presented by the following differential equation:

$$\frac{dl_k(W)}{dt} = -\gamma l_k(W)$$  \hfill (13)

Considering the inventory level at the wholesaler at $(k-1)T_k$ which is $Q_k$, the inventory level for the specified period will be:

$I_k(W) = Q_k e^{\frac{-\gamma (k-1)T_k - t}{c + \alpha}}, \quad (k-2)T_k \leq t \leq (k-1)T_k$  \hfill (14)

In a similar way, the inventory level at the wholesaler can be obtained for the period starts at $(k-2)T_k$ using the differential equations presented in (13), considering the boundary condition derived from (14) at $t = (k-2)T_k$:

$I_k(W) = Q_k e^{\frac{\gamma (k-2)T_k - t}{c + \alpha}}, \quad (k-3)T_k \leq t \leq (k-2)T_k$  \hfill (15)

The inventory level at the wholesaler during ith interval can be calculated as follows:

$$l_k(W) = Q_k \left(\sum_{i=1}^{k} e^{\frac{-\gamma (k-i)T_k}{c + \alpha}}\right) e^{\frac{\gamma (k-i)T_k - t}{c + \alpha}}, \quad i = 1, 2, \ldots, k$$  \hfill (16)

3.3. Costs at the retailer

The retailer has four types of cost: purchasing, carrying, deterioration and shortage costs. The purchasing cost for the retailer is:

$PC_{rk} = A_k + p_b Q_{rk}$  \hfill (17)

Carrying cost in the retailer consists of two different elements; cost at the OW and at the RW.

Carrying cost at the RW = $ICC_{RW} = h_r \int_{0}^{t_r} I_1(t) dt$

$$= h_r \frac{cW e^{\frac{-tr}{\beta + \alpha}} - d e^{\frac{tr}{\beta + \alpha}} - 1}{\beta} + d e^{\frac{tr}{\beta + \alpha}} - 1 - t_r$$

Carrying cost at the OW = $ICC_{OW} = h_r \int_{t_r}^{t_f} I_1(t) dt + h_r \int_{t_f}^{t_k} I_1(t) dt$

$$= \frac{h_r W}{\beta} (1 - e^{\frac{tr}{\beta + \alpha}}) + \frac{h_r d}{\beta} \left(e^{\frac{tr}{\beta + \alpha}} - 1\right) - \left(t_0 - t_r\right) - t_r$$

Fig. 1. Graphical presentation of the inventory level at the RW and the OW at the retailer.
Based on the calculation done, total carrying cost at the retailer during one inventory period is:

$$IC_C = ICCRW + ICCOW$$

(18)

Deterioration cost at the retailer includes both deterioration at the RW and the OW:

$$DC_R = DC_RW + DC_OW$$

(19)

Considering inventory levels at the RW and the OW in (5)–(7), deterioration cost is as follows:

$$DC_RW = p_W(e^{\mu t} - 1) \left[ \frac{CW e^{-\mu t}}{\beta - \alpha} + \frac{\beta C W}{\alpha (\beta - \alpha)} (e^{-\mu t} - 1) - dt, \right.$$ \n$$DC_OW = p_C(e^{\mu t} - 1) \left[ \frac{CW e^{-\mu t}}{C + \alpha} + \frac{\alpha C W}{(C + \alpha) (\beta - \alpha)} (e^{-\mu t} - 1) - dt, \right.$$

During the shortage period, the demand is partially backlogged. There are two different types of shortage cost; one is based on per unit for the lost sale and the second is for backlogged demand which is per unit per unit of time.

$$SC_C = c_f \left[ \int_0^{t_1} (1 - \delta) dt + c_f \int_{t_1}^{t_2} \delta dt \right]$$

$$= c_f (1 - \delta) t_1 + \frac{1}{2} c_f \delta t_1^2$$

(20)

3.4. Costs at the wholesaler

There are three different types of cost at the wholesaler; namely purchasing, inventory carrying and deterioration costs. Considering the purchasing price for the wholesaler ($p_W$) and the order quantity ($Q_W$), the purchasing cost will be:

$$PC_W = A_W + p_W Q_W$$

(21)

Inventory carrying cost for the wholesaler during the $i$th interval is:

$$ICC_W = \frac{h_W Q_R}{\gamma} \left( e^{\gamma t_s} - 1 \right) \sum_{m=i+1}^{k} e^{\gamma(m-i) t_s}$$

Hence, carrying cost during one period is (consider that there is no carrying cost during $4$th interval):

$$ICC_W = \frac{h_W Q_R}{\gamma} \left( e^{\gamma t_s} - 1 \right) \sum_{m=i+1}^{k} e^{\gamma(m-i) t_s}$$

$$= \frac{h_W Q_R}{\gamma} \left( e^{\gamma t_s} - 1 \right) \left( e^{\gamma t_s} - 1 - k \right)$$

(22)

Based on (16), the amount of the deterioration in each interval can be calculated as follows:

$$Time period: [(k-1)T_k, kT_k] \quad Deterioration: 0$$

$$[(k-2)T_k, (k-1)T_k] \quad Q_R(e^{\gamma t_k} - 1)$$

$$[(k-3)T_k, (k-2)T_k] \quad Q_R(e^{\gamma t_k} - 1)(1 + e^{\gamma t_k})$$

$$[(k-4)T_k, (k-3)T_k] \quad Q_R(e^{\gamma t_k} - 1)(1 + e^{\gamma t_k} + e^{2\gamma t_k})$$

$$[(i-1)T_k, iT_k] \quad (i\text{th Interval}) \quad Q_R(e^{\gamma t_k} - 1) \left( \sum_{m=i+1}^{k} e^{\gamma(m-i) t_k} \right)$$

Considering deterioration in each interval, the total deterioration cost at the wholesaler during $T_w$ can be obtained:

$$DC_W = p_W Q_R(e^{\gamma t_k} - 1) \sum_{i=1}^{k} \sum_{m=i+1}^{k} e^{\gamma(m-i) t_k}$$

$$= p_W Q_R(e^{\gamma t_k} - 1) \left( \sum_{m=1}^{k} \frac{e^{\gamma m t_k}}{e^{\gamma t_k} - 1} - k \right)$$

(24)

Based on (12) and the calculation done in (23), the wholesaler’s order quantity is:

$$Q_W = kQ_R + Q_R(e^{\gamma t_k} - 1) \sum_{i=1}^{k} \sum_{m=i+1}^{k} e^{\gamma(m-i) t_k} = Q_R \left( \frac{e^{\gamma t_k} - 1}{e^{\gamma t_k} - 1} - k \right)$$

(25)

Considering the cost function of the wholesaler and the retailer, the total cost of the system per unit of time is as follows:

$$TC = \frac{1}{T_w} (PC_R + ICCRW + DC_RW + SC_R) + \frac{1}{T_w} (PC_W + ICCW + DC_W)$$

(26)

As mentioned before, $TC$ is a function of $t_r, t_i$ and $k$. Therefore the problem is:

Minimum $TC(t_r, t_i, k)$ subject to $t_r, t_i, k \geq 0, k \text{ is an integer}$

(27)

4. Optimisation

In order to solve problem (27) analytically, enumeration can be used as $k$ cannot take a very large value. In this case the problem can be solved for different values of $k$ (e.g., 1, 2, 30) and the best optimal solution and the corresponding $k$ can be considered as the optimal solution for the problem. Appendix A presents an analytical solution for this problem. This method however can result in an exhaustive search if $k$ tends to get large values. In studies with similar models and objective function in terms of complexity, researchers develop a heuristic to solve the problem to avoid time consuming solution process (see Pal et al., 2005; Yan, Banerjee, & Yang, 2011; Yang & Wee, 2002). In this research a heuristic is suggested that combines genetic algorithm (GA) and a neighbouring search which can search the feasible area, solve the problem in a short time and give a near-optimal solution. The steps of the GA are as follow:

Step 1. Deciding about parameters in GA: population, number of generations, the percentage of the next generation which should be generated by mutation, reinsertion and crossover and when to stop the algorithm

Step 2. Producing the first generation and calculating the fitness function (gen = 1)

Step 3. Saving the best solution in the population

Step 4. If the planned number of generations has been produced go to Step 9

Step 5. Using the previous generation to produce new generation by mutation, reinsertion and crossover and calculating the fitness function
Step 6. Saving the solution in the population
Step 7. In case of no improvement compare to the previous generation go to Step 9
Step 8. Go to Step 4
Step 9. Saving the optimal solution
Step 10. Stop

In order to implement this genetic algorithm similar to Maiti, Bhunia, and Maiti (2006) and Gupta, Bhunia, and Goyal (2007) the following components are considered:

- Parameters of genetic algorithm (Population size, maximum number of generations, the probabilities used in the genetic operations).
- Chromosome representation.
- Initial population.
- Fitness function.
- Selection process.
- Genetic operations (crossover, mutation and reinsertion).

4.1. Parameters of genetic algorithm

First all the parameters of genetic algorithm should be defined. These parameters are the population size (MPOP), maximum number of generations (MGEN), probability of reinsertion (PREIN), probability of mutation (PMUT) and probability of cross over (PCROS). In this research the values of the introduce parameters are as follow:

- MPOP = 200, MGEN = 200, PREIN = 0.1, PMUT = 0.15, PCROS = 0.75.

4.2. Chromosome representation

A three-dimensional vector, \( X = (t_1, t_2, k) \), is used to represent a person in the population (a solution). In this vector \( t_1 \) and \( t_2 \) are real numbers and \( k \) is an integer number.

4.3. Initial population

The first generation is generated by assigning random values to decision variables considering the bounds of the relevant decision variable. This process continues until the desired population is obtained. In this research for the first two decision variables a uniformly distributed number is used and the value for the third decision variable is chosen randomly from a set of integer values.

4.4. Fitness function

After obtaining the first generation and following generation, the quality of the solutions should be checked. The value of the objective function for each solution indicates the quality of the solution. In this research the total cost of the system corresponding to each chromosome is the objective function. In order to use the conventional selection process as in Maiti et al. (2006) and Maiti and Maiti (2007) there is a need for modification as the objective function for each solution indicates the quality of the solutions should be checked. The value of the objective function for each solution indicates the quality of the solutions. The value of the objective function for each solution indicates the quality of the solutions should be checked.

4.5. Selection process

Considering the fitness function, it is guaranteed that the solution with lower cost function has a higher chance to be selected. In this study the selection process used in Maiti and Maiti (2007) is considered.

4.6. Genetic operations

Solutions/persons in the second and the following generations are generated by the ones in the previous generation using three main operations; crossover, mutation and reinsertion. Genetic algorithm continues until one of the termination condition is held; a specific number of generations are produced or there is no improvement in the optimal solution from one generation to the next.

4.7. Crossover operation

The majority of the population in generations are produced by crossover operator (predetermined percentage). In this study one point crossover is used. After choosing the parents by selection process, crossover operation generates a random value which shows what part of these two solutions should be exchanged in order to have new solutions. The fitness function of the newly generated solutions is found. This process is repeated until reaching the percentage of the population for new generation.

4.8. Mutation operation

A small percentage of new generation is generated using mutation operation. After choosing a chromosome using the selection process, mutation randomly picks one of the decision variables in the chosen solution and assigns a new value to the decision variable within the bounds. In the next step the corresponding fitness function is found. These steps are repeated until the desired population generated by mutation operation is obtained.

4.9. Reinsertion operation

Reinsertion operation creates a very small number of solutions in the new generation. Reinsertion selects one of the solutions randomly using the selection process from the last generation and moves that to the new generation with no change.

Using GA with appropriate generating tools guarantees feasible area coverage but not optimality. Therefore in this study after producing the last generation, each individual in the last generation is studied to see if there is any local optimal in their neighbourhood. To do this, all the neighbour solutions to each individual are searched and evaluated. Each solution has three decision variables and a neighbouring solution is exactly the same as the base solution except for one of the decision variables which has a difference of one unit compare to the base solution. In this study each solution has six neighbouring solutions. After examining all the neighbours, the best neighbour is considered as the base and in a similar way all the neighbour solutions to new base are evaluated. This algorithm is repeated until a local optimum is reached. This neighbourhood searching is conducted for each individual in the last generation. The best local optimum is considered as the problem solution.

5. Numerical example

To illustrate the result of the analysis conducted in session 3, the following numerical examples are considered and the heuristic method is applied. In order to see the advantage of supply chain perspective and the benefits that the retailer and the wholesaler gain through integrated planning, in these examples the inventory systems are optimised both with supply chain and single company point of view.

In order to optimise the system of the retailer and the wholesaler separately (non-integrated approach), first the inventory
system in the retailer is optimised and having the inventory policies calculated for the retailer, the inventory system is optimised for the wholesaler using an analytical approach (see Appendix B). The results of this optimisation are also presented in the following numerical examples.

**Example 1.** Let time unit be day, demand function, $D(t) = 0.2I_0(t)/(-t + 200)$, warehouse capacity in OW, $W = 200$, deterioration rate in OW, $\alpha = 5\%$, deterioration rate in RW, $\beta = 8\%$, deterioration rate in wholesaler's warehouse, $\gamma = 3\%$, percentage of backlogged demand during shortage time in retailer, $\delta = 50\%$, ordering cost for retailer, $A_R = 1500$, ordering cost for wholesaler, $A_W = 2500$, purchasing price for retailer, $p_R = 8$, purchasing price for wholesaler, $p_W = 3.5$, - carrying cost in OW, $h_W = 0.4$, carrying cost in RW, $h_R = 0.5$, carrying cost in wholesaler, $h_{OW} = 0.3$, fixed shortage cost per unit, $c_s = 20$, - and variable shortage cost, $c_v = 4$. As can be seen in Table 1, the total cost of the system is decreased by 14% when optimising the system with a supply chain perspective. As a result of this integration, the total cost of the retailer and the wholesaler are dropped by 6% and 33% respectively.

The optimal values for $t_t$ and $t_r$ are 0 and 1.9 respectively which means the model does not suggest using RW. The shortage period is 1.9 unit of time during which 380 units are demanded but 50% are backlogged. Therefore the demand for 190 units of the product is backlogged waiting for the new replenishment. As soon as new replenishment is received (order quantity = 390), 190 units are used to cover the backlogged demand from previous period and the rest of the order quantity is stored in the warehouse.

**Example 2.** Let time unit be day, demand function, $D(t) = 0.1I_0(t)/(-t + 100)$, warehouse capacity in OW, $W = 50$, deterioration rate in OW, $\alpha = 5\%$, deterioration rate in RW, $\beta = 8\%$, deterioration rate in wholesaler's warehouse, $\gamma = 3\%$, percentage of backlogged demand during shortage time in retailer, $\delta = 40\%$, ordering cost for retailer, $A_R = 1000$, ordering cost for wholesaler, $A_W = 2500$, purchasing price for retailer, $p_R = 8$, purchasing price for wholesaler, $p_W = 3.5$, - carrying cost in OW, $h_W = 0.4$, carrying cost in RW, $h_R = 0.5$, carrying cost in wholesaler, $h_{OW} = 0.3$, fixed shortage cost per unit, $c_s = 30$, - and variable shortage cost, $c_v = 4$.

In order to study the effect of changes in parameters on the total cost and decision variables, a sensitivity analysis is conducted. In order to carry out the analysis, the all parameters in the model have been set to two different levels apart from the value in the basic example (20% decrease and increase compared to the original level) and the change (CTC) in total cost has been calculated.

$$CTC = \frac{TC^{new} - TC}{TC} \times 100\%$$

In the following section, Table 2 illustrates the results of the solution for the example and Table 3 presents the changes in different parameters and the effects of these changes on the solution.

**Table 1**
Results of the numerical example.

<table>
<thead>
<tr>
<th></th>
<th>$k$</th>
<th>$Q_R$</th>
<th>$T_e$</th>
<th>$Q_W$</th>
<th>$T_W$</th>
<th>$TC$</th>
<th>$TC_W$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before integration</td>
<td>200</td>
<td>0.9</td>
<td>1520</td>
<td>6.3</td>
<td>5021</td>
<td>3522</td>
<td>1499</td>
</tr>
<tr>
<td>After integration</td>
<td>50</td>
<td>2.8</td>
<td>1275</td>
<td>8.4</td>
<td>4290</td>
<td>3299</td>
<td>1000</td>
</tr>
</tbody>
</table>

The results presented in Table 2 suggest that the inventory period in retailer should be 5.0 days and that a quantity of 381 units of the product should be ordered at the beginning of each period. During the first 2.0 days of each cycle, the inventory stored in RW is used to cover the demand. After this time the inventory in OW is used as there is no inventory in RW. The retailer's inventory in OW is depleted completely at the time 2.4 after which there is a shortage period of 2.6 days.

The model also suggests that wholesaler's inventory cycle should be 10.0 days and that at the beginning of each period an order of 825 units should be placed. As can be seen in Table 2, the total cost of the system per unit of time is 2838.

After examining the effects of the changes on total cost function, the inventory policies in both retailer and wholesaler are considered to investigate how these values react in the case of any changes in parameters (Table 4).

Studying the output of the sensitivity analysis, the following points can be made:

1. As the results in Table 3 show, the total cost is most sensitive to $d$ of which a 20% change can make a change around 17% in the same direction. It should be noted that this model is based on the total cost and not profit; therefore this increase in cost cannot be interpreted without considering the total profit of the system.

2. The total cost shows the second highest sensitivity to $p_b$, $C_s$ and $h_{OW}$. According to Table 3, all these factors change the total cost in the same direction. As the total cost of the system is sensitive to the purchasing price that both the retail and the wholesaler pay, applying discounting models in their business is highly recommended. Nevertheless buying larger amount necessitates both retailer and wholesaler to focus more on marketing in order to increase the demand. As mentioned before, the cost function is highly sensitive to demand, which means that the retailer can decrease its margins on the product and influence the demand (in case the product is price elastic) and from the other side, purchasing in larger quantities and enjoying lower prices.

3. The cost of an unmet demand which becomes lost sale is very complicated to quantify as it depends on many factors. If there are competitor products on shelf, lost sale may cause the product to lose the competition to those which are present on shelf. In some cases when the product obtains especial characteristics, the customer may go to another store to look for the desired product and they may do the whole shopping in the second store and this could be considered as a lost profit for the first store. The shortage of the product may have a negative affect on the image of the product brand or the retailer. Usually the cost of lost sale is a managerial decision as it needs experience and knowledge about the market and the competition.

4. The total cost is least sensitive to $W$, $h_R$ and $h_{OW}$. In the case of a 20% change in these parameters the change in total cost is almost zero. As warehouse capacity is a strategic decision which may cause a huge amount of cost or saving for a retailer, studying the influence of warehouse capacity on total cost needs a more detailed sensitivity analysis. Considering the same product with the same market, a change in warehouse capacity from 50 to 150 decreases the total cost by 9%. In case the distribution system is
to deliver a product of which the service level is relatively high (from a modelling point of view the cost of shortage per unit of product is considered to be a large number), the total cost of the system would be 3014. In this example a change of warehouse capacity from 50 to 150, make a decrease of 14% in the total cost. However evaluating the trade-offs between the cost of expanding the warehouse capacity and the decrease in total cost in order to make this strategic decision is a managerial task.

5. The effects of changes in inventory policies in both retailer and wholesaler are presented in Table 4. As can be seen a 20% increase in $p_R$ shows 40% and 22% decrease in $Q_p$ and $T_R$ respectively which means that the model suggests the retailer a higher frequency with smaller order quantity.

6. An increase in $W$, $c_w$, and $\beta$ warehouse capacity, make a change in $Q_p$ and $T_R$ in the same direction as $p_R$ does. In case of an increase in $W$, the model tries to minimise the shortages and
the usage of RW by using extra capacity through higher frequency in replenishments. Higher variable shortage cost \((c_{sw})\) has the same influence on the optimal solution as the model tries to avoid shortages. Due to deterioration and the capacity limits however the model cannot increase the order quantity, therefore it suggests higher frequency with smaller order quantities. Higher level of deterioration in RW also motivates the model to adopt lower order quantities and inventory periods for the retailer to minimise the use of RW and the corresponding costs.

7. A decrease in \(A_{0}\), \(\gamma\) and \(h_{W}\) suggests lower order quantity and shorter inventory period for the retailer. These changes are intuitive when decreasing \(A_{0}\). In case of a decrease in \(\gamma\) and \(h_{W}\), the model tends to keep the inventory in the wholesaler’s warehouse and avoid the higher holding cost and deterioration in the retailer’s level. This however is the case when holding cost and deterioration rate in the retailer remains unchanged.

8. As presented in Table 4, in general the decision variables related to the wholesaler show less sensitivity to the changes in the parameters. In case of a 20% decrease in \(h_{W}\), \(h_{W}\) and \(\gamma\) the model suggests the wholesaler to order larger order quantities with less frequency.

9. Some parameters that are related to the retailer may indirectly influence the decision variables of the wholesaler. In case of an increase in \(c_{sf}\) for instance, the model suggests lower shortages in the retailer. Lower shortage level is obtained through 17% and 36% decreases in the retailer’s order quantity and inventory period respectively. In order to cover this demand (which is higher than the initial example) the wholesaler is suggested to have shorter inventory period (a decrease of 7%) but larger order quantity (an increase of 25.6%).

6. Conclusions

In this study an analytical model for deteriorating inventory is developed considering a limit for warehouse capacity in retailer. One of the main features which characterises this model is the supply chain perspective which means that, in determining optimal policies, costs of the wholesaler and the retailer are considered and minimised simultaneously. Most of the studies done in this area are from a single company’s point of view. In this model purchasing cost, inventory carrying cost and deterioration cost for both wholesaler and retailer and shortage cost only for retailer are taken into account. In the solution part, a heuristic method is applied to find a fairly good solution to avoid time consuming calculations. Another feature of this model is that it is more generic compared to other models: by setting \(c = 0\) the model will be converted to a model with constant demand. In addition, this model can change to a model with complete backlogging or lost sale.

As this model considers a percentage of in-hand inventories getting deteriorated, it cannot be applied to the distribution systems delivering products with fixed life-time. Although this two-echelon model differs greatly from real-world problems, it requires complex mathematical calculations. In order to make this model more realistic, some extension opportunities are suggested. For further research the inflation rate is suggested to be considered in cost calculation which is part of the real-life problems. In some cases wholesalers accept delays in payments by retailers in order to motivate them to increase their order quantity. This model considers and analyses a distribution system which consists of one retailer and one wholesaler. Normally that is not the real case and distribution chains consist of retailers and wholesalers. Therefore an extension that is suggested is to consider multi-suppliers and multi-retailers. Few studies have been conducted on deteriorating inventory models for multi-product systems, and as there is a competition between products for warehouse capacity, this assumption is also suggested for further research. Finally, quantity discount models and transportation models between echelons are recommended for future research.

Appendix A

In order to solve the problem analytically, first the objective function should be simplified using some approximation. The error of this approximation is acceptable only when \(kt_{R}\) and \(T_{R}\) have relatively small values. Using Taylor expansion the value of the wholesaler order quantity is as follows:

\[
Q_{W} = XQ_{k}
\]

where:

\[
X = \left(\frac{e^{\gamma T_{R}} - 1}{e^{\gamma T_{R}} - 1}\right)^{1 + (k - 1)T_{R}} \left(1 + \frac{1}{2}k^{2}T_{R} - \frac{1}{2}k^{2}T_{R}\right)
\]

The second derivatives of the total cost function with respect to \(t_{R}\) is as follows:

\[
\frac{d^{2}TC}{dt_{R}^{2}} = \frac{2}{T_{R}} \left(\frac{A}{T_{W}} - B\right) + c_{sf}d + \frac{2}{kT_{R}} \left(\frac{C}{T_{R}} - D\right) + \frac{1}{kT_{R}} \frac{dD}{dt_{R}}
\]

where:

\[
A = (PC_{R} + ICC_{R} + DC_{R} + SC_{R})
\]

\[
B = p_{W}d + c_{sf}(1 - \delta)\]d + c_{sw}d
\]

\[
C = (PC_{W} + ICC_{W} + DC_{W})
\]

\[
D = \left(\frac{dX}{dt_{R}} + Q_{k} \frac{dX}{dt_{R}}\right) \left(2p_{W} + \frac{h_{W}}{\gamma}\right) - k\delta \left(p_{W} + \frac{h_{W}}{\gamma}\right)
\]

\[
\frac{dX}{dt_{R}} \approx \frac{1}{2}k\gamma(k - 1) + k^{2}T_{R}(\frac{k^{2}}{2} - k - 1)
\]

By having \(k\) and \(t_{R}\) set and changing \(t_{R}\) from zero to a large number that it can possibly take, the second derivative of the total cost function with respect to \(t_{R}\) presented in (30), takes a positive value. It however necessitates complex analysis to show the convexity of the total cost function.

An exhaustive search can be conducted to find the global optimum as an upper bound can be assumed for each of the decision variables. There is however a trade-off between the accuracy of the decision variables and the solution time.

In order to conduct an exhaustive search it is assumed that \(k\) can get a maximum value of 30. \(t_{R}\) and \(t_{W}\) are assumed to be less than 5 years with the accuracy of \(e = 0.01\) (the heuristics is solved with the same accuracy). The search results in the same optimal solution but the solution time is much longer. Conducting the exhaustive search with a higher accuracy will increase the solution time while this increase in the genetic algorithm time is small (see Table 5).

Appendix B

In this part two separated problems are analysed. Firstly the retailer’s inventory system is studied and the inventory policies are set. In the next step using the retailer’s optimal inventory policies, the wholesaler’s inventory system is analysed and optimised.

Based on analyses done in Section 3.3, the inventory cost for retailer per unit of time is as follows:

\[
TC_{R} = \frac{1}{T_{R}} (PC_{R} + ICC_{R} + DC_{R} + SC_{R})
\]

The cost function contains \(T_{R}\) and \(Q_{k}\), therefore partial derivatives of these two terms with respect to \(t_{R}\) and \(t_{W}\) are also needed. Based on (10) and (11) the derivatives are as follow:
The total cost of the retailer is a function of \( t_s \) and \( t_r \). To minimise the total cost, the derivatives of the cost function with respect to \( t_s \) and \( t_r \) are found and set equal to zero.

\[
\frac{dTC_R}{dt_s} = 0, \quad \frac{dTC_R}{dt_r} = 0
\]

The optimal values of \( k \) can be derived from the following set of inequalities:

\[
TC_W(k') \leq TC_W(k' - 1)
\]

\[
TC_W(k) \leq TC_W(k + 1).
\]

References


