Simulation-based inventory management of perishable products via linear discrete choice models

Computers & Operations Research 157 (2023) 106270

Contents lists available at ScienceDirect

Computers and Operations Research

journal homepage: www.elsevier.com/locate/cor

Simulation-based inventory management of perishable products via linear discrete choice models

Daniele Giovanni Gioia *, Leonardo Kanashiro Felizardo, Paolo Brandimarte

Department of Mathematical Sciences "Giuseppe Luigi Lagrange", DISMA, Politecnico di Torino, Corso Duca degli Abruzzi 24, Turin (TO), 10129, Italy

ARTICLE INFO

Dataset link: https://github.com/DanieleGioia/PerishableDCM

Keywords:
Multi-item inventory systems
Perishable products
Inventory control
Simulation-based optimization
Discrete choice models

ABSTRACT

Retail inventory management of perishable items, like fresh food, is a relevant and complex problem. It is relevant in the light of trends towards the reduction of food waste, and because of potential cross-sales interaction with other item categories. It is complex, because of multiple sources of uncertainty in supply, demand, and quality, and other complicating factors like seasonality within the week, FIFO/LIFO consumer behavior, and potential substitutions between items, possibly because of a stockout. Similar items may be vertically differentiated due to intrinsic quality, which is also related with item age, or brand image, as it could be the case when a retail chain stocks both a brand item and a private label one. In the paper, we adapt a simple discrete choice model to represent consumers' heterogeneity and different tradeoffs between price and quality, and apply simulation-based optimization to learn simple ordering rules for two vertically differentiated items, adapted to a seasonal case, in order to maximize long-term average profit under a lost sales assumption. While well-known constant and base-stock policies need not be optimal, they are simple to communicate and apply. We explore combinations of such rules for the two items, obtaining some useful managerial insights.

1. Introduction and motivation

Inventory management of food, and more generally perishable items, is gaining importance due to the increased attention paid to food waste and its implications for environmental footprint and social responsibility. Supermarkets deal with substantial losses due to moisture loss, spoilage, and other causes (Buzby et al., 2009), and a better understanding of how to manage perishable products may reduce the amount of waste, while improving profits at the same time. A detailed timeline of the recent major institutional efforts aimed at reducing food waste is available in Akkaş and Gaur (2022), where they identify the perishable inventory management problem as an active research opportunity to achieve more acceptable sustainability, especially because of the new types of data, technologies, and business models available nowadays.

Since the literature on inventory management of perishable items is quite vast, a common preliminary way to navigate deteriorating inventory issues concerns a classification through the shelf-life characteristics. The classification proposed by Goyal and Giri (2001), and then adopted by Bakker et al. (2012) and Janssen et al. (2016), distinguishes between:

- items where deterioration occurs in continuous time, at a given rate;
- items with a predetermined fixed shelf-life, so that items become unfit for sales at a given time instant;
- items with a probabilistically distributed shelf-life.

In this paper we assume a *discrete-time* inventory model with a *single* retailer and a *multi-product* setup, where a *fixed shelf-life* hypothesis holds, together with *deterministic lead-times* for each product. The assumption of deterministic lead-times and shelf-lives may be more or less realistic, depending on the specific case, but it is not necessarily a simplification. Indeed, it prevents the application of renewal theory (Karaesmen et al., 2011), which would otherwise be possible with exponentially distributed characteristics. Moreover, when we consider multiple products, featuring different ages on the shelves and different times to delivery, the dimensionality of the state space makes the problem analytically prohibitive. Because of that, simulation-based optimization may provide insightful results (Deng et al., 2023; Haijema and Minner, 2016; Jalali and van Nieuwenhuyse, 2015). In the model, we consider only uncertainty on the consumer's side, related to demand, which is modeled as a discrete random variable, and to the random utility associated with different available products.

The standard multi-product inventory management problem is a well-debated topic in the production research community, but the

* Corresponding author.
E-mail addresses: daniele.gioia@polito.it (D.G. Gioia), leonardo.kanashiro@polito.it (L.K. Felizardo), paolo.brandimarte@polito.it (P. Brandimarte).

https://doi.org/10.1016/j.cor.2023.106270
Received 24 August 2022; Received in revised form 12 March 2023; Accepted 2 May 2023
Available online 6 May 2023

literature on multiple deteriorating items is scarce (Janssen et al., 2016), even though the practical benefits of a control strategy relying on substitutable products are well known (Fisher and Raman, 2010). Actually, the substitution between similar perishable products regularly happens, especially during a stock-out (Van Woensel et al., 2007), providing a promising way to reduce the waste of food with a short shelf-life (van Donselaar et al., 2006). Investigating further the multi-product perishable inventory problem would concretely benefit the way the issue is addressed by exploring management policies that take into account the joint effects of different items. In particular, the multi-product setting may also apply to a single product, if we consider the availability of items with different residual shelf-life and its impact on consumers' behavior (Tsiros and Heilman, 2005). The central focus of our endeavors is on the construction of a flexible model relying on utility-based preferences for each client, leading to a *discrete choice model* (Train, 2009) in a perishable setting, which enables modeling endogenous substitution rates *between and within* products. We pursue this goal by adopting a *linear* discrete choice model, widely used for vertically differentiated products (Transchel, 2017; Pan and Honhon, 2012; Transchel et al., 2021). This flexible framework can be also applied to deal with deteriorating items, where age has an impact on the perceived utility and may be used to model age-based substitution in a multi-item setting.

The use of an approach based on discrete choice models also implicitly addresses another standard feature of deteriorating inventories, i.e., the inventory depletion mechanism. Extreme assumptions are either a pure Last-In–First-Out (LIFO) or a pure First-In–First-Out (FIFO), or a prespecified hybrid. When the inventory issuing is under the control of an inventory manager, as is the case of blood platelets in hospitals (Haijema et al., 2005; Nahmias and Pierskalla, 1976), a FIFO pattern may be expected, whereas when consumers make the choice themselves, one could expect a pattern leaning towards LIFO. In the multichannel case described in Deng et al. (2023), a hybrid case is obtained due to the presence of an online channel, where an inventory manager is in charge of picking the item, and a physical channel, where consumers pick the selected item. Also in a case where the consumer is in charge of picking, a mixed LIFO/FIFO or possibly a purely random selection may be observed, as some environmentally conscious consumers may behave according to a FIFO mechanism. In this paper, we consider a case in which heterogeneous consumers can choose between vertically differentiated items, featuring different stock ages and, possibly, discounted prices. Hence, we cannot assume a mechanism fixed *ex ante*, and the LIFO/FIFO behavior must be made endogenous. Dynamics are complex and call for a simulation-based approach to tune the parameters of standard replenishment policies.

In this paper we do not deal with assortment and pricing decisions explicitly. We refer to Transchel (2017) or Ferguson and Koenigsberg (2007) for examples of investigations on these problems by linear price-based utility models. When optimizing the parameters of the replenishment policies, prices are given. Nevertheless, since they are included in the choice models, the approach can be used to explore the impact of different prespecified pricing policies.

1.1. Contributions of the paper

The parameterized replenishment policies that we apply in the paper are not optimal, but their performance has been compared with exact stochastic dynamic programming in Haijema and Minner (2019), where it is shown that they can achieve a performance not too far from the optimal one. Moreover, they are more easily accepted by practitioners. Simulation-based optimization is a widely used approach to fine tune such policies (Haijema and Minner, 2016), and we use standard search algorithms in this paper. So, we do not claim to provide any methodological contribution in this respect. Our contribution is in the application of these policies within a more complicated environment, featuring seasonality, vertically differentiated items, and heterogeneous consumers trading off perceived quality (which is both product- and age-dependent) and price. We stress the fact that the interaction between these factors and a non-zero delivery lead time increases state dimensionality, precluding not only the application of exact dynamic programming, but also the search for a large number of parameters of more sophisticated replenishment rules. We also pay attention to practically relevant issues, like order size variability, which may create difficulties with suppliers. We consider strategies by which two substitute items may help in managing seasonality. To the best of our knowledge, this combination of factors has not been investigated, but it represents a useful step towards interesting applications.

It is also worth noting that the developed simulation-based framework is publicly available and may be employed as an experimental search tool for optimal pricing and optimal assortment problems as well, filling the lack of open-source libraries in the perishable inventory literature.[1]

1.2. Plan of the paper

The paper is organized as follows. In Section 2 we discuss the literature related to this work. In Section 3 we illustrate the simulation framework, the selected discrete choice model, and the role of uncertainty in our simulation experiments. Section 4 connects the simulation framework with sequential optimization concepts and introduces the heuristic replenishment policies. In Section 5 we present the experimental results and provide some managerial insights obtained through the selected price/quality business scenarios. Section 6 concludes with some final remarks and topics for future research.

2. Literature review

The literature about inventory management problems is massive and, although focusing on perishable products reduces its size, we still have to deal with a huge amount of studies. To provide a readable literature overview, we start by narrowing down the list of topics to better fit our specific research aim. In particular, we will concentrate on models where the inventory is controlled over a horizon, products have a fixed shelf life without obsolescence risk (e.g., abrupt reduction in value due to market competition reasons), and there is a single retail point, under the assumption of an efficient supply chain so that fixed ordering charges are shared among several items and are negligible at the single item level. It follows that we will treat neither single-period models, commonly tackled by newsvendor-based approaches, nor multi-echelon systems. We refer to Goyal and Giri (2001), Bakker et al. (2012), Janssen et al. (2016) and Karaesmen et al. (2011) for a detailed review and additional references for models on perishable products under alternative hypotheses.

We assume a discrete-time model with periodic review and focus on parameterized replenishment policies that are tuned by simulation-based optimization. However, continuous review models are common as well. A significant contribution in this area is due to the work of Perry (see, e.g., Perry and Stadje, 2000), which is also discussed in detail by Karaesmen et al. (2011). Another example of a continuous review approach based on a Markovian renewal model is Kalpakam and Shanthi (2001).

Some analytical approaches under discrete-time review hypotheses rely on stochastic modeling (Ishii and Nose, 1996) or MILP modeling (Pauls-Worm et al., 2014). Nevertheless, our work is closer to dynamic programming and heuristic approaches like (Hendrix et al., 2019; Haijema and Minner, 2019). In this context, assumptions on the lead-time and the shelf-life are essential features that affect the

[1] The code of the simulation environment and the related dynamic setting are freely available at https://github.com/DanieleGioia/PerishableDCM.

complexity of the problem, quickly creating unmanageable state spaces of the system because of an increase in problem dimensionality.

Early examples of state aggregation, oriented to develop a tractable approximation, solve dynamic programming problems of reduced dimension (Nahmias, 1977), where they collapse the state vector noticing that the ordering function is more sensitive to changes in newer inventory than older inventory. However, it is important to observe that they assume a FIFO depletion mechanism, which supports such an observation. Similarly, Brodheim et al. (1975) adopts a Markov chain model with a manageable number of states by making a decision only based on the number of new items of the system, again assuming a FIFO issuing policy in a blood inventory management application. The effect of a dimensionality reduction on the performance of a dynamic programming approach has been recently explored by Haijema and Minner (2019), where they investigate the value of stock-age information through a simulation-based optimization procedure, following a stream of research based on traditional stock-level dependent ordering policies and optimal disposal policies for perishables (Haijema and Minner, 2016; Haijema, 2014).

The presence of more than one product definitely increases the complexity of the model, making it really hard to investigate analytically. The research on perishable inventories has largely been confined to single product contexts (Karaesmen et al., 2011), especially when non-zero lead times or seasonality are involved.

As the foremost characteristics of a multi-item setting in inventory decisions, we can identify the actor in charge of the substitution between items and the substitution direction (Shin et al., 2015). If the substitution is supplier-driven, the supplier decides what to offer to the consumers, providing suitable models for blood oriented studies (Haijema et al., 2005; Nahmias and Pierskalla, 1976). Specifically, in Nahmias and Pierskalla (1976), only one of the available products is subject to deterioration in a blood-related application. On the contrary, we address a grocery oriented multi-product setting, closer to Hendrix et al. (2019) and Buisman et al. (2020), where the substitution is customer-driven and it is up to the client to decide what to buy. However, the aforementioned works assume an *exogenous substitution* between the available products, whereas we develop a utility-based *endogenous* one. Moreover, Hendrix et al. (2019) deal with a dynamic programming formulation that suffers from the curse of dimensionality, limiting the analysis to inventories where a small number of items is held. Similarly to us, Buisman et al. (2020) rely on simulation-based optimization and study *Order-Up-To* policies, with and without consideration of stock-out based substitutions. Nonetheless, they have to fix an issuing policy (LIFO/FIFO) and they operate by batches of clients rather than by single customers. As far as the direction of substitution is concerned, it can be managed by a model based on consumers' utility.

Our framework represents consumer behavior by a *Discrete Choice Model* (DCM), which describes their preferences among a set of alternatives and may be estimated by consolidated statistical methodologies (McFadden, 1973). However, since different assumptions may lead to distinct models (Train, 2009), the characteristics of the problem must be accurately considered. We preliminary distinguish between horizontally and vertically differentiated items. Horizontal differentiation allows for idiosyncratic preferences that are not purely based on a quality–price ranking (Transchel et al., 2021), whereas the vertical one considers cases where the choice would be driven only by quality, if the price of the products were the same. Vertical differentiation fits well our grocery application, where we are interested in the effect of substitution for very similar products, especially in the event that products only differ in their *Residual Shelf-Life* (RSL) or, for example, when private labels coexist with external brands. To this aim we employ a *linear* DCM, extensively studied in Pan and Honhon (2012) and Transchel (2017), and then (Transchel et al., 2021) from an analytic point of view, giving us several advantages in our simulations. Specifically:

- Despite the absence of an optimal policy benchmark, due to the model complexity and generality, we can obtain an essential understanding of the resulting performance.
- Even though we have no real data currently, it is possible to set meaningful parameters, identifying practical insight by the comparison of what we expect and what we get.

However, in contrast to our focus on replenishment policies for perishable items, the aforementioned papers also investigate pricing and assortment problems and their applications do not consider deterioration. Furthermore, seasonality and significant lead times are not dealt with in the former analytical works, because they would make the exact computation prohibitive.

3. The simulation framework

We have developed an open-source flexible simulation-based environment that models a single retailer who has to make daily decisions on the quantity to order of J vertically differentiated perishable products. Each product $j \in \mathcal{J} = \{1, \dots, J\}$ is characterized by the following features:

- LT_j: fixed discrete delivery lead time.
- SL_j: fixed discrete shelf-life (at the time of delivery).
- The market economics of the product, i.e., selling price (p_j^d), purchase cost (c_j), perceived quality (q_j^d) and implied margin ($p_j^d - c_j = m_j^d$). Some of these features explicitly depend on the residual life of each item on-hand of that product $d \in \{1, \dots, \mathrm{SL_j}\}$.

Each simulation step t corresponds to a day and is subject to weekly seasonality, under which the uncertainty on demand depends only on $k = (t+1)_{\mathrm{mod}7}$. At the beginning of each day, after the respective lead time has elapsed, incoming orders are handed over to the store and inventory is updated accordingly. After delivery, the business day starts and we sample the total number of clients that will interact with the store by a loop that simulates their individual utility, according to given discrete choice model. If a consumer has a positive utility for any product among those available, she will pick one item, choosing the one maximizing her utility and removing that particular item from the inventory. When the business day ends, the retailer checks if any item has reached the end of the shelf-life and scraps it accordingly. Finally, the state of the system is observed to determine the next order size per product, according to the specified ordering policy, completing the simulation step for the day.

3.1. Demand simulation

The model allows for any kind of distribution to sample the number of consumers visiting the store per day. It is worth stressing again that their behavior will be simulated individually in sequence, on the basis of the discrete choice model and the available inventory. As to seasonality, we adapt it to a Poisson distribution, which is characterized by its first moment. Specifically, we define a multiplicative seasonality factor η_k, such that the expected number of consumers for day k is

$$\mu_k = \eta_k \mu, \quad k \in 0, \dots, 6, \tag{1}$$

for a selected value of μ.

3.2. The discrete choice model

After sampling the number of consumers N for the present day, the simulation processes each consumer $n \in \{1, \dots N\}$ by sampling its random utility $\mathbf{U}_n$ for each of the available products (as we specify later in this section, the utility for each consumer depends on a random coefficient associated to that individual consumer). Practically speaking, we could think that the inventory is updated like a shelf of a supermarket, where several products are available for the consumer, who associates a utility with each of them: $\mathbf{U}_n = \{U_{nj}, j \in \mathcal{J} \cup 0\}$. We also fictitiously set to 0 the utility for the products currently unavailable and include a dummy product $j = 0$ that serves as a no-choice option when the utility is negative for all the other possibilities. Theoretically, the framework allows for any discrete choice model and we refer to Train (2009) for a complete overview. However, we are especially interested in the substitution effects of similar products on grocery applications, where the vertical differentiation hypothesis is a sensible assumption.

For each consumer n, we compute a linear utility, widely used to model vertically differentiated non-perishable products (Pan and Honhon, 2012; Transchel, 2017; Transchel et al., 2021). To adapt the idea to perishable items, we treat items with different age as different ones, enlarging the set of available products $\mathcal{J} = \{1, \dots, J\}$ to $\mathbb{J} = \{1, \dots, J_{\mathsf{P}}\}$, where:

$$J_{\mathsf{P}} = \sum_{j=1}^{J} \mathsf{SL_j}. \tag{2}$$

For instance, if we consider a scenario with two products (A and B) with different shelf life of 4 and 2 periods, respectively, the original number of possible choices would be $J = 2$, and a non-perishable utility model would allow for A, B, or nothing. By applying this stock-age adjustment, the dimensionality increases to $J_{\mathsf{P}} = 4+2$, where both products A and B may be chosen with different residual shelf-lives, where the depreciated qualities may be counterbalanced by discounts on the respective price. We may consider each product at different stock age as a specific item of that product. Analytically, according to the linear discrete choice model, consumer n evaluates the utilities

$$U_{nj} = \theta_n q_j - p_j, \quad \forall j \in \mathbb{J} \cup \{0\}, \tag{3}$$

where q_j is the perceived quality of item j, p_j is its selling price, and θ_n is a stochastic factor accounting for how consumer n trades off quality and price, which we model by a beta random variable. It follows that we allow for heterogeneous agents with a individual quantitative evaluation of the price–quality (p_j/q_j) ratio of the available products, implying an age-based substitution that interacts with the stock-out based one. In particular, the age-based substitution is a well-known issue in deteriorating inventories, especially studied in pricing problems (Ferguson and Koenigsberg, 2007). Hence, our model is a realistic generalization.

4. Dynamics of the policy and sequential decisions

After specifying the demand model, we aim to develop a policy that makes (near-)optimal decisions for each simulation step. Following the typical paradigms of dynamic programming (Powell, 2021; Brandimarte, 2021), we first define a sequential decision framework based on the observed state of the system, the immediate rewards, and the consequent actions. Then we select and optimize alternative heuristic policies by a simulation-based strategy.

4.1. State, actions and reward of the sequential approach

State variables. When a business day ends, the decision-maker observes the state of the system, which depends on the on-hand inventory and the in-transit orders. Specifically, we can formally define it by introducing the following quantities for the products in $\mathcal{J}$ (i.e., the actual set of products, without accounting for age):

- $O^l_{t,j}$: number of items of product $j \in \mathcal{J}$ to be delivered in $l \in \{0, \dots, \mathsf{LT_j} - 1\}$ days at time t.
- $I^d_{t,j}$: on-hand inventory of product $j \in \mathcal{J}$ with a residual life of $d \in \{1, \dots, \mathsf{SL_j}\}$ days at time t.
- Day $k \in \{0, \dots, K = 6\}$.

The observed state variable at time t for each product $j \in \mathcal{J}$ will be:

$$S_{t,j} = \left[O^{\mathsf{LT_j}-1}_{t,j}, \dots, O^0_{t,j} | I^{\mathsf{SL_j}}_{t,j}, \dots, I^1_{t,j}\right]. \tag{4}$$

Hence, the complete state variable in our simulations will be

$$S_t = \left[S_{t,1} | \dots | S_{t,J} | k\right].$$

The dimension of the state space is $\sum_j ((\mathsf{SL}_j - 1) + \mathsf{LT}_j) + 1$, because when the state is observed, the expired items with zero residual life have been already scrapped and the time to delivery of the most recent order is reduced by 1 to $\mathsf{LT_j} - 1$, for each item j.

Action space. For each product, the decision-maker must decide the order size, so that the control action is an array that determines the next shipments of goods. When the order is placed, it is subject to the full lead time, thus we define the time t action as:

$$x_t = \left[O^{\mathsf{LT}_1}_{t,1} | \dots | O^{\mathsf{LT}_J}_{t,J}\right]. \tag{5}$$

To optimize the replenishment, we would like to design a policy π such that, for each observed state S_t,

$$x_t = X^\pi(S_t). \tag{6}$$

We provide an overview of the complete dynamics of the simulation framework in Fig. 1, where the state variable and the actions are emphasized.

Reward function. Our simulation environment implements an offline learning approach. Moreover, the immediate reward is myopic due to the gap between the moment when the items are purchased and the one where they are actually sold. In fact, the further complexity introduced by seasonality produces an additional gap between what is sold on a day and what is bought to be ready for the next days. Practically speaking, this generates significant negative rewards when the seasonality factor is low and substantially favorable ones during peaks of the weekly demand.

On the one hand, the retailer pays

$$\sum_{j=1}^{J} c_j O^{\mathsf{LT}_j}_{t,j} \tag{7}$$

when a replenishment order is issued. On the other hand, she earns a revenue from items sold during a business day,

$$\sum_{j=1}^{J} \sum_{d=1}^{\mathsf{SL}_j} p^d_j L^d_{t,j}(S_{t,j}), \tag{8}$$

where $L^d_{t,j}(S_{t,j})$ denotes the sold items of product j at time t with residual shelf-life d, subject to inventory availability and demand uncertainty.[2] It follows that the reward on day t is:

$$C(x_t, L_t) = \sum_{j=1}^{J} \left(-c_j O^{\mathsf{LT}_j}_{t,j} + \sum_{d=1}^{\mathsf{SL}_j} p^d_j L^d_{t,j}(S_{t,j}) \right). \tag{9}$$

4.2. Implemented policies

Exact solution approaches that consider the entire state space by dynamic programming and explicitly represent a value function by lookup tables suffer from the curse of dimensionality. Hence, the complexity of the problem requires using approximated methods and heuristics, often based on a direct approximation of the policy function, rather than a value function.

[2] The residual shelf-life superscript d can be avoided by modifying the set of on which the index j ranges from $\mathcal{J}$ to $\mathbb{J}$. We use this explicit notation for the sake of clarity.

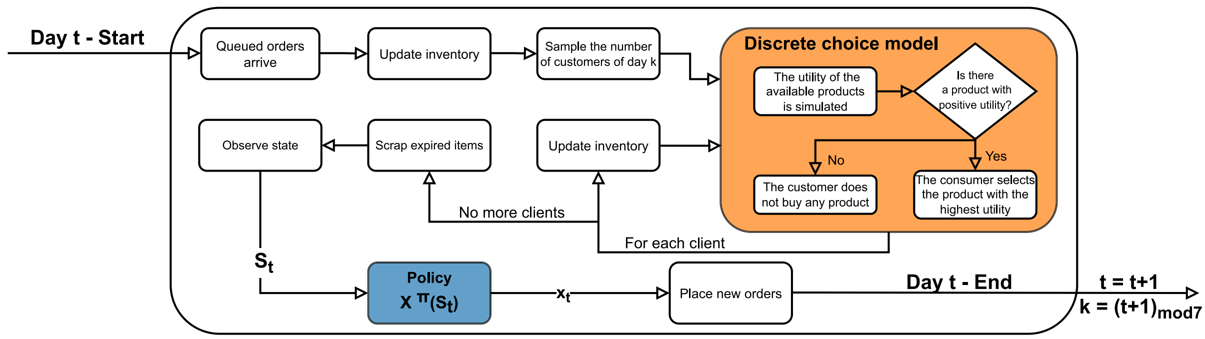

Fig. 1. Simulation framework and interaction with the ordering policy.

Base-stock policy. Under the hypotheses of a periodic review and discrete shelf-life, the classical heuristic approach to discrete-time inventory management of perishable items relies on *Base-Stock policies* (also known as *order-up-to rules*). Various examples include applications to single-product settings (Haijema and Minner, 2019, 2016) and multi-product ones (Buisman et al., 2020).

In this paper, we employ a Base-Stock Policy (BSP) by defining, similarly to (Gioia et al., 2022),

$$X_j^{\text{BSP}}(S_t|\mathbf{z}) = \left(z_{j,k} - \left(\sum_{l=1}^{\text{LT}_j} O_{t,j}^l + \sum_{d=1}^{\text{SL}_j} I_{t,j}^d\right)\right)^+, \qquad j \in \mathcal{J}. \tag{10}$$

The policy depends on the observed state of the system and reduces the state dimension by aggregating items across their stock age. The BSP policy, as well as the alternative ones that we will investigate hereafter, is a parametric *Policy Function Approximation* (PFA) (Powell, 2021), which depends on $z_{j,k}$. This vector is optimized by a simulation-based approach and it has a dimension of $J \times (K+1)$. For example, if there are two products, since we have a 7-days based seasonality, $J \times (K+1) = 2 \times 7 = 14$. Notice that, since we can only order brand new items, the stock-age adjustment of (2) is only considered in the discrete choice model, whereas it is not in the replenishment policies.

Constant order policy. The simplest way to generate meaningful orders and to take the seasonality pattern into account is a weekly *Constant Order Policy* (COP). Unlike the BSP policy, it relies on a smaller vector of parameters, $\mathbf{z} \in \mathbb{R}^{J\times(K+1)}$, since replenishment orders are issued without evaluating the state of the system, but we may choose a different order size for each day of the week. Haijema and Minner (2016) examine several examples of constant order policies in a perishable setting. Formally:

$$X_j^{\text{COP}}(\mathbf{z}) = z_{j,k}, \qquad j \in \mathcal{J}. \tag{11}$$

Correlated base-stock policy. The presence of substitutions between similar products in our model calls for a modification of the traditional base-stock policies, where the interaction among items is entirely disregarded. In order to keep complexity to a manageable level, we can define a parameter vector $\mathbf{z}$, similarly to the BSP case, by aggregating the state components for all of the products subject to substitution. In our specific case, since all of them are vertically differentiated, we shall compute:

$$X_j^{\text{BSPC}}(S_t|\mathbf{z}) = \left(z_{j,k} - \sum_{r=1}^{J}\left(\sum_{l=1}^{\text{LT}_r} O_{t,r}^l + \sum_{d=1}^{\text{SL}_r} I_{t,r}^d\right)\right)^+, \qquad j \in \mathcal{J}, \tag{12}$$

where $r \in \{1,\dots,J\} = \mathcal{J}$ ranges over products that can be substitutes of each other.

Semi-seasonal policies. In the preliminary work reported in Gioia et al. (2022), a Soft Actor-Critic technique (Haarnoja et al., 2018) is adopted, based on artificial neural networks to approximate both a greedy value function and a direct policy. A policy was obtained, achieving similar performance as the BSP policy, even though it tracked seasonality with only one product out of two, keeping the order quantity for the other one constant. Based on this finding, we suggest here a similar idea, based on a simple parameterized ordering policy, which is computationally much cheaper to fine tune. At present, the approach has only been implemented for the case of two products (A and B), but it can be generalized when it is only possible to manage seasonality by a single product, whereas others arrive in fixed batches due to specific contracts or constraints on the supply chain. Note that, unlike the COP policy, a single order size is applied across the days of the week, for items whose supply is inflexible, disregarding seasonality. Indeed, it is worth noting that an understudied research problem in the inventory of perishable items literature concerns contracting (Karaesmen et al., 2011). Specifically, in the case in which an easily controllable private label coexists with an externally supplied alternative, a constant quantity for a contract with the supplier would be easier to manage, especially when a continuous sharing of information about the observed state of the system is not feasible. In our setting, featuring two products A and B, we will consider a first case in which orders for A can be adapted to the seasonal pattern, wheres product B is supplied by an external producer with constant order sizes (CB_BSPA), which suggests a possible interpretation of A as a private label item. Symmetrically, and for the purpose of studying business scenarios featuring different margins for the product replenished by seasonally adapted orders, we next consider the case where orders for B are seasonally adapted, whereas A can only be ordered according to a constant policy (CA_BSPB). In order to specify the CB_BSPA policy, we define:

$$X_j^{\text{CB\_BSPA}}(S_t|\mathbf{z}) = \begin{cases} z_B, & j = \text{B} \\ \left(z_{A,k} - \left(\left(\sum_l O_{t,A}^l + \sum_d I_{t,A}^d\right) + \sum_d I_{t,B}^d\right)\right)^+, & j = \text{A}. \end{cases} \tag{13}$$

The dimension of the vector $\mathbf{z}$ is further reduced and it now only takes into account:

- a single constant value for the each product ordered in constant quantities, regardless of the day of the week;
- $K+1$ values for the each product that allows for seasonally adapted orders.

Hence, a policy where one item is seasonally managed and the other ones are not, involves a parameter vector $\mathbf{z} \in \mathbb{R}^{J+(K+1)}$.

Table 1
Weekly seasonality pattern employed.

Weekday k	Mon 0	Tue 1	Wed 2	Thr 3	Fri 4	Sat 5	Sun 6
Seasonality Factor (η_k)	0.68	0.76	0.76	0.76	0.99	1.52	1.52

When B allows for seasonally adapted orders and A does not, the symmetric policy is:

$$X_j^{\mathtt{CA\_BSPB}}(S_t|\mathbf{z}) = \begin{cases} \left(z_{B,k} - \left(\left(\sum_l O_{t,B}^l + \sum_d I_{t,B}^d \right) + \sum_d I_{t,A}^d \right) \right)^+, & j = \mathrm{B}. \\ z_A, & j = \mathrm{A} \end{cases} \tag{14}$$

4.3. Optimization of policy parameters

All of the policies we presented so far depend on a vector of policy parameters $\mathbf{z}$, whose dimension depends on the complexity of the associate observed state. We aim to maximize the average daily reward (9) for a determined time horizon, where we simulate the inventory replenishment policy. In practice, whatever the parametric policy is, we aim at solving the maximization problem

$$\max_z \mathbb{E} \left\{ \frac{1}{T} \sum_{t=1}^{T} C(X_1^\pi(S_t|\mathbf{z}), \dots, X_J^\pi(S_t|\mathbf{z}), L_t) \right\}. \tag{15}$$

Note that, unlike other model formulations, there is no discount factor involved, as we consider average reward. We leave the time horizon T unspecified, as we will rely on sample averages over some suitably long horizon. A thorough comparison of alternative computation strategies to optimize (15) is not the focus of this work. Instead, we selected an off-the-shelf global optimization solver implemented by Eriksson et al. (2019), based on a stochastic surrogate optimization method proposed by Regis and Shoemaker (2007), relying on radial basis functions.

5. Simulation experiments

The experimental section of this work will focus on a particular case study that concerns only two products. However, the open-source framework may handle as many products as desired. The design of experiments relies on the four business scenarios specified in Table 2, where we divide the features into those related to market economics and those concerning lead-time and shelf-life. We keep a fixed setting for the distribution of demand and the stochastic parameter θ in the discrete choice model, while varying the market economics of the products, following some theoretical insights that hold in a non-perishable context (Pan and Honhon, 2012; Transchel et al., 2021) and help us to test meaningful configurations. We elaborate on this topic in Section 5.1, to justify our experimental design in terms of business and market settings. As to demand uncertainty, recalling the seasonality transformation (1), we assume a Poisson distributed demand for each day k within the week:

$$D_k \sim \mathrm{Poiss}(\mu_k), \qquad k \in \{0, 1, \dots, 6\},$$

where we set $\mu = 300$ for the whole simulation, applying the seasonality factors η_k of Table 1. Concerning the random utility, the distribution of the stochastic parameter θ_n is beta,

$$\theta_n \sim \mathrm{Beta}(2, 3),$$

where n refers to the nth simulated consumer visiting the store in the simulated time horizon. These random variables are independent and identically distributed across consumers.

Table 2
Design of experiments. For each product, a shelf life SL and a lead time LT are set. Furthermore, the market economics are selected (purchase cost c, selling price p, and corresponding margin m). The perceived quality q is also specified, for each value of remaining shelf-life. The first scenario has the same margin for fresh items of both products, while the other ones are more favorable to product A.

Scenario	Product	Market economics		Other features
1	A	p = [6,6,6,6] c = 4	q = [22.5,23,23.5,24] m = [2,2,2,2]	SL = 4 LT = 3
	B	p = [4,4] c = 2	q = [18,20] m = [2,2]	SL = 2 LT = 2
2	A	p = [6,6,6,6] c = 3	q = [22.5,23,23.5,24] m = [3,3,3,3]	SL = 4 LT = 3
	B	p = [4,4] c = 2	q = [18,20] m = [2,2]	SL = 2 LT = 2
3	A	p = [6,6,6,6] c = 3	q = [22.5,23,23.5,24] m = [3,3,3,3]	SL = 4 LT = 3
	B	p = [3.3,4] c = 2	q = [18,20] m = [1.3,2]	SL = 2 LT = 2
4	A	p = [5,6,6,6] c = 3	q = [22.5,23,23.5,24] m = [2,3,3,3]	SL = 4 LT = 3
	B	p = [3.3,4] c = 2	q = [18,20] m = [1.3,2]	SL = 2 LT = 2

5.1. Insights on the selected business scenarios

The purpose of this section is twofold: on the one hand, we want to shed more light on consumer behavior under the specified discrete choice model of Section 3.2; on the other one, we want to justify the selection of numerical parameters in our simulation experiments.

The selected business scenarios are outlined in Table 2, where we report the shelf life SL and the lead time LT, alongside with purchase cost c, selling price p, and the resulting margin m. Of particular interest is the perceived quality q, which decays with age. In business scenarios 1 and 2, the decay is not compensated by a reduction in price. On the contrary, discounts are offered in business scenarios 3 and 4. Though our aim is neither optimal pricing nor assortment planning, it is important to explore the effect of a discount by varying prices. Furthermore, we set the margin of the products to the same value as the first business scenario, while modifying the cost of product A in the other ones to make it the most economically profitable one.

No-discount business scenarios. Taking advantage of the linearity of the discrete choice model, it is possible to visualize the interaction of the price and the perceived quality of each product with the consumer's random utility coefficient θ. In particular, in Fig. 2(a) we illustrate graphically business scenarios 1 and 2, which differ in the cost of product A, but are identical from the consumers' viewpoint. We plot utility as a function of θ, and observe a reduction in the slope of the utility for each product when the *Residual Shelf Life* (RSL) decreases. There are four red lines for product A and two blue lines for product B. Specifically, product A will have an initial perceived quality value of $q_A^{\mathrm{SL}_A} = 24$, which gradually decreases to 22.5 as the remaining life decreases. On the other hand, B, being a lower margin and lower quality product, has an initial value of 20, which reaches 18 before it is discarded. If we disregard the stock-age adjustment, by considering only brand new items and the two lines with the largest slope for each product, the linear utilities for products A and B are

$$U_A = 24\theta - 6, \qquad \text{and} \qquad U_B = 20\theta - 4,$$

respectively. These two lines intersect for $\theta = 0.5$, where a consumer is indifferent between the two alternatives. Moreover, the utility for product B is zero when $\theta = 0.2$; hence, the consumer will not buy anything when $\theta < 0.2$. Therefore, we may represent areas corresponding to each choice, under the probability density function (PDF) of the stochastic parameter θ (which is beta distributed, with support

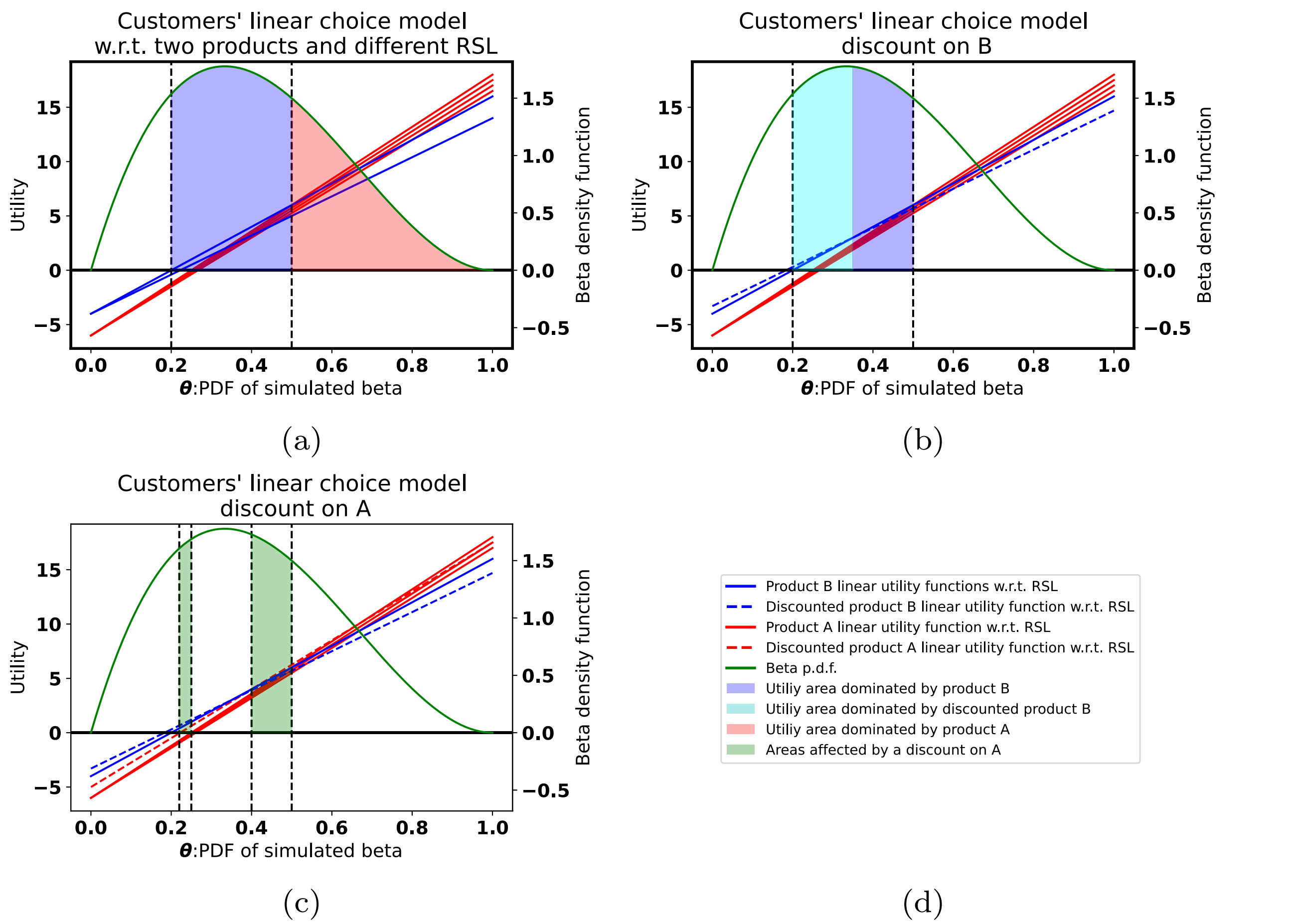

Fig. 2. Linear interactions between available products and the utilities of the consumers sampled from a Beta distribution. (a) no discount strategies (scenario 1 and 2); (b) discount on B (scenario 3); (c) discount on both A and B (scenario 4); (d) legend of the plot.

on the interval $[0,1]$): in the red area, product A dominates B, and roles are reversed in the blue area. The blank area corresponds to the no-purchase region, where both products have negative utility. If we still disregard deterioration, it is possible to compute the choice probabilities analytically (Pan and Honhon, 2012):

$$\mathbb{P}(\text{Choose B}) = F_\theta\left(\frac{p_A^{\mathrm{SL}_A} - p_B^{\mathrm{SL}_B}}{q_A^{\mathrm{SL}_A} - q_B^{\mathrm{SL}_B}}\right) - F_\theta\left(\frac{p_B^{\mathrm{SL}_B}}{q_B^{\mathrm{SL}_B}}\right)$$

$$= F_\theta(0.5) - F_\theta(0.2) \approx 50\%,$$

$$\mathbb{P}(\text{Choose A}) = 1 - F_\theta\left(\frac{p_A^{\mathrm{SL}_A} - p_B^{\mathrm{SL}_B}}{q_A^{\mathrm{SL}_A} - q_B^{\mathrm{SL}_B}}\right) = 1 - F_\theta(0.5) \approx 30\%,$$

$$\mathbb{P}(\text{No purchase}) \approx 20\%,$$

where $F_\theta(x) \doteq \mathbb{P}(\theta \leq x)$ denotes the cumulative distribution function (CDF) of the selected beta distribution for the random coefficient θ. A similar calculation can be carried out when the stock-age adjustment is accounted for and products are not necessarily brand new. The procedure should consider all of the intersections of the linear utilities, but it is also worth noting that, unless a discount policy is applied, the consumer will always choose the freshest item, if available.

In Pan and Honhon (2012, Lemma 1), the authors present the conditions under which products may be included in an optimal assortment, for preferences described by a linear discrete choice model and assuming no deterioration in items. In a case with two items, these conditions are:

$$p_B^{\mathrm{SL}_B} < p_A^{\mathrm{SL}_A}, \tag{16}$$

$$0 \leq \frac{p_B^{\mathrm{SL}_B}}{q_B^{\mathrm{SL}_B}} < \frac{p_A^{\mathrm{SL}_A}}{q_A^{\mathrm{SL}_A}} < 1, \tag{17}$$

$$m_B^{\mathrm{SL}_B} \leq m_A^{\mathrm{SL}_A}. \tag{18}$$

Although we satisfy conditions (16)–(18) when deterioration is disregarded, we relax them when adjusting for perishability. One reasonable assumption that we follow is a non-dominance condition within the prices of a single product, i.e.,

$$p_j^1 \leq \cdots \leq p_j^{\mathrm{SL}_j}, \qquad j \in \mathcal{J}. \tag{19}$$

However, due to discount strategies on aged items, product prices can now intersect according to complex patterns, making analytical conclusions more difficult to state.

When deterioration is considered, analytical conditions on optimal assortment may be difficult to meet, due to the possible inability to get rid of aged items in a profitable way. Furthermore, the interaction with demand seasonality, delivery lead times, and possible stockouts make the overall picture complex, calling for a simulation-based optimization approach. For example, the value of the flexibility of a product with a shorter lead-time and the adverse effect of a shorter shelf-life are extremely difficult quantities to assess.

Discount business scenarios. The third and the fourth business scenarios assume discount policies when the quality of the products deteriorates because of the ageing process. When no discounts are applied, as shown in Fig. 2(a), preferences are not so different from those of a LIFO consumer behavior within each product category, except for the interaction with other products (Gioia et al., 2022). On the contrary, if we apply a discount, we split the preference regions between new and old items within the same product.

In scenario 3, we only discount product B, and we can graphically appreciate the consequent effects in Fig. 2(b). In particular, we set:

$$p_B^1 = 4 \to p_B^{1,\text{disc}} = 3.3,$$

splitting the choice regions between the two possible residual shelf-lives of the items within product B. The new linear utility for the aged product B is now subject to an upward shift, together with the slope variation due to the deterioration, which leads to a different interaction with the beta sampled consumers' utility. The discounted product B in the third scenario eliminates the utility domination of the freshest product B over the aged one. We also notice that, thanks to the discount strategy, the optimal assortment conditions (16)–(18) now hold within the B product:

$$\begin{aligned} p_B^{1,\text{disc}} &< p_B^2, \\ 0 \le \frac{p_B^{1,\text{disc}}}{q_B^1} &< \frac{p_B^2}{q_B^2} < 1, \\ m_B^{1,\text{disc}} &\le m_B^2. \end{aligned}$$

The overall effect resembles a half-FIFO/half-LIFO mixture, but we have to bear in mind the interaction effect with other products, which generates new substitution patterns. In Fig. 2(b), we divide the original blue area of Fig. 2(a) into two new regions, where either the 1-RSL or the 2-RSL item corresponding to product B (i.e., product B with a residual shelf-life of 1 and 2 days, respectively) now dominates. Analytically, we can compute the new area (which is related with choice probabilities) just like before:

$$\begin{aligned} \mathbb{P}(\text{Choose B/2-RSL}) &= F_\theta\left(\frac{p_A^{\mathrm{SL}_A} - p_B^2}{q_A^{\mathrm{SL}_A} - q_B^2}\right) - F_\theta\left(\frac{p_B^2 - p_B^1}{q_B^2 - q_B^1}\right) \\ &= F_\theta\left(\frac{1}{2}\right) - F_\theta\left(\frac{0.7}{2}\right) \approx 25\%, \\ \mathbb{P}(\text{Choose B/1-RSL}) &= F_\theta\left(\frac{p_B^2 - p_B^1}{q_B^2 - q_B^1}\right) - F_\theta\left(\frac{p_B^1}{q_B^1}\right) \\ &= F_\theta\left(\frac{0.7}{2}\right) - F_\theta\left(\frac{3.3}{18}\right) \approx 25\%. \end{aligned}$$

Hence, this choice of values allows us to study the effects of discounting for product B, when the two demand fractions are roughly equal and sufficiently large to have an impact on the adopted replenishment policies. It should also be pointed out that, from a practical point of view, it is quite complex to estimate the parameters of the discrete choice models accurately. This difficulty questions the validity of the multi-product intersected areas when they are too small (e.g., really slow quality deterioration) and the consumers' choices depend on very small differences in utility. Possible approaches to deal with unreliable data and parameters have been discussed in Gioia et al. (2022). Here, we assume reliable data and, in the simulations, consumers' choices are made according to the exact utility values of available items.

In business scenario 4, we consider the application of a discount for both products. We apply a reduced price to the oldest items of product A as well, affecting B in the green areas of Fig. 2(c). The aged A, when available, dominates the freshest B in a new portion of the utility distribution, accounting for roughly an additional 16% of the demand. However, such an aged item of product A cannot be ordered by the retailer, since only fresh items are delivered by the suppliers, according to our assumptions. As we will see later in the experimental results, the most significant effect is the creation of a lower threshold on the minimum positive utility (i.e., the smallest value of θ for which the consumer purchases something):

$$\min\left\{\frac{p_j^d}{q_j^d} \ge 0 \text{ s.t. } d \in \{1,\dots,\mathrm{SL_j}\}, j = A\right\} \\ =: \theta_{\min}^A = \frac{6}{24} = 0.25 \to \frac{5}{22.5} \approx 0.22$$

In particular, since the consumer buys an item when its utility is positive, the substitution between similar products can occur as long as there exists an item that has a positive utility for the consumer, if her preferred item is not available. If the zero-utility crossing thresholds of the products are closer (the one of product B is $4/20 = 0.20$), the retailer can force the substitution more easily, to the point of generating an artificial modification of the assortment. Once the lost demand diminishes due to possible substitutions, it is conceivable to reduce the ordered amount of B, being the low-profitable (but easy to substitute) product. Despite that, all these observations are subject to complex interactions with lead-time, shelf-life, and seasonality, so that assessing the performance of replenishment policies requires a careful simulation experiment.

5.2. Performance of the replenishment policies

We have evaluated all of the policies presented in Section 4.2, solving problem (15) by simulation-based optimization. In order to tune the decision parameters in each policy, we adopted a simulation horizon of 60 weeks (420 steps), setting five different seeds for pseudo-random number generation for each business scenario of Table 2. Then, we tested each optimized policy out of sample, on a much longer horizon of 600 weeks (4200 steps). The key performance measures are:

- The *Average profit* per day (Avg.Profit), computed by averaging the cumulative profit of Eq. (9) over the test horizon.
- The *Average waste* per day (Avg.Waste), computed by averaging the number of scrapped items per day over the test horizon.
- The cumulative *Unmet Demand*, computed by summing the number of clients to whom nothing was offered over the test horizon.

As far as unmet demand is concerned, we must point out that it is something different from a common lost sale variable. Many discrete choice models allow for a no-purchase option that may be selected when all the available options feature a negative utility. Here, unmet demand only records when the inventory is empty and no choice can be made at all. It can be considered as a metric of how much is available on the shelves, and a lower value would suggest an higher number of stored items. Since well-replenished shelves can be considered as standard of quality for some firms (Akkaş and Gaur, 2022), getting a glimpse about such stock-outs may be insightful.

We provide the mean value and the standard deviation for all the selected metrics, computed with respect to the five replications for all of the policies and all of the business scenarios in Table 3. We also plot the size of the orders issued by the policies on each business scenario for the same three weeks (11–14 out of 600), on the same test conditions in Fig. 3. The solid line is the mean across the five seeds, while the area identifies the 95% confidence interval.[3]

The case of identical margins. In the first business scenario, gathering more detailed state information (however aggregate) does not seem to provide any advantage with respect to an open-loop, constant ordering policy. The best rule in terms of average profit (Table 3) and average waste is the constant policy `COP`, suggesting that the using knowledge about state of the system, as done in `BSP`, could even have an adverse effect on the replenishment choices in some settings. The order sizes for the `COP` policy (Fig. 3) resemble what we would expect from the analytical insights. Orders follow the seasonal pattern and cover about 50% of the demand with product B, while the rest is met with A. The analysis of the number of peaks and their size is less straightforward because they are subject to the joint effects of lead-time and shelf-life.

Concerning the other policies, `BSPC` is the best in terms of unmet demand. This follows from a better information about what is available in all of the inventories. `CA_BSPB`, since seasonality is ignored when

[3] Since these confidence intervals are obtained with five replications, they should be taken with care and are presented just to provide a glimpse of the underlying uncertainty.

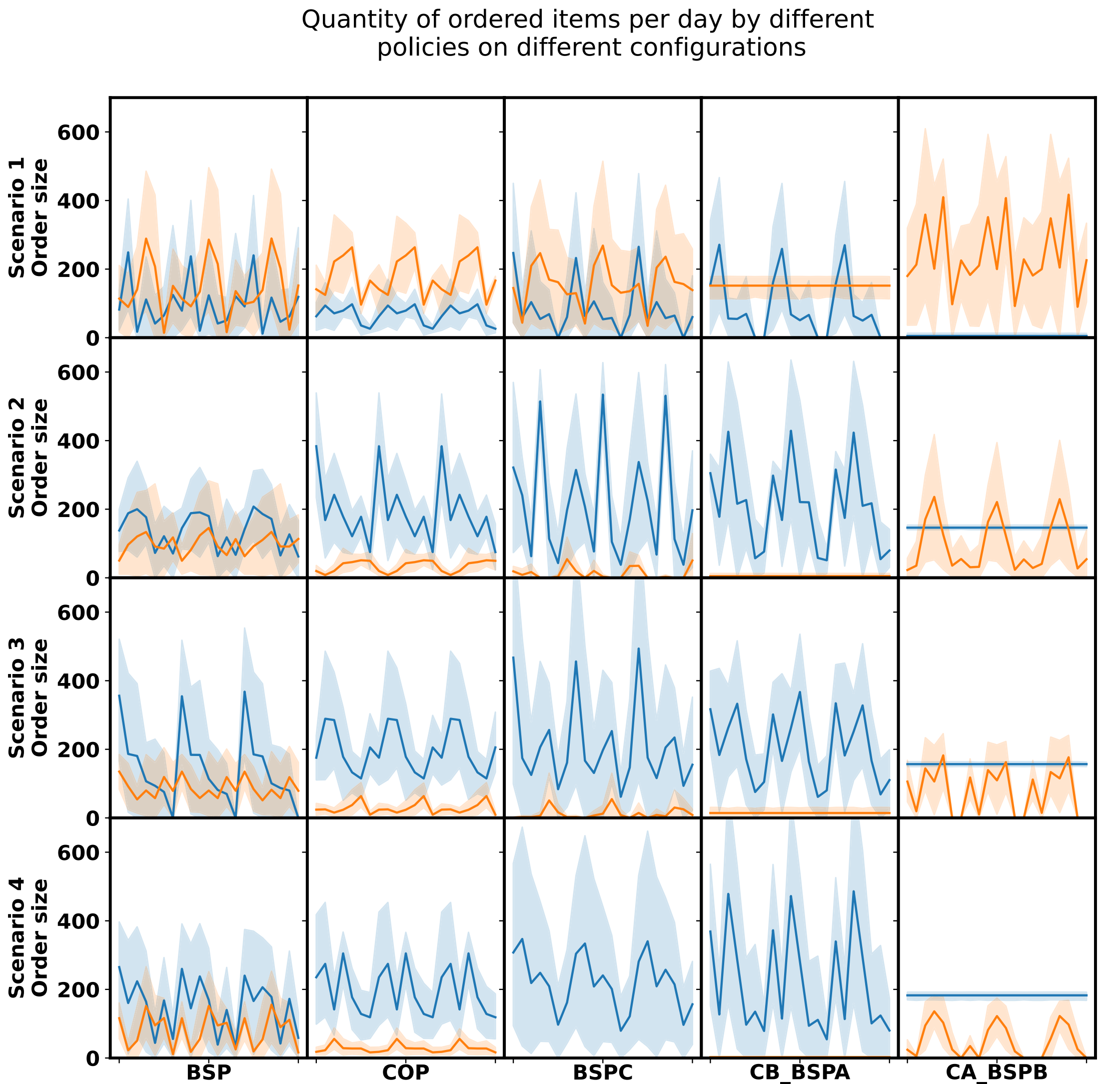

Fig. 3. Size of the orders issued by different policies on the four investigated business scenarios. Product A is colored in blue, product B in orange.

Table 3
Quantitative results of the heuristic policies on the four investigated scenarios. For each metric, we report the mean value and the standard deviation computed on five different seeds.

	Scenario		BSP	COP	BSPC	CB_BSPA	CA_BSPB
Same initial margin	1	Avg.Profit	452.53 ± 4.86	466.36 ± 2.34	457.12 ± 2.91	450.34 ± 10.04	450.18 ± 2.38
		Avg.Waste	5.27 ± 1.51	4.77 ± 0.92	4.49 ± 1.29	5.12 ± 3.20	8.53 ± 2.04
		Unmet demand	18592 ± 3638	18414 ± 4810	16448 ± 4074	32617 ± 14726	34573 ± 9210
Different initial margin	2	Avg.Profit	583.17 ± 14.23	618.88 ± 4.61	627.64 ± 7.42	623.61 ± 8	587.15 ± 10.24
		Avg.Waste	4.34 ± 2.21	3.35 ± 0.9	4.03 ± 0.82	1.47 ± 0.33	7.15 ± 2.64
		Unmet demand	25442 ± 6955	19382.4 ± 6328	18423 ± 4824	51283 ± 8747	14835 ± 5017
	3	Avg.Profit	575.29 ± 6.98	623.28 ± 4.84	625.78 ± 9.98	623.28 ± 9.17	592.35 ± 10.14
		Avg.Waste	2.96 ± 1.16	3.27 ± 0.41	3.24 ± 0.77	2.31 ± 1.47	3.91 ± 1.67
		Unmet demand	26543 ± 11050	17641 ± 3036	21614 ± 4625	37059 ± 9856	20219 ± 9478
	4	Avg.Profit	592.86 ± 4.56	625.6 ± 7.57	640.47 ± 2.22	629.65 ± 11.46	609.54 ± 9.99
		Avg.Waste	0.06 ± 0.03	0.06 ± 0.04	0.04 ± 0.02	0.98 ± 1.15	0. ± 0.
		Unmet demand	15210 ± 1748	8960 ± 1530	6037 ± 2784	10634 ± 5034	4808 ± 2546

ordering product A and the lead time of product B is shorter, prefers to force a substitution in favor of product B by more nervous and numerous peaks, but it produces the highest amount of waste. Furthermore, both CA_BSPB and CB_BSPA tend to miss the peaks of demand of the product whose orders they cannot adapt to the seasonal pattern, due to a limited supply, thus generating the highest values of unmet demand.

A final comment regards the true average optimal profit, whose evaluation would require a computationally expensive dynamic programming approach. To get a feeling for the quality of the proposed policies, we can use a simple back-of-the-envelope calculation. We adapt a Poisson distribution to reflect seasonality, but since the seasonality factors add up to 7, the average number of consumers per week is the same as in a week with no seasonality, which implies an average of 300 consumers per day. Of these, 20% have no positive utility from any item, so $300.8 = 240$ consumers should buy an item. Since the margin is 2 for both items, this suggests that, on average, we should expect a daily profit of no more than 480. Clearly, this is just a rough-cut calculation, but a comparison with the results in Table 3 for business scenario 1 provides a glimpse about the quality of the adopted policies in the simulated setting.[4]

The case of different margins. When the cost of product A drops by one unit, as is the case in business scenarios 2, 3, and 4, it becomes the more profitable for the retailer. We can therefore expect that ordering policies will try to enforce the substitution of B with A whenever possible. In Fig. 3, the last three scenarios emphasize this effect and its interaction with each ordering policy. However, discounts in the third and fourth scenarios have more significant consequences on the performance (see Table 3).

The policies where the order size of one product is fixed (CB_BSPA and CA_BSPB) must rely on the other one to deal with seasonality. Since product A is the most profitable one, when its order sizes are flexible, the product B almost disappears in the second and the fourth scenario, having a larger but still marginal role in the second scenario, where it is the only one discounted. It is also worth noticing that, in the fourth scenario, product B has a reduced price, but the effect on A makes it a minor advantage. On the one hand, due to the full reliance on product A, CB_BSPA does not care about the consumers that are not willing to substitute A in the place of B. This generates a limited supply of the latter product that entails very high values of unmet demand, reducing the average waste though. On the other hand, CA_BSPB produces opposite results in terms of waste and unmet demand in the second and third scenarios, maintaining a large supply of A to encourage a substitution effect, but being forced to handle the seasonality only through B. It follows that waste increases and unmet demand decreases. The fourth scenario is somehow different, because of the discount on A that completely compensates the oversupply of A, reducing to basically zero the number of scrapped items per day. Overall, the discount strategy on product A does significantly affect the performance, especially in reducing waste, but also in boosting the average profit as well.

The Correlated Base-Stock Policy (BSPC) yields the best average profit in all of the scenarios where the margins are different. In particular, we immediately notice that this is the only policy where a seasonal pattern is not evident in B (for business scenarios 2–4). It rather appears to meet a fraction of demand for B that cannot be substituted by A, when needed, in the second and third scenarios, while removing the less profitable product from the assortment in the fourth scenario. Unlike COP and BSP, which do not consider a detailed information about inventories, it is better able to take advantage of possible substitutions, revealing a fundamental feature that was not noticeable in the first business scenario. Specifically, in these three scenarios, the BSP policy does an even worse job with average profit, while COP seems to be still a reasonable approach, which improves when the substitution opportunity diminishes (third scenario, due to discount on B), reducing the gap with BSPC.

Similarly to the first scenario, we can estimate an upper bound on the average performance by a simplified calculation. However, we have to bear in mind that the discount strategies produce more intricate effects. Considering 300 consumers a day on average, we assume that the retailer is able to force them to buy the product with the highest margin in the subset of products for which they have a positive utility. In our two products case, this means that the retailer can force consumers who are willing to buy A, since they this has positive utility for them, to buy this higher margin product, even though they might prefer B. Clearly, this is an idealization, as a physical retailer cannot personalize the assortment for each specific consumer. In the case of consumers that have negative utility for A, but positive for B, the retailer might offer B, which yields a lower margin but covers an additional demand portion. To quantify the profit in this setting, let us observe that product A crosses the zero-utility line at point

$$\theta_1 = \frac{p_A^{\mathrm{SL}_A}}{q_A^{\mathrm{SL}_A}} = \frac{6}{24} = 0.25.$$

Hence, the retailer would like to force all consumers with $\theta \geq 0.25$ to buy A. The demand portion corresponding to consumers who have positive utility from B but not A features a value of θ between this critical value θ_1 and value of θ where B crosses the zero-utility, i.e.,

$$\theta_2 = \frac{p_B^{\mathrm{SL}_B}}{q_B^{\mathrm{SL}_B}} = \frac{4}{20} = 0.2.$$

Hence, the retailer would like to offer only B to consumers with $\theta \in [0.2, 0.25)$. Given the cumulative distribution function of the beta distribution F_θ of the random coefficient θ, the ideal average profit in scenario 2, assuming again an average of 300 consumers per day, is

$$300 \cdot \left[3 \cdot \left(1 - F_\theta(0.25)\right) + 2 \cdot \left(F_\theta(0.25) - F_\theta(0.2)\right)\right] \approx 712.$$

In business scenarios 3 and 4, we should consider consumers who are only willing to buy an aged but discounted item. In scenario 3, a discount in offered on aged items of B, and the retailer is able to cover an additional portion of demand, consisting of consumers who are not willing to buy either A or fresh and more expensive items of B, but have positive utility from discounted B. The value of θ where utility from discounted B crosses the zero-utility line is

$$\theta_3 = \frac{p_B^{1,\mathrm{disc}}}{q_B^1} = \frac{3.3}{18} \approx 0.1833.$$

Hence, the retailer can collect a margin of 1.3 from consumers featuring $\theta \in [0.1833, 0.2)$, yielding an upper bound of

$$300 \cdot 1.3 \cdot \left(F_\theta(0.2) - F_\theta(0.1833)\right) + 712 \approx 722.$$

In scenario 4, a discount is also offered on aged items of A but, since their margin is the same as fresh items of B, the analysis of scenario 3 can be applied to scenario 4 as well, from the retailer's point of view. These upper bounds can be compared with the performance of replenishment policies in Table 3, and we may observe a larger gap than in scenario 1. However, we should observe that here the assumption of a retailer able to customize assortment for each specific consumer is strongly idealized, and the interaction with seasonality and delivery lead time, as well as demand variability, makes this ideal target very hard to achieve.

6. Conclusions and directions for further research

In this paper we have considered parameterized ordering policies for a rather stylized inventory control problem with two vertically differentiated items, under a choice model based on a simple linear utility. Despite this simplicity, the model deals with complicating factors like demand seasonality, inventory deterioration, discounted prices, delivery lead times, and consumer heterogeneity. The ensuing complexity does not make the problem amenable to elegant analytical modeling, and we have pursued a simulation-based approach, relying on standard surrogate models. The aim is to gain managerial insight and to check the behavior of alternative ordering policies, rather than

[4] A careful comparison with dynamic programming has been reported, in a simpler setting, by Haijema and Minner (2019).

actually providing decision support. Nevertheless, the accompanying software allows for a far more general analysis.

The adoption of simple parameterized policies may be justified on the ground of managerial acceptance. Furthermore, variability in order sizes may be a source of concern when interacting with a supplier, especially in a setting with demand seasonality and short shelf-life. Indeed, a retailer may have to deal with items provided by outside suppliers as well as private label items, featuring different degrees of ordering flexibility. The analysis of the proposed rules suggests a potential approach to cope with seasonality by taking advantage of product substitutions.

The simulations rely on a discrete choice model based on a linear utility trading off price and quality. Whatever choice model we adopt, there are two sources of modeling errors:

1. *Model misspecification*, i.e., the model is not able to capture actual consumer behavior.
2. *Estimation errors*, i.e., the parameters characterizing the model are not quantified correctly.

The proposed approach is clearly subject to both sources of difficulty, which may be even more relevant when discounts are used to counter the LIFO consumer behavior in the face of deteriorating items. However, we do not have a descriptive aim, but rather a prescriptive one. The investigated approach is parsimonious and, hopefully, accurate enough to derive a satisfactory ordering policy. A fundamental step for further investigation should be a check of robustness against model errors. Furthermore, we have assumed given prices, but the model could be used for pricing purposes as well (Pan and Honhon, 2012). Since the interaction between complex problem features and joint ordering and pricing policies is not easy to analyze, a simulation-based optimization approach would be a useful tool. The shared code accompanying the paper may also be useful to pursue such an investigation.

CRediT authorship contribution statement

Daniele Giovanni Gioia: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Leonardo Kanashiro Felizardo:** Data curation, Investigation, Software. **Paolo Brandimarte:** Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Data availability

The code that supports the findings of this study is openly available at https://github.com/DanieleGioia/PerishableDCM.

Acknowledgments

A preliminary version of this manuscript was presented at the 10th IFAC Conference on Manufacturing Modelling, Management and Control, June 22–24, 2022, Nantes, France. We would like to thank the reviewers of The International Federation of Automatic Control (IFAC) for the comments on the preliminary work

Funding

This research was financed in part by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (CAPES - Coordination for the Improvement of Higher Education Personnel, Finance Code 001, grant 88882.333380/2019-01), Brazil

References

Akkaş, A., Gaur, V., 2022. OM Forum—Reducing food waste: An operations management research agenda. Manuf. Serv. Oper. Manag. 24, 1261–1275.
Bakker, M., Riezebos, J., Teunter, R.H., 2012. Review of inventory systems with deterioration since 2001. European J. Oper. Res. 221, 275–284.
Brandimarte, P., 2021. From Shortest Paths to Reinforcement Learning: A MATLAB-Based Tutorial on Dynamic Programming. Springer.
Brodheim, E., Derman, C., Prastacos, G., 1975. On the evaluation of a class of inventory policies for perishable products such as blood. Manage. Sci. 21, 1320–1325.
Buisman, M.E., Haijema, R., Hendrix, E.M., 2020. Retailer replenishment policies with one-way consumer-based substitution to increase profit and reduce food waste. Logist. Res. 13 (7), 1–15.
Buzby, J., Wells, H., Axtman, B., Mickey, J., 2009. Supermarket Loss Estimates for Fresh Fruit, Vegetables, Meat, Poultry, and Seafood and Their Use in the ERS Loss-Adjusted Food Availability Data. Economic Information Bulletin 58313, United States Department of Agriculture, Economic Research Service.
Deng, Y., Zhang, X., Wang, T., Wang, L., Zhang, Y., Wang, X., Zhao, S., Qi, Y., Yang, G., Pengb, X., 2023. Alibaba realizes millions in cost savings through integrated demand forecasting, inventory management, price optimization, and product recommendations. Inf. J. Appl. Anal. 53, 32–46.
Eriksson, D., Bindel, D., Shoemaker, C.A., 2019. Pysot and POAP: An event-driven asynchronous framework for surrogate optimization. http://dx.doi.org/10.48550/ARXIV.1908.00420, arXiv preprint.
Ferguson, M.E., Koenigsberg, O., 2007. How should a firm manage deteriorating inventory? Prod. Oper. Manage. 16, 306–321.
Fisher, M., Raman, A., 2010. The New Science of Retailing: How Analytics are Transforming the Supply Chain and Improving Performance. Harvard Business Review Press.
Gioia, D.G., Felizardo, L.K., Brandimarte, P., 2022. Inventory management of vertically differentiated perishable products with stock-out based substitution. IFAC-PapersOnLine 55, 2683–2688, 10th IFAC on Manufacturing Modelling, Management and Control MIM 2022.
Goyal, S., Giri, B., 2001. Recent trends in modeling of deteriorating inventory. European J. Oper. Res. 134, 1–16.
Haarnoja, T., Zhou, A., Abbeel, P., Levine, S., 2018. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In: International Conference on Machine Learning. PMLR, pp. 1861–1870.
Haijema, R., 2014. Optimal ordering, issuance and disposal policies for inventory management of perishable products. Int. J. Prod. Econ. 157, 158–169.
Haijema, R., Minner, S., 2016. Stock-level dependent ordering of perishables: A comparison of hybrid base-stock and constant order policies. Int. J. Prod. Econ. 181, 215–225.
Haijema, R., Minner, S., 2019. Improved ordering of perishables: The value of stock-age information. Int. J. Prod. Econ. 209, 316–324.
Haijema, R., van der Wal, J., van Dijk, N.M., 2005. Blood platelet production: a multi-type perishable inventory problem. In: Operations Research Proceedings 2004. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 84–92.
Hendrix, E., Ortega Lopez, G., Haijema, R., Buisman, M., García Fernandez, I., 2019. On computing optimal policies in perishable inventory control using value iteration. Comput. Math. Methods 1, e1027.
Ishii, H., Nose, T., 1996. Perishable inventory control with two types of customers and different selling prices under the warehouse capacity constraint. Int. J. Prod. Econ. 44, 167–176.
Jalali, H., van Nieuwenhuyse, I., 2015. Simulation optimization in inventory replenishment: a classification. IIE Trans. 47, 1–19.
Janssen, L., Claus, T., Sauer, J., 2016. Literature review of deteriorating inventory models by key topics from 2012 to 2015. Int. J. Prod. Econ. 182, 86–112.
Kalpakam, S., Shanthi, S., 2001. A perishable inventory system with modified $(S-1, S)$ policy and arbitrary processing times. Comput. Oper. Res. 28, 453–471.
Karaesmen, I.Z., Scheller-Wolf, A., Deniz, B., 2011. Managing perishable and aging inventories: Review and future research directions. In: Kempf, K.G., Keskinocak, P., Uzsoy, R. (Eds.), Planning Production and Inventories in the Extended Enterprise: A State of the Art Handbook, Volume 1. Springer, US, pp. 393–436.
McFadden, D., 1973. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.), Frontiers in Econometrics. Academic Press, New York, pp. 105–142.
Nahmias, S., 1977. Higher-order approximations for the perishable-inventory problem. Oper. Res. 25, 630–640.
Nahmias, S., Pierskalla, W.P., 1976. A two-product perishable/nonperishable inventory problem. SIAM J. Appl. Math. 30, 483–500.
Pan, X.A., Honhon, D., 2012. Assortment planning for vertically differentiated products. Prod. Oper. Manage. 21, 253–275.
Pauls-Worm, K.G., Hendrix, E.M., Haijema, R., van der Vorst, J.G., 2014. An MILP approximation for ordering perishable products with non-stationary demand and service level constraints. Int. J. Prod. Econ. 157, 133–146.

Perry, D., Stadje, W., 2000. An inventory system for perishable items with by-products. Math. Methods Oper. Res. 51, 287–300.
Powell, W.B., 2021. Reinforcement Learning and Stochastic Optimization: A Unified Framework for Sequential Decisions. Wiley.
Regis, R.G., Shoemaker, C.A., 2007. A stochastic radial basis function method for the global optimization of expensive functions. INFORMS J. Comput. 19, 497–509.
Shin, H., Park, S., Lee, E., Benton, W., 2015. A classification of the literature on the planning of substitutable products. European J. Oper. Res. 246, 686–699.
Train, K.E., 2009. Discrete Choice Methods with Simulation. Cambridge University Press.
Transchel, S., 2017. Inventory management under price-based and stockout-based substitution. European J. Oper. Res. 262, 996–1008.
Transchel, S., Buisman, M.E., Haijema, R., 2021. Joint assortment and inventory optimization for vertically differentiated products under consumer-driven substitution. European J. Oper. Res. 301, 163–179.
Tsiros, M., Heilman, C.M., 2005. The effect of expiration dates and perceived risk on purchasing behavior in grocery store perishable categories. J. Market. 69, 114–129.
van Donselaar, K., Van Woensel, T., Broekmeulen, R., Fransoo, J., 2006. Inventory control of perishables in supermarkets. Int. J. Prod. Econ. 104, 462–472.
Van Woensel, T., van Donselaar, K., Broekmeulen, R., Fransoo, J., 2007. Consumer responses to shelf out-of-stocks of perishable products. Int. J. Phys. Distrib. Logist. Manage. 37, 704–718.