University of Tehran Press

Industrial Management Journal

Online ISSN: 3115 - 7386

Homepage: https://imj.ut.ac.ir

Integrated Assortment–Shelf Optimization under Substitution and Space Elasticity: A Hybrid Memetic Algorithm

Ali Ebrahimi Kordler✉ [1], **Saleh Mansouri** [2], **and Farzad Bahrami** [3]

1. Corresponding author, Associate Prof., Department of Accounting and Finance, College of Management, University of Tehran, Tehran, Iran. E-mail: aebrahimi@ut.ac.ir
2. Ph.D. Candidate, Department of Industrial Management, School of Management, Kish International Campus, University of Tehran, Kish Island, Iran. E-mail: saleh.mansouri@ut.ac.ir
3. Assistant Prof., Department of Industrial Management, Faculty of Administrative Sciences and Economics, Arak University, Arak, Iran. E-mail: f-bahrami@araku.ac.ir

Article Info

Article type:
Research Article

Article history:
Received January 11, 2026
Received in revised form February 21, 2026
Accepted March 14, 2026
Available online April 01, 2026

Keywords:
Assortment planning, shelf-space optimization, substitution effects, metaheuristics, local search.

ABSTRACT

Objective: This study maximizes expected retail profit by jointly optimizing product assortment and shelf-space allocation, considering substitution effects and space-elastic demand. The problem's NP-hardness, compounded by category-level bounds and store capacity, renders exact methods impractical for large-scale instances. Therefore, we develop a computationally efficient framework to generate near-optimal shelf plans aligned with real-world constraints.

Methodology: We propose a hybrid Memetic Algorithm embedded with Iterated Local Search (ILS), combining evolutionary global exploration with local hill-climbing refinement. A two-phase initialization ensures every candidate planogram satisfies capacity constraints. Each chromosome encodes item-to-shelf mappings, with facings emerging endogenously. Mechanisms, including crossover, mutation, and diversity control, preserve solution validity and mitigate premature convergence. The framework was validated using real data from the Iranian retail chain Ofoq Kourosh, encompassing 39 product categories and 21,000 cm of total shelf length.

Results: The algorithm consistently converged toward feasible solutions. Under the original category bounds, profit reached 3.04 million; relaxing these bounds improved profit by 6.2% to 3.25 million. Allocation outcomes aligned with demand elasticity: impulse-driven categories reached upper limits, while low-elasticity staples stabilized near minima. Pareto analysis confirmed that roughly 20% of categories generated over 80% of profit. Notably, the optimized solution resulted in a 37% increase in profit compared to the current store configuration.

Conclusion: Results confirm the efficacy of hybrid metaheuristics for complex retail optimization. The framework consistently achieved near-optimal solutions under realistic constraints. Managerially, shelf-space allocation should prioritize high-elasticity categories while maintaining a minimal representation of staples. Future research should extend this framework to multi-store and omni-channel contexts with dynamic demand modeling.

Cite this article: Ebrahimi Kordler, A., Mansouri, S. & Bahrami, F., (2026). Integrated assortment–shelf optimization under substitution and space elasticity: A hybrid memetic algorithm, ***Industrial Management Journal***, 18(2), 1-33. https://doi.org/10.22059/imj.2026.404854.1008268

Publisher: University of Tehran Press.

DOI: https://doi.org/10.22059/imj.2026.404854.1008268

Introduction

In contemporary grocery retail, assortment planning is high-stakes because shelf space is a scarce and expensive resource, and "a large number of products compete with each other to obtain more shelf space," making space-aware optimization central to sales and shopper satisfaction (Ziari & Sajadieh, 2025). At the same time, the operational scale is massive—a typical supermarket carries roughly 33,000 SKUs—so even small listing/delisting moves ripple across perception, operations, and performance (Sethuraman et al., 2022). Moreover, adding items indiscriminately can inflate decision effort, as larger assortments increase the information-processing load and choice difficulty, which may undermine the gains. Therefore, breadth must be curated to ensure navigation effort remains manageable (Pham et al., 2025).

Today's retailers must allocate scarce shelf space while coping with approximately 30% growth in the number of items compared to a decade ago, making assortment choices directly consequential for both profit and shopper experience (Bianchi-Aguiar et al., 2021). Empirical syntheses further show that "more variety" can overshoot: across 42 categories, item reductions of up to 54% delivered an average sales lift of +11 %, indicating that curated breadth can outperform maximal breadth. Crucially, substitution softens delisting risk (Hübner & Kuhn, 2024). Formal assortment theory under explicit substitution also shows the problem is NP-complete, and that capacity utilization and the identity of optimal SKUs depend on substitution levels, so naïve "add more" or greedy rules may fail (Çömez-Dolgan et al., 2021). Mohaghar et al. (2020) propose a four-echelon omni-channel supply chain model for a seasonal product under stochastic demand, demonstrating that centralized coordination yields higher overall profit compared to decentralized decision-making. Their findings highlight the strategic value of integrated planning in complex retail networks. On the implementation side, real planograms are bound by four classes of constraints—shelf, product, multi-shelf, and category—so fast heuristics are required at realistic sizes where exact solvers often time out. Contemporary decision support, therefore, integrates assortment, shelf-space, and replenishment with space-elastic demand and substitution. Computational tests on retail data show profit improvements resulting from this integration.

Moreover, because demand and elasticity inputs are often noisy or missing, Distributionally Robust Optimization (DRO) offers principled stability by treating uncertainty via discrepancy-based ambiguity sets that shrink toward the truth as data accumulate—making recommendations less brittle in practice (Kuhn et al., 2025). In a two-echelon vendor-managed inventory setting with Poisson-distributed retail demand, system costs are explicitly sensitive to the demand rate—higher demand levels lead to increased costs—underscoring how demand parameters directly influence inventory control and replenishment decisions (Haji et al., 2009).

Recent methodology has shifted decisively toward fast metaheuristics that jointly select the assortment and allocate facings under space-elastic demand and explicit substitution, as exact solvers scale poorly under real-world planogram constraints. A 2024 systematic review maps this trend and highlights GA-based and local-improvement hybrids (e.g., GA + tabu/ILS) as practical workhorses for large instances and integrated objectives (Heger & Klein, 2024). Within this stream, Hübner et al. (2020) optimize two-dimensional (tilted) shelves under stochastic demand, space elasticity, and substitution, using a specialized GA that achieves approximately 99% average solution quality compared to exact benchmarks and reports profit uplifts of up to 15% in a retail case. Omitting demand effects can significantly reduce profits, underscoring the importance of modeling substitution and price elasticity within the formulation rather than addressing them ex post. Complementing this, Czerniachowska et al. (2021) construct a GA around four families of merchandising constraints—shelf, shelf-type/level, product, and virtual segments with capping/nesting— demonstrating that a chromosome encoding facings, capping, and nesting, combined with a repair step, yields efficient and feasible planograms with short runtimes relative to CPLEX. Finally, Czerniachowska et al. (2023) propose lightweight, feasibility-first heuristics that prune the search space to admissible sequences only; across 45 instances, they find feasible solutions far more often than CPLEX and achieve, on average, 95% of the best observed profit with minute-level compute—evidence that simple, industrializable rules can rival heavy optimization under tight time budgets.

Our method operates at the category–shelf level, where retailers actually define assortments. Feasibility is governed by group-level minimum and maximum bounds, along with a strict store-wide capacity constraint, reflecting the real limits within which assortments must be selected. The profit objective combines direct contribution margins with substitution-driven recapture and space-elastic demand, while item-share inputs enter the objective as demand weights that shape the expected mix without acting as hard constraints. Since the problem is NP-hard, we employ a memetic algorithm that combines extensive genetic exploration with targeted local improvement, supported by repair and diversification mechanisms to maintain feasibility and diversity. The result is a pipeline that generates high-margin, operationally credible assortments aligned with category bounds and total store capacity.

The remainder of this paper develops and validates this framework as follows: the literature review synthesizes five decades of research—from early space-elasticity experiments to recent robust and omni-channel formulations—and identifies the methodological gaps that motivate our hybrid approach. The problem definition section formalizes the objective function, substitution structure, and capacity constraints, followed by a detailed exposition of the Memetic/ILS algorithm, including initialization, crossover, mutation, hill-climbing, and diversity control mechanisms. The computational experiments section presents results from four algorithmic configurations, utilizing real data from 39 product categories, to examine the impact of population size and constraint relaxation on solution quality and convergence behavior. Finally, the discussion interprets these findings in light of theoretical expectations and managerial implications, while acknowledging limitations and charting directions for future research.

Literature Background

Against this contemporary backdrop, it is helpful to trace how the literature has evolved. Research on assortment and shelf-space optimization has evolved over the past five decades, beginning with early empirical studies of space elasticity in the 1970s and gradually advancing toward integrated, stochastic, and robust optimization models. The shelf-space literature originates from in-store experiments documenting space elasticity. Early work demonstrated that allocating more facings to a product tends to increase its sales, albeit with diminishing returns. Even with manual data collection, regression analyses confirmed a systematic but tapering sales lift, laying the empirical foundation for subsequent space-elasticity models (Anderson & Amato, 1974; Curhan, 1972). Later studies began to couple product selection with space allocation in supermarkets, anticipating today's integrated planning perspective (Hansen & Heinsbroek, 1979). Marketing–operations research subsequently quantified the effects of horizontal and vertical positioning, demonstrating the practical magnitude of shelf-driven demand. For example, moving a product from the worst to the best horizontal location could materially increase sales (Dreze et al., 1994).

As the field transitioned from descriptive to prescriptive approaches, researchers embedded space-elasticity and positioning effects into decision models and showed how to linearize diminishing-returns profit curves to maintain solvability at retail scale (Hansen et al., 2010). Building on this foundation, Hwang et al. (2005) integrated vertical and horizontal shelf positions into a joint space-allocation and inventory-control model. Drawing on retailer case studies, they demonstrated that shifting a brand from lower to higher shelf levels could yield double-digit sales gains, and they proposed a genetic algorithm heuristic that scaled effectively to realistic category sizes. Concurrently, category-management surveys clarified the tactical division of assortment selection, shelf-space planning, and in-store replenishment, and advocated for their integration (Hübner & Kuhn, 2012).

Later reviews observed that assortment models often omit space elasticity, while shelf-space models typically ignore substitution effects—highlighting the need for integrated formulations (Hübner, 2011; Hübner & Kuhn, 2012). Kök and Fisher (2007) were pivotal in this regard: by combining assortment, substitution, and space elasticity in a newsvendor-type model, they demonstrated that ignoring substitution could result in a misestimation of profits by more than 10%. Their heuristic, tested on categories of up to 25 products, revealed both structural properties—such as a preference for smaller items under tight space constraints—and the limitations of earlier models.

Murray et al. (2010) extended this conversation into retail practice, demonstrating through interviews and surveys how shelf planning, assortment, and supplier negotiations are intertwined. Their results emphasized that optimization models must be reconciled with managerial judgment and planogram feasibility. Similarly, Baron et al. (2011) examined consumer responses to shelf configurations in experimental settings, demonstrating that block placements and adjacency cues could shift brand-switching probabilities by measurable margins. These findings underscored the importance of modeling not only elasticity but also perceptual effects.

Lotfi et al. (2011) advanced methodological integration by proposing a weighted goal programming model that jointly optimized replenishment planning and shelf allocation under budget, storage, and perishability constraints. In a supermarket case, their LINGO-based model showed that reallocating just 5% of shelf space between brands could improve service levels without reducing margins. Hübner and Kuhn (2012) further examined substitution effects within assortment optimization, while Farias et al. (2013) added a theoretical perspective by casting assortment optimization as a robust choice-modeling problem, providing bounds in the absence of complete substitution data. Together, these works pushed the literature toward more rigorous formulations.

A key advancement was the integration of assortment choice with shelf allocation under space-elastic demand and substitution, demonstrating how facings and vertical/horizontal placement jointly shape demand. This research stream also emphasized block placement and "planogram feasibility" constraints (Hübner & Schaal, 2017). Hübner et al. (2016) formalized this approach through a capacitated assortment model under stochastic demand and substitution. Using datasets of up to 50 products, they demonstrated that ignoring substitution could result in a reduction of expected profits by more than 15%. Their heuristic closed much of the computational gap left by earlier models.

Subsequent work extended the focus to two-dimensional shelves, emphasizing rectangular blocks, adjacency, and the need to co-optimize item selection, facings, assignment, and arrangement—not merely counts of facings (Hübner et al., 2020). Parallel studies examined revenue maximization under vertical placement effects, revealing the computational burden of realistic planogramming (Geismar et al., 2015). Bianchi-Aguiar et al. (2021) provided a comprehensive review, codifying decisions (space, vertical/horizontal location, facings), demand effects (elasticity, cross-space, position), and constraints (integer facings, capacity, bounds). They stressed that optimization-grade models still exceed the capabilities of commercial software. Hübner et al. (2020) further argued that two-dimensional shelves are intrinsically integrated—encompassing selection, quantity, assignment, and arrangement—and that one-dimensional heuristics are inadequate once vertical blocks and adjacencies become relevant.

The computational complexity of assortment optimization was clarified by Aouad et al. (2018), who proved tight hardness-of-approximation results under ranking preference models. Their reduction from the maximum independent set problem demonstrated that even approximate solutions can be intractable in the general case. However, simple revenue-ordered assortments also provide the best possible guarantees with respect to price parameters. This complexity lens helps explain why scalable heuristics and matheuristics dominate practice.

As instance sizes increased, the literature shifted toward model reduction, exact–approximate hybrids, and matheuristics. Leitner et al. (2024) showed that compact formulations and decomposition techniques can quickly close gaps in assortment MILPs, although exactness deteriorates beyond a few hundred items, making matheuristics and rounding essential. Accordingly, a two-phase mathematical approach for integrated assortment–shelf allocation achieved near-optimal solutions, with an average gap of less than 3.8%. The optimized plan concentrated profit in high-visibility areas, where the most visible families accounted for about 80% of total store profit (Abbaszadeh et al., 2025)

Brandimarte et al. (2024) advocated a reduce-and-solve matheuristic, solving manageable subproblems first and then invoking a commercial solver to keep industrial-scale problems tractable. Heger and Klein (2024) confirmed this trend by classifying contributions according to customer models, constraints, and solution concepts, and highlighting the rise of hybrid approaches that blend metaheuristics with MIP-based repair. In parallel, product portfolio research in supply chains has shifted beyond decisions to delete or revitalize products to focus on proactive performance preservation. Abbasnia et al. (2025) applied Inverted DEA sensitivity analysis to identify strategic items and derive improvement trajectories, showing that small input reallocations (e.g., 5%) can significantly enhance both margin and efficiency. Such approaches complement assortment optimization by informing which product groups should remain prioritized when space or capital constraints become binding (Abbasnia et al., 2025).

Where assortment is central, the realism of choice models is crucial. While the multinomial logit (MNL) model remains popular for its tractability, nested and mixed logit models, as well as Markov chain choice models (MCCM), capture richer substitution patterns (Désir et al., 2024). Agarwal et al. (2019) bridged these approaches with their "MNL-Bandit," a dynamic learning algorithm that simultaneously estimates choice parameters and maximizes revenue. Tested on hundreds of products, their policy achieved near-optimal regret bounds and avoided the poor performance of simpler explore-then-exploit methods, underscoring the feasibility of adaptive, data-driven assortment planning.

On the distributionally robust side, Yu et al. (2024) framed assortment optimization as a worst-case CVaR or chance-constrained problem over a moment set, deriving SOCP characterizations. They demonstrated that revenue-ordered solutions remain effective under realistic conditions, delivering substantial revenue increases and more stable worst-case performance. Most recently, Hense and Hübner (2022) extended the modeling horizon to omni-channel retail. Their integrated assortment–space–inventory model incorporated both in-channel and cross-channel substitution under stochastic, space-elastic demand. In test cases with up to 80 items across store and webshop, they found that coordinated omni-channel planning could significantly raise profits, with in-channel substitution effects typically dominating. Their heuristic consistently outperformed benchmarks and delivered near-optimal solutions—marking a first step toward tactical omni-channel planogramming.

Beyond static planograms, the field has expanded into online and personalized settings. Chen et al. (2024) studied recommendation-at-checkout with add-on items and demonstrated that a robust protection-in-expectation policy achieved provable competitive ratios under worst-case arrivals, performing strongly in large-scale simulations. For clarity, we complement this review with a summary table (Table 1) synthesizing the key studies discussed, along with a compact glossary of terms and a brief caption for the accompanying figure.

In summary, the literature demonstrates a clear evolution—from early empirical studies of space elasticity, to integrated assortment–shelf models, and more recently to robust and omni-channel formulations. Over time, the field has progressed from descriptive observations to prescriptive optimization, and now toward scalable, hybrid, and data-driven algorithms. While state-of-the-art research highlights the power of integrated, stochastic, and robust models, commercial tools still lag behind, leaving significant opportunities to bridge the gap between theory and practice in retail shelf-space and assortment planning.

Despite the progress, the existing literature still reveals several gaps at the interface of assortment optimization and shelf-space planning. Many models either abstract away from practical feasibility constraints (e.g., group-level bounds, total store capacity) or rely on solution methods that do not scale reliably to supermarket-sized instances. Moreover, although metaheuristics and matheuristics have gained traction, relatively little work has focused on hybrid approaches that explicitly balance exploration with feasibility repair and targeted local improvement. Our study addresses this gap by developing a memetic algorithm that operates at the category–shelf level, blending genetic exploration with repair and diversification to enforce feasibility, and hill-climbing for local refinement. Unlike prior models, our formulation integrates direct margins, substitution-driven recapture, and space-elastic demand within realistic capacity and grouping bounds—producing high-margin assortments that are both computationally efficient and operationally credible. In doing so, we position our contribution at the intersection of theoretical rigor and practical decision support for large-scale grocery retailing.

Table 1. Chronological map of the literature linking empirical space effects to integrated, substitution-aware assortment and 2D shelf-space optimization—highlighting demand models, constraints, algorithms, and the specific takeaway for a scalable, robust pipeline

Reference	Focus	Choice/Demand Model	Substitution	Space Elasticity/ Position	Key Constraints	Algorithm/ Method	Scale/Context	Key Finding
Agarwal et al. (2019)	MNL-bandit	MNL (learning)	Yes	—	Inventory/capacity (implicit)	UCB bandit	Online large-scale	Regret near lower bounds; learn assortment on the fly.
Anderson and Amato (1974)	Integrated selection+ space	Empirical/Analytical	Yes (brand-level)	Own	Shelf capacity	Optimization (early OR)	Small analytical	Early joint treatment of assortment and display space.
Aouad et al. (2018)	Hardness/ approximation	Rank-based choice	Yes	—	—	Complexity results	—	The assortment problem is often NP-hard to approximate tightly.
Baron et al. (2011)	Inventory-dependent demand	Inventory-sensitive	Partial	Own	Service levels; shelf stock	Analytical + algorithms	Category-scale	Ignoring stock dependence harms profit.
Bianchi-Aguiar et al. (2021)	Shelf-space review	Survey	—	—	Min/max; capacity; position	Synthesis	—	Codifies decision types/effects/constraints; urges scalability.
Brandimartet al. (2024)	Model reduction + metaheuristic	General retail	—	—	Problem-specific	Reduce-and-solve (MIP repair)	Industrial	Reducing instances and then repairing with MIP

Reference	Focus	Choice/Demand Model	Substitution	Space Elasticity/ Position	Key Constraints	Algorithm/ Method	Scale/Context	Key Finding
								improves both time/quality.
Chen et al. (2024)	Check out the add-on assortment	Choice + inventory	Yes	—	Inventory protection	Robust online policy	Large-scale sims	Robust protection-in-expectation boosts worst-case performance.
Curhan (1972)	Empirical space elasticity	Empirical	Implicit	Own (diminishing returns)	—	Field study; regression	Supermarket categories	More space boosts sales with diminishing returns, exhibiting heterogeneous effects.
Désir et al. (2024)	Robust assortment under MCCM	MCCM	Yes		Uncertainty sets	Max–min duality + algos	Large choice sets	Efficient robust algorithms under MCCM.
Drèze et al. (1994)	Positioning & space elasticity	Empirical	Implicit	Own + Position (H/V)	—	Field experiments	Multiple categories	Horizontal/vertical position often dominates extra facings.
Farias et al. (2013)	Nonparametric/ robust revenue	Nonparametric choice	Yes	—	—	Robust optimization	Data-limited	Worst-case revenue over consistent choices; reduces overfitting.
Geismar et al. (2015)	2D shelf revenue	Demand with vertical	Partial	Own + Vertical	Rectangular blocks	IP + graph reduction	Hundreds	Vertical effects & 2D packing are critical.
Hansen and Heinsbroek (1979)	Selection+ space in supermarkets	Analytical	Implicit	Own	Capacity; integer facings	Mathematical programming	Store category	Selection and facings are intertwined decisions.
Hansen et al. (2010)	Shelf allocation methods	Space-elastic profit	Implicit	Own + Position	Capacity; integer facings	Heuristics vs metaheuristics	Retail-scale	Metaheuristics outperform basic heuristics and linearization tricks.
Hense (2022)	Category management trends	Conceptual/empirical	Partial	Own	Policy/role constraints	Review/cases	Retail practice	Role-based category rules shape feasible space/assortment.
Hübner and Schaal (2017a)	Integrated assortment + space	Space-elastic	Yes	Own + Cross-space	Min/max; capacity	MIP + heuristics	Large categories	Joint selection + facings+positioning improves outcomes.
Hübner and Schaal (2017b)	Stochastic, space-elastic	Stochastic demand	Yes	Own	Capacity; integer	MIP + sampling	Retail-scale	Uncertainty changes optimal facings/arrangements.
Hübner et al. (2020)	2D shelves—	Space-elastic + position	Partial	Own + Position	2D rectangular; adjacency	Optimization frameworks	Large planograms	Selection, quantities, assignment, and

Reference	Focus	Choice/Demand Model	Substitution	Space Elasticity/ Position	Key Constraints	Algorithm/ Method	Scale/Context	Key Finding
	integrated view							arrangement are inseparable.
Hwang et al. (2005)	Space+ inventory + position	Demand with position	Partial	Own + Vertical	Inventory, planogram	GA + gradient	Category-scale	Joint space-inventory-location improves profit vs. siloed.
Kök and Fisher (2007)	Assortment under substitution	MNL & variants	Yes	—	Cardinality/capacity	Heuristics + decomposition	Hundreds of SKUs	Scale-aware assortment under substitution; industrial evidence.
Leitner et al. (2024)	Exact assortment w/ product costs	MNL-like MILP	Implicit	—	Capacity; cost	Exact + decomposition	Large MILPs	Exact works to a point; hybrids are needed at scale.
Lotfi et al. (2011)	Multi-objective shelf planning	Aggregate	Implicit	Own	Budget; freshness; capacity	Goal programming	Category-scale	Trade-offs profit vs. service.
Murray et al. (2010)	2D layout + pricing	Demand with layout	Partial	Own + Position	2D blocks; orientation	MINLP	Dozens–hundreds	2D planogram realism matters for revenue.
Yu et al. (2024)	Data-driven robust assortment	MNL (DRO / CVaR)	Yes	—	Moment/CVaR constraints	SOCP; data-driven	Real datasets	Robustness improves lower-tail revenue with a small mean loss.
Abbaszadeh et al. (2024)	Integrated assortment–shelf optimization	Space-elastic demand	Yes	2D layout	Shelf, product, and category bounds	Two-phased metaheuristic (column generation + single-shelf optimization)	Retail-scale (real data)	Hybrid metaheuristic achieved <3.8% optimality gap; high-visibility items drove ~80% of profit.

Notes: "own" = effect of a product's own shelf space on its own sales

Materials and Methods

Problem definition and Solution method

This paper addresses the problem of maximizing the expected profit from arranging product variety on retail shelves. As a combinatorial and NP-hard problem, its inherent complexity —compounded by numerous physical and variety-related constraints—renders exact solution approaches infeasible for real-world scenarios (Heger & Klein, 2024). Hybrid and multi-stage metaheuristic frameworks have recently gained significant attention (Bahrami, 2025). To tackle this computational challenge, we propose a memetic algorithm hybridized with Iterated Local Search (ILS). The choice of a hybrid metaheuristic approach, which combines the global search capabilities of memetic algorithms with the local intensification of ILS, is motivated by its proven effectiveness in solving complex NP-hard problems with intricate constraint structures.

The integrated assortment–shelf optimization problem involves multiple stakeholders with interconnected interests. Category managers seek to maximize profit while balancing variety against shelf constraints; our framework provides optimal SKU selection and facing allocation under realistic bounds. Store operations managers prioritize planogram feasibility, addressed through strict enforcement of capacity constraints. Retail executives focus on overall store profitability; the proposed model quantifies the profit impact of constraint relaxation and supports scenario planning. Consumers benefit indirectly, as the algorithm preserves variety and prevents over-rationalization of staple categories. From a managerial perspective, the framework serves as a decision support tool that reduces reliance on intuition and enables data-driven trade-off analysis.

For each item i, the following data is available: base demand share s_i, remaining profit margin in percent m_i, consumer price p_i, shelf width w_i, space elasticity ϵ_i, assortment level (e.g., categories such as 'Chocolate' or 'Canned Fruits'). Substitution effects between items are modeled using a sparse matrix $P_{IAU} = \{p_{ij}\}$, where each element p_{ij} denotes the probability that the demand for an off-the-shelf item j is transferred to an on-shelf item i. This probability p_{ij} is computed based on the following formula:

$$p_{ij} = \frac{\text{Feature Similarity}(i, j) \times \text{Total Sales Volume}(i)}{\text{Number of Stores with Sales}(i)} \tag{1}$$

Here, Feature Similarity (i, j) quantifies shared attributes —such as brand, color, dimensions, and quality level —between items i and j. Total Sales Volume (i) represents the aggregate sales of the item i, reflecting its market popularity. The Number of Stores with Sales (i) serves as a normalization factor, indicating the market breadth of the item i. This approach integrates both intrinsic product characteristics and observed market presence to estimate the likelihood of substitution.

Shelf-space elasticity is incorporated to capture the responsiveness of sales to the amount of display allocated. Following Curhan (1972), it is defined as the ratio of the percentage change in unit sales to the percentage change in shelf space:

$$E = \frac{\Delta Q / Q}{\Delta S / S} \tag{2}$$

A large meta-analysis reports an average elasticity of approximately 0.17, with substantial heterogeneity across categories—lowest for commodities, higher for staples, and highest for impulse items (Eisend, 2014). Field studies further confirm that location and space effects are non-linear, exhibiting diminishing returns once sufficient facings are allocated (Dreze et al., 1994). In line with this evidence, we calibrated category-level elasticities within the empirically observed band (≈0.13–0.29): assigning lower values to basic commodities such as flour, sugar, and salt; moderate values to staples like rice and tomato paste; and higher values to impulse-oriented categories such as chocolate, potato chips, and soft drinks.

A solution $Y = \{Y_{\mathrm{L}}\}$ assigns to each assortment level L a list of items placed on that level. Each chromosome gene corresponds to one level, and its content is the list of items selected for that level. Facings are not predetermined: the number of faces for an item is simply the number of times it appears in that list. In this way, facings emerge endogenously during optimization rather than being externally imposed.

The profit function has two components. For each item i selected on the shelf, we define:

$$\mathrm{base}_i = m_i \, p_i \, f_i^{\epsilon_i} \tag{3}$$

The direct profit is:

$$R_1 = \sum_{i \in Y} \mathrm{base}_i \, s_i \tag{4}$$

Moreover, the substitution profit is:

$$R_2 = \sum_{i \in Y} \sum_{j \notin Y} \mathrm{base}_i \, s_j \, p_{ij} \tag{5}$$

The objective is therefore:

$$\max_{Y} \; R(Y) = R_1 + R_2 \tag{6}$$

To ensure planogram feasibility, the optimization is subject to the following capacity constraints:

1. Level constraints:

$$\text{MinWidthAssort}_L \leq \sum_{i \in Y_L} w_i \leq \text{MaxWidthAssort}_L \tag{7}$$

2. Store constraint:

$$\sum_L \sum_{i \in Y_L} w_i \leq \text{MAX\_STORE\_WIDTH} \tag{8}$$

Memetic Framework with ILS

The proposed memetic framework combines global search with local exploitation through hill climbing embedded in an ILS scheme. The main steps are here (Figure 1):

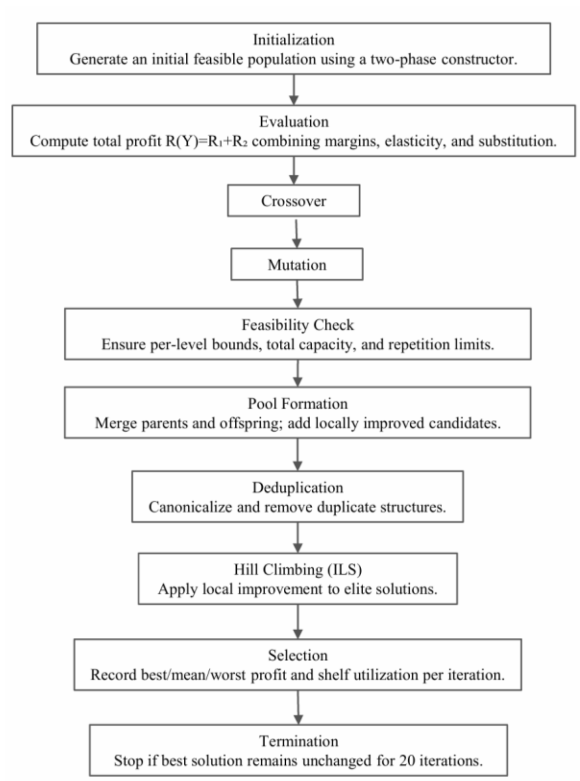

Figure 1. Flowchart: Memetic Algorithm

Having specified the objective function and capacity constraints, we address this NP-hard problem using a memetic algorithm (Neri et al., 2011) that integrates genetic exploration—via crossover and mutation—with targeted local refinement through Iterated Local Search (ILS). This hybrid approach leverages the global search capabilities of evolutionary methods while intensifying promising regions of the solution space. To maintain diversity and avoid premature convergence, the framework incorporates deduplication and diversification mechanisms during pool formation. An adaptive stopping criterion is employed, terminating the search when the best-of-generation profit remains unchanged over a predefined number of iterations.

Initialization

A two-phase generator ensured that all solutions were feasible: (1) for each level, random items were added until the minimum width was reached, without exceeding maxima, and (2) the remaining store capacity was filled with random items across levels until the global width limit was nearly saturated.

We generate a diverse set of feasible chromosomes before the memetic cycle begins. Each chromosome $Y = \{Y_L\}$ maps assortment levels $L$to item lists. The constructor is repair-light and enforces feasibility during construction, ensuring that individuals are valid at birth.

Inputs and Notation

- For item i: width w_i, level $L(i)$, demand share s_i, price p_i, margin m_i, elasticity ϵ_i.
- For each level L: lower/upper width bounds $(\text{MinWidth}_L, \text{MaxWidth}_L)$.
- Global store capacity: MAX_STORE_WIDTH.

Define:

$W(Y_L) = \sum_{i \in Y_L} w_i$, where $W(Y_L)$is the total shelf width occupied by items assigned to level L. Similarly, $W_{\text{tot}}(Y) = \sum_L W(Y_L)$,where $W_{\text{tot}}(Y)$is the total shelf width used across all levels in solution Y.

Phase 1 — Satisfy all per-level minima

For each level L, we randomly sample items $i \in \mathcal{I}_L$(the pool of items assigned to level L) and greedily accept any item that maintains feasibility and contributes toward reaching the MinWidth_L. Formally, if the current partial solution is $Y$and candidate $i \in \mathcal{I}_L$, we accept i if:

$$W(Y_L) + w_i \leq \text{MaxWidth}_L, W_{\text{tot}}(Y) + w_i \leq \text{MAX\_STORE\_WIDTH} \quad (9)$$

Phase 2 — Pack remaining capacity (near-cap fill)

After all levels meet their minimum width requirements, the constructor opportunistically adds items across levels to approach the MAX_STORE_WIDTH without breaching any per-level maxima. At each step, it selects a random level $L$and a random candidate $i \in I_L$, accepting $i$only if all feasibility checks remain satisfied; otherwise, it resamples. The process terminates once the residual capacity MAX_STORE_WIDTH $- W_{\text{tot}}(Y)$ falls below a small tolerance τ.

Mutation

We use a two-by-two, width-preserving mutation designed to maintain shelf balance while exploring new item combinations.

1. Random removals. Uniformly sample two distinct assortment levels L_a, L_b. From each, randomly select one item—denoted $i_a \in L_a$and $i_b \in L_b$—and mark them for removal. Let the target width be $W^\star = w_{i_a} + w_{i_b}$.

2. Random insertions under tolerance. Uniformly sample two candidate items $j_1, j_2$from the entire feasible item pool (with no level restriction). Accept this pair if their combined width falls within the tolerance band.ɣ denote the mutation tolerance.

$$| \ (w_{j_1} + w_{j_2}) - W^\star \ | \leq ɣ \times \max(W^\star, \varepsilon) \tag{10}$$

With ε a small positive constant (numerical guard). If the candidate pair violates the tolerance condition, a new pair is resampled—up to 10 attempts.

3. Constraint screening (hard feasibility). If a width-feasible pair is found, tentatively apply the move $\{i_a, i_b\} \to \{j_1, j_2\}$and check all constraints:

- Per-level width bounds: for every level L, $\text{MinWidth}_L \leq \sum_{i \in Y_L} w_i \leq \text{MaxWidth}_L$.
- Per-level item-count bounds (if active).
- Global store width: $\sum_L \sum_{i \in Y_L} w_i \leq$ MAX_STORE_WIDTH.

4. Rollback on failure. If no candidate pair satisfies both the tolerance and the constraint screening within 10 attempts, the mutation is cancelled and the removed items $\{i_a, i_b\}$ are restored. The individual remains unchanged.

Crossover

Crossover is performed at the assortment-level granularity (a one-cut operation across levels), which may transiently disturb level widths. To keep the algorithm repair-light yet feasible, we apply a simple, in-group post-adjustment immediately after crossover:

1. If a level falls below its minimum width, randomly add items from that level until the MinWidth_L is satisfied, provided that global width and item count/facing caps are not violated.

2. If a level exceeds its maximum width, randomly remove items from that level until the MaxWidth_L is met.

3. After these per-group adjustments, we re-check the global store width and the item repetition cap. If the child solution remains infeasible, the crossover outcome is rejected and the parent is retained.

Deduplication

To prevent clones from dominating and to sustain exploration in Hill-Climbing, we canonicalize each candidate and retain only the first occurrence of each unique structure. If uniqueness falls below a predefined diversity threshold, the population is refilled up to a minimum size. This serves as a pragmatic mechanism for maintaining population diversity (Morrison & De Jong, 2001). Placing deduplication before hill climbing focuses the local-search budget on unique, high-quality representatives rather than expending effort on near-identical clones. This approach reduces redundant evaluations, improves coverage of distinct basins of attraction, and yields more stable improvements per iteration.

Hill-Climbing (ILS local search)

We apply hill-climbing (Russell & Norvig, 2010) to improve a given feasible solution Y. Let the neighborhood of Y be defined by two operators:

1. 1→1 replacement:

Select $i \in Y$, remove it, and insert $j \notin Y$. The new solution is:

$$Y' = (Y \setminus \{i\}) \cup \{j\} \quad (11)$$

2. 1→2 replacement:

Select $i \in Y$, remove it, and insert two items $j_1, j_2 \notin Y$. The new solution is:

$$Y' = (Y \setminus \{i\}) \cup \{j_1, j_2\} \tag{12}$$

Each neighbor Y' is only considered if it is feasible, i.e.,

$$\forall L: \text{MinWidth}_L \leq \sum_{k \in Y_L'} w_k \leq \text{MaxWidth}_L, \sum_L \sum_{k \in Y_L'} w_k \leq \text{MAX\_STORE\_WIDTH} \tag{13}$$

Let $R(Y)$ denote the profit function (from earlier). Then:

$$Y_{t+1} = \begin{cases} Y', & \text{if } R(Y') > R(Y_t),\ Y' \in \mathcal{N}(Y_t) \text{ feasible}, \\ Y_t, & \text{otherwise}. \end{cases} \tag{14}$$

This process is repeated until one of the following termination conditions is met:

- No improving neighbor is found after HC_LOCAL_TRIES = 5 attempts per removed item, or
- HC_GLOBAL_PATIENCE = 10 consecutive non-improving steps occur.

Diversity control

To mitigate genetic drift and premature convergence in this multimodal search, we maintain population diversity throughout the run. This preserves exploration across distinct regions of the solution space, yielding more stable and higher-quality solutions. In our implementation, diversity is sustained through elitism, fitness-proportionate roulette selection, and duplicate elimination with diversity floor maintenance. The latter two mechanisms are explained in detail below.

Fitness-proportionate (roulette-wheel) selection: Parents are sampled with probability proportional to their fitness,

$$p_i = \frac{f_i}{\sum_j f_j} \left(\text{or } p_i = \frac{f_i - f_{\min} + \varepsilon}{\sum_j (f_j - f_{\min} + \varepsilon)} \text{ for numerical stability}\right) \tag{15}$$

which biases selection toward higher-fitness individuals while retaining a nonzero probability for all candidates (Shakir Hameed et al., 2023).

Termination

We employ an adaptive stabilization rule: the run terminates as soon as the best-of-generation profit has remained unchanged for a predefined number of consecutive iterations. A hard iteration cap is maintained only as a safety fallback.

Results

Computational results and Parameter settings

For the empirical study, we selected one branch of the Ofoq Koorosh chain stores (OfoqKourosh, 2025), located in Tehran, with a floor area of approximately 250 m². Within this store, the dry-food warehouse was the focus of analysis. A total shelf length of 21,000 cm was considered as the global capacity distributed across 39 product categories. The product categories include Beverages, Snacks & Sweets, Staples and Dry Goods, Canned and Preserved Foods, Condiments & Cooking Ingredients, Breakfast & Spreads, Pasta & Grains.

The sales and inventory data for these categories cover October 2022 to September 2023. For space elasticity, we used the empirically observed band (≈0.13–0.29) described in the methodology: lower values for commodities (e.g., flour, sugar, salt), moderate values for staples (e.g., rice, tomato paste), and higher values for impulse categories (e.g., chocolate, potato chips, soft drinks). Substitution inputs were derived from product attributes and sales-based relationships, as outlined in the methodology. Category-level minimum and maximum width bounds were set using managerial experience and store-specific constraints.

Algorithmic settings followed the implementation described earlier: mutation tolerance of $\gamma = \pm 10\%$ (maximum 10 attempts); HC_LOCAL_TRIES = 5 and HC_GLOBAL_PATIENCE = 10 for hill climbing; immigrant rate 0.05; and a deduplication floor of 50%. Hill climbing proceeds by selecting a random non-empty level and attempting to remove one item, with the choice of removable items capped at 50 tries. It terminates when no improvement is found in 10 consecutive outer attempts, each allowing up to 5 local replacement trials. Memetic termination is adaptive: the run stops if the best-of-generation profit remains effectively unchanged for 20 consecutive iterations; otherwise, it continues until a hard cap of 5,000 iterations is reached. In practice, we first experimented with fixed 100-generation runs and gradually increased the budget, observing that performance stabilized within this range. We also tested several population sizes (10, 20, 30, and 40) and found that 20 or 30 individuals offer a good balance between diversity and runtime. The combination of duplicate elimination and immigrant injection maintained population diversity, avoided stagnation in local optima, and supported stable convergence under the specified stopping rules. Table 2 summarizes the parameter configuration adopted for the Memetic/ILS algorithm in this study.

Table 2. Memetic/ILS parameter configuration

Parameter	Setting
Generations	<=5,000
Population	20 or 30
Global store width	21,000
Mutation tolerance	±10%, max 10 attempts
Dedup keep ratio	50%
Immigrant rate	5%
Selection	Top-5 + Elite + 30% Roulette + Fill

The empirical results are reported for four algorithmic configurations using real retail data covering 39 product categories, with a global shelf capacity of 21,000 cm. The experiments examined the impact of population size and category-level maximum bounds on solution quality, allocation patterns, and convergence behavior.

The computational results are summarized across four experimental runs in Table 3. Run 1 employed a population of 20 with original minimum–maximum bounds, while Run 2 increased the maximum by ×1.5 for categories initially capped below 1000 cm. Run 3 used a population of 30 with original bounds, and Run 4 used a population of 30 with increased bounds. As shown in Table 3, moving from Run 1 to Run 3 (20→30 individuals, original bounds) yields a modest objective gain (+17,130; ≈+0.6%) at the cost of longer compute time (~14→16 hours). Relaxing maxima at population 20 (Run 2) delivers a larger improvement over Run 1 (+186,367; ≈+6.2%) with a smaller number of generations (3,718→2,817) and comparable runtime. The best objective value is achieved in Run 4 (3,249,058) under population 30 with relaxed maxima, albeit with the most extended runtime (~29 hours) and the highest number of generations (4,089). These results highlight the trade-off between exploration (population size) and constraint flexibility (relaxed maxima), on the one hand, and solution quality and computational cost, on the other. Accordingly, subsequent analyses focus on the population-30 configurations (Runs 3–4) as they provide the most stable and superior outcomes.

Table 3. Summary of computational performance across four experimental runs

Number of runs	Seconds	Hours	Number of generations	Profit
Run 1	51,652	14	3718	3,020,329
Run 2	53,178	15	2817	3,206,696
Run 3	58,857	16	2855	3,037,459
Run 4	103,852	29	4089	3,249,058

Allocation outcomes of Run 3 and Run 4 are presented in Figures 2a and 2b, respectively. Under the original bounds (Figure 2a), impulse-oriented categories, such as Chocolate and Soft Drinks, reached their maxima, while staples (e.g., Flour, Salt) remained near their minima. After relaxing the maximum (Figure 2b), additional space was absorbed by Jam & Preserves, Sauces & Dressings, and Non-Alcoholic Beverages, resulting in both an increase in the number of distinct SKUs and a rise in total facings.

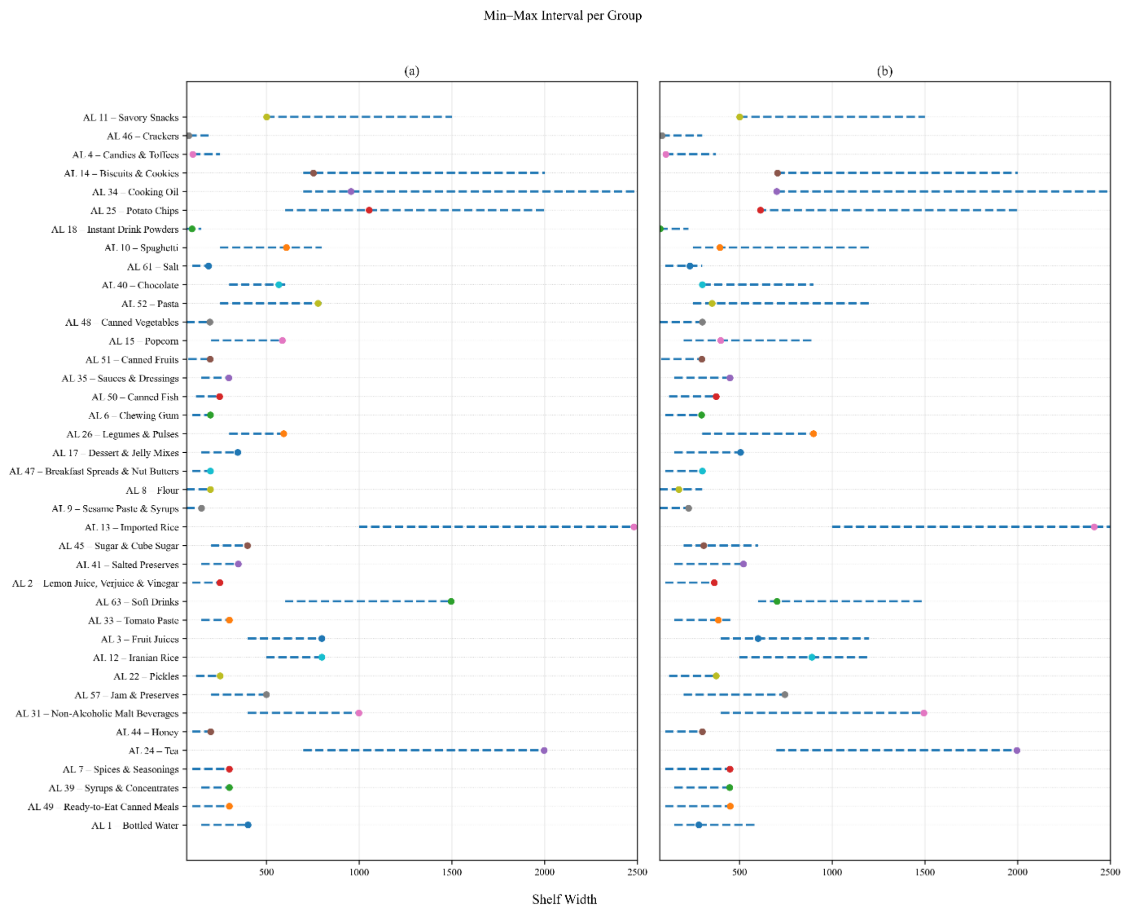

Figure 2. Category-level shelf allocation under the 30-individual configuration. (a) Original min–max bounds. (b) Relaxed maxima scenario.

The convergence behavior of the 30-individual population is reported in Figures 3a and 3 b. Under original bounds (Figure 3a, Run 3), the best profit curve stabilizes around 3.04M after approximately 2,855 generations. Under relaxed maxima (Figure 3b, Run 4), the algorithm converges to 3.25M after 4,089 generations. Comparing the two scenarios, relaxing category maxima improved the final best profit and yielded smoother and slightly superior convergence.

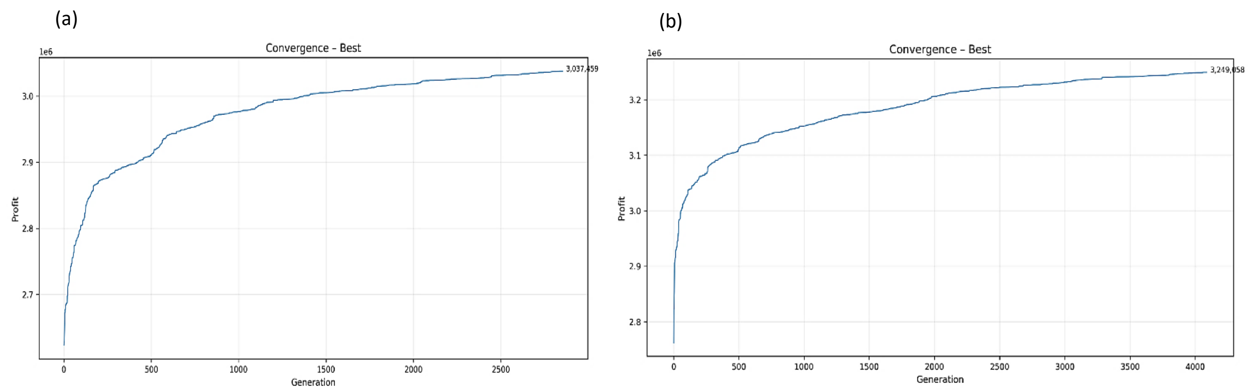

Figure 3. Convergence for population = 30. (a) Original min–max bounds. (b) Relaxed maxima scenario.

Cumulative coverage analysis under original bounds (Figure 4a) highlights an intense concentration of shelf allocation. When categories are sorted by their final allocated width, the curve shows that roughly 20% of the 39 categories (approximately eight groups) capture around 70% of the total shelf space. Elastic and impulse-driven groups, such as those in the Chocolate, Soft Drinks, and Tea sectors, dominate the allocation. At the same time, staple and commodity categories with lower profits (e.g., Flour, Salt) remain clustered at the lower end. This pattern confirms the existence of a pronounced head–tail structure, where a minority of categories absorb most of the capacity, underscoring their strategic importance in retail space planning. Cumulative coverage under relaxed maxima (Figure 4b) displays a slightly flatter curve compared to the base case, indicating reduced concentration. With higher upper bounds, mid-tier categories, such as Jam & Preserves, Sauces & Dressings, and Non-Alcoholic Beverages, absorb additional space, thereby distributing coverage more evenly across these groups. Although the top approximately 20% of categories still account for a dominant share, the relative contribution of secondary groups increases— suggesting that flexibility in maximum bounds can enhance variety and balance without undermining profit concentration in high-elasticity groups.

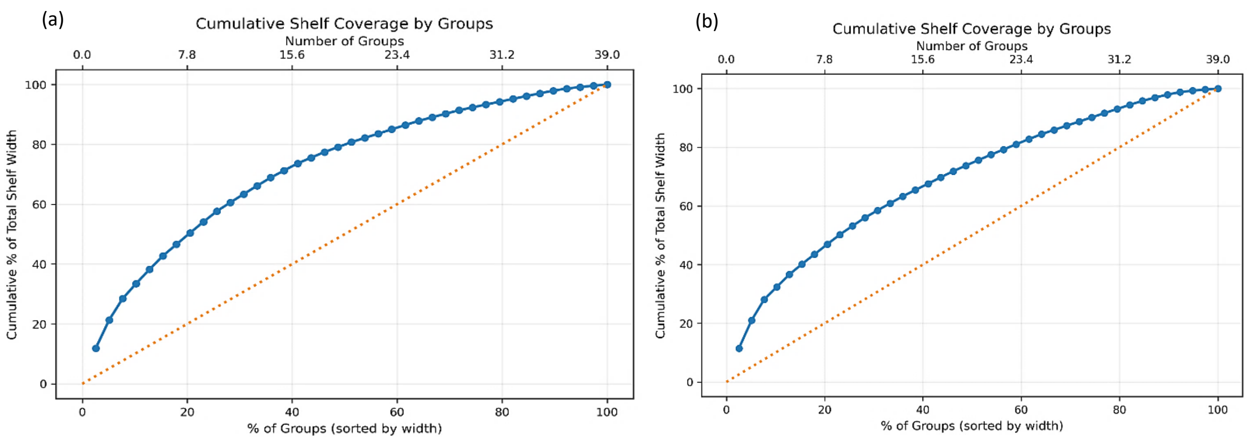

Figure 4. Cumulative coverage for population = 30. (a) Original min–max bounds. (b) Relaxed maxima scenario.

Table 4 summarizes the allocation outcomes for the 30-individual configuration under the original min–max bounds. Results show that highly elastic and impulse-oriented categories approach their maximum width allocations and exhibit high average facings, reflecting strong responsiveness to shelf space. Conversely, basic commodities—such as Flour, Salt, and Sugar—remain close to their minima with low SKU variety and facings, confirming their limited contribution to incremental profit. Mid-range categories (e.g., Pasta, Biscuits & Cookies, Cooking Oil) occupy moderate allocations consistent with their elasticity levels. Overall, this table demonstrates that the optimal planogram under the original bounds is highly skewed toward elastic categories, while commodity categories are underrepresented. Table 5 reports the results for the 30-individual configuration when maximum category bounds are relaxed. Compared to the normal-bounds scenario, several mid-tier categories—such as Jam & Preserves, Sauces & Dressings, Canned Vegetables, and Non-Alcoholic Beverages—occupy significantly more space, resulting in both an increase in the number of unique SKUs and the average facings. This reallocation enhances variety and shelf visibility across a broader range of groups, while maintaining high-elastic categories (e.g., Tea, Soft Drinks) near their maximum levels. Commodities remain constrained at low levels, indicating limited responsiveness to market space. Thus, the relaxed-bounds scenario provides a more balanced allocation, enhancing category variety without diminishing the dominance of the most profitable groups.

Table 4. Final allocation outcomes under 30-individual configuration (original min–max bounds)

Product Category	Initial Number of SKUs	Minimum Shelf Width (cm)	Maximum Shelf Width (cm)	Final Allocated Width (cm)	Unique SKUs Selected	Total SKUs Selected including) repeated (facings	Average Facings per SKU
Biscuits & Cookies	152	700	2000	753.9	39	135	3.46
Bottled Water	16	150	400	400	16	57	3.56
Breakfast Spreads & Nut Butters	17	100	200	198.4	11	21	1.91
Candies & Toffees	8	100	250	103	4	7	1.75
Canned Fish	19	120	250	246.8	10	29	2.9
Canned Fruits	8	80	200	196.5	8	23	2.88
Canned Vegetables	10	70	200	195	10	26	2.6
Chewing Gum	40	100	200	197.7	19	38	2
Chocolate	42	300	600	566	21	163	7.76
Cooking Oil	29	700	2500	955.5	24	91	3.79
Crackers	11	80	200	81.4	4	7	1.75
Dessert & Jelly Mixes	41	150	350	345.7	25	41	1.64
Flour	16	70	200	198.3	12	16	1.33
Fruit Juices	131	400	800	798	84	155	1.85
Honey	14	100	200	199.8	12	28	2.33
Imported Rice	59	1000	2500	2482	34	62	1.82
Instant Drink Powders	24	70	150	98.1	5	6	1.2
Iranian Rice	28	500	800	798.7	21	36	1.71
Jam & Preserves	33	200	500	498.9	26	71	2.73
Legumes & Pulses	27	300	600	593.5	26	51	1.96
Lemon Juice, Verjuice & Vinegar	19	100	250	248.5	18	26	1.44
Non-Alcoholic Malt Beverages	56	400	1000	998.3	48	132	2.75
Pasta	33	250	800	778	29	64	2.21
Pickles	17	120	250	249.5	17	31	1.82
Popcorn	13	200	600	585.5	11	29	2.64
Potato Chips	36	600	2000	1054	26	62	2.38
Ready-to-Eat Canned Meals	21	100	300	300	17	40	2.35
Salt	11	100	200	187.3	7	14	2
Salted Preserves	17	150	350	347.7	13	41	3.15
Sauces & Dressings	23	150	300	296.2	14	44	3.14
Savory Snacks	43	500	1500	500.5	19	28	1.47
Sesame Paste & Syrups	9	70	150	149	7	19	2.71
Soft Drinks	55	600	1500	1494.9	54	243	4.5
Spaghetti	24	250	800	607.7	17	31	1.82
Spices & Seasonings	44	100	300	299.9	30	47	1.57
Sugar & Cube Sugar	42	200	400	397.6	16	29	1.81
Syrups & Concentrates	7	150	300	300	7	28	4
Tea	55	700	2000	1997.6	47	172	3.66
Tomato Paste	7	150	300	299.2	7	30	4.29
Total	1257	10180	26400	20998.6	815	2173	100.64

Table 5. Final allocation outcomes under 30-individual configuration (relaxed maxima)

Product Category	Initial Number of SKUs	Minimum Shelf Width (cm)	Maximum Shelf Width (cm)	Final Allocated Width (cm)	Unique SKUs Selected	Total SKUs Selected including) repeated (facings	Average Facings per SKU
Biscuits & Cookies	152	700	2000	705	45	91	2.02
Bottled Water	16	150	600	281.7	15	41	2.73
Breakfast Spreads & Nut Butters	17	100	300	299.8	12	31	2.58
Candies & Toffees	8	100	375	102.5	5	7	1.4
Canned Fish	19	120	375	374.3	13	44	3.38
Canned Fruits	8	80	300	297	8	35	4.38
Canned Vegetables	10	70	300	300	10	40	4
Chewing Gum	40	100	300	296.3	21	57	2.71
Chocolate	42	300	900	300	14	94	6.71
Cooking Oil	29	700	2500	701	24	63	2.63
Crackers	11	80	300	82.2	5	7	1.4
Dessert & Jelly Mixes	41	150	525	505	30	60	2
Flour	16	70	300	172.8	12	14	1.17
Fruit Juices	131	400	1200	600.8	65	126	1.94
Honey	14	100	300	299.8	11	41	3.73
Imported Rice	59	1000	2500	2412.5	31	60	1.94
Instant Drink Powders	24	70	225	72.1	5	5	1
Iranian Rice	28	500	1200	890	19	42	2.21
Jam & Preserves	33	200	750	745.4	27	106	3.93
Legumes & Pulses	27	300	900	899.5	26	78	3
Lemon Juice, Verjuice & Vinegar	19	100	375	364.3	19	35	1.84
Non-Alcoholic Malt Beverages	56	400	1500	1494	45	191	4.24
Pasta	33	250	1200	353	24	29	1.21
Pickles	17	120	375	374	17	46	2.71
Popcorn	13	200	900	399.5	10	19	1.9
Potato Chips	36	600	2000	614	21	36	1.71
Ready-to-Eat Canned Meals	21	100	450	450	18	60	3.33
Salt	11	100	300	232.7	5	18	3.6
Salted Preserves	17	150	525	522	13	62	4.77
Sauces & Dressings	23	150	450	448.9	15	68	4.53
Savory Snacks	43	500	1500	501.5	24	29	1.21
Sesame Paste & Syrups	9	70	225	225	7	29	4.14
Soft Drinks	55	600	1500	701.9	41	107	2.61
Spaghetti	24	250	1200	394.2	13	22	1.69
Spices & Seasonings	44	100	450	449.4	36	67	1.86
Sugar & Cube Sugar	42	200	600	307.6	16	19	1.19
Syrups & Concentrates	7	150	450	447	7	41	5.86
Tea	55	700	2000	1996.3	43	169	3.93
Tomato Paste	7	150	450	386.5	7	38	5.43
Total	1257	10180	32600	20999.5	779	2127	112.62

To assess the practical value of the optimization framework, we compared the algorithm's best solution (Run 4) against the store's current planogram configuration. The current configuration was extracted from the retailer's operational records, reflecting the existing assortment and facing decisions implemented in practice. Expected profit for both configurations was computed using the same objective function (Eq. 6), ensuring a consistent basis for comparison under identical demand, substitution, and elasticity assumptions.

Table 6 summarizes the comparison across three key metrics: the average number of active SKUs per category, the average facings per category, and the expected profit. The optimized solution increases the average SKU count by approximately 27% (from 15.7 to 20.0 items per category) and total facings by approximately 12%, resulting in a 37% improvement in expected profit under the model assumptions.

Table 6. Comparison of optimized solution versus current store configuration

Metric	Current	Memetic/ILS	Change (%)
Avg. SKUs per category	15.7	20	27.4%
Avg. facings per category	48.8	54.5	11.7%
Expected profit	2,373,380	3,249,058	36.9%≅37%

These results suggest that the current store configuration underutilizes available shelf capacity in terms of product variety. The optimization algorithm capitalizes on this opportunity by activating additional SKUs—particularly smaller-width items that fit within the same shelf space—and by leveraging substitution relationships to capture demand that would otherwise be lost when items are out of stock. It should be noted that the profit improvement is conditional on the validity of the estimated elasticity and substitution parameters; the comparison, therefore, indicates potential rather than guaranteed gains.

Discussion and Conclusion

In this study, we formulated and solved an integrated assortment–shelf optimization problem that simultaneously considers substitution effects and space-elastic demand under realistic retail constraints. We developed a hybrid Memetic Algorithm embedded with Iterated Local Search (ILS) to address this NP-hard problem. Practically, we constructed a substitution structure based on product attributes and sales relationships to capture demand recapture when an item is off-shelf; calibrated category-level space elasticities within empirically grounded bands to reflect the diminishing returns of facings; encoded feasibility through per-category minimum and maximum bounds and a tight global capacity constraint, ensuring that every candidate planogram is operationally valid; and designed a memetic search in which chromosomes map items to shelf levels and facings emerge endogenously, incorporating feasibility-first initialization, repair-light crossover and mutation, deduplication and diversity control, and local hill-climbing to refine promising solutions. The proposed framework was validated using real retail data.

The hybridization strategy was intentionally designed to balance exploration and intensification throughout the search process. The Memetic layer governed population-based global exploration through crossover and mutation operators, while the embedded ILS component applied local hill-climbing to refine elite individuals at each generation. Mutation operations were width-preserving, preventing unbalanced allocations, and crossover was followed by an in-group adjustment to maintain per-level bounds. A deduplication mechanism avoided population cloning and preserved diversity, while an adaptive stopping criterion terminated the run once objective improvement plateaued. This methodological integration provided both computational stability and practical realism, making the algorithm suitable for large-scale retail data where exact optimization is infeasible.

The present study proposed a hybrid framework incorporating several innovations that collectively distinguished it from prior shelf-space optimization research. First, unlike conventional approaches, where facings were explicit decision variables requiring separate optimization (Hübner & Schaal, 2017; Hübner et al., 2020)The chromosome encoding allowed facings to emerge endogenously through repeated item occurrences within shelf-level genes, thereby reducing the solution space dimensionality. Second, the two-phase initialization guaranteed 100% feasibility across all runs, unlike previous GA-based methods that relied on expensive post-hoc repair (Czerniachowska et al., 2021) .Third, whereas standard mutation operators often destabilize constraint satisfaction (Hansen et al., 2010), the width-preserving 2×2 mutation maintained approximate capacity neutrality within ±10% tolerance bands, eliminating repair overhead. Fourth, to prevent premature convergence, the framework incorporated deduplication before hill climbing, combined with immigrant injection and diversity floor maintenance, thereby focusing computational effort on structurally distinct solutions. Fifth, although hybrid metaheuristics had been identified as promising directions (Heger & Klein, 2024), few studies have operationalized this integration for shelf-space problems; the Memetic/ILS framework explicitly combines evolutionary global search with systematic hill-climbing intensification. Together, these design choices enabled the algorithm to achieve high-quality, feasible solutions at problem scales where exact solvers were computationally infeasible.

The computational experiments confirmed that the hybrid Memetic/ILS algorithm achieved stable, high-quality solutions across runs. The combination of global exploration and local hill-climbing refinement maintained diversity in the solution pool, avoided premature convergence, and ensured continuous improvement while preserving feasibility under strict shelf and category constraints. This hybrid structure enabled the algorithm to reach near-optimal regions efficiently, providing practical convergence within a realistic computational time.

The observed allocation patterns across product categories are consistent with theoretical expectations from the literature. Impulse-driven and high-margin categories reached or approached their upper space limits, reflecting their high profit density and strong responsiveness to additional shelf exposure. Conversely, low-elasticity staples stabilized near their lower bounds. Relaxing category-level constraints resulted in structural changes to assortment balance. Greater flexibility in maximum bounds allowed a broader and more balanced distribution of shelf space without diminishing overall profitability. This outcome suggests that moderate constraint relaxation supports both variety and visual appeal, resulting in more realistic and consumer-friendly shelf configurations.

Pareto analysis confirmed that roughly 20% of SKUs accounted for more than 80% of total profit. This aligns with the long-tail phenomenon in assortment planning research, where a small number of items dominate category performance. From a managerial standpoint, these results support SKU rationalization strategies: retailers can confidently focus on high-performing, high-elasticity categories while maintaining limited representation of lower-performing staples to preserve variety and shopper satisfaction. Thus, an optimal shelf strategy should retain minimal representation of low-elasticity commodities for completeness while dedicating incremental space to categories that exhibit strong responsiveness and higher marginal returns.

Beyond the direct optimization outcomes, several implicit findings emerged from the computational experiments with important managerial implications. (a) The 6.2% profit improvement from relaxing category maxima (Run 2 vs. Run 1) demonstrated that retailers could use constraint adjustments as a low-cost strategic lever—rather than expanding physical store capacity, simply revising internal category policies unlocked substantial gains. (b) When upper bounds were relaxed, mid-tier categories such as Jam & Preserves and Sauces & Dressings absorbed significant additional space, indicating that these "middle performers" possessed untapped elasticity; retailers typically focused on optimizing top-tier categories, yet the findings suggested that mid-tier groups warranted greater attention in assortment reviews. (c) The algorithm maintained minimal representation of low-elasticity staples (Flour, Salt) without eliminating them, preserving perceived variety and category completeness—addressing a common concern that optimization led to over-rationalization. (d) The modest gain (+0.6%) from increasing population size (20→30) at the cost of additional runtime suggested that retailers with tighter planning windows could adopt smaller populations with minimal quality loss, enabling faster replanning cycles.

The comparison with the current store configuration revealed actionable insights for retail managers. The 37% profit improvement was primarily attributable to two factors: (a) an increase in the number of active SKUs (+27%), suggesting that the current planogram underutilized available shelf capacity in terms of product variety; and (b) improved exploitation of substitution relationships, whereby adding complementary items captured demand that would otherwise have been lost. Notably, the optimized solution favored smaller-width items that occupied less shelf space per unit, allowing for a higher variety within the same total capacity. From an operational perspective, this implied that expanding the store's back-room storage or replenishment frequency could have supported the recommended assortment expansion. However, the profit improvement was conditional on the model's demand assumptions. Managers were advised to interpret these figures as indicative potential rather than guaranteed outcomes and to consider pilot testing on a subset of categories before implementing them on a full scale.

While this study makes meaningful contributions to integrated assortment–shelf optimization, several limitations should be acknowledged. From a methodological standpoint, the NP-hard nature and scale of the problem (~1,000+ item-level decisions) precluded comparison against exact optimal benchmarks. Results may not generalize directly to stores with different layouts, customer demographics, or product types such as perishables. Computationally, the best-performing configuration required approximately 29 hours, which is acceptable for weekly planning cycles but prohibitive for real-time optimization. Finally, the Iranian retail context may limit direct transferability to Western or other emerging markets. Despite these limitations, the proposed framework provides a robust foundation for practical shelf-space optimization and offers clear directions for future research.

Building on these results, the proposed framework offers strong practical transferability and extensibility. It serves as a scenario-planning engine for category managers: by tuning per-category ceilings/floors and total capacity, they can quantify assortment breadth, facings, and profit trade-offs before operational rollout. The demand layer is modular—alternative elasticity and substitution estimators (including data-driven or machine-learning updates) can be plugged in without redesigning the optimization core. Looking ahead, the same pipeline can be deployed across multi-store networks and omni-channel contexts, enriched with promotion/seasonality calendars or lightweight learning components for continuous parameter refresh—amplifying performance while preserving interpretability and operational credibility.

Future studies could extend the analysis in several directions. One promising avenue is to test the algorithm across multiple stores with heterogeneous layouts and consumer demographics, enabling richer generalization. Another direction is to incorporate dynamic aspects—such as seasonality, promotions, or competitive reactions—into the optimization framework. Advances in demand modeling, particularly with machine learning methods, could further improve elasticity and substitution estimation, yielding more precise inputs. Finally, real-world implementation studies would be valuable for assessing managerial acceptance and the operational feasibility of algorithmic recommendations.

Data Availability Statement

Data available on request from the authors.

Acknowledgements

The authors gratefully acknowledge the management and data analytics team of Ofoq Kourosh Retail Chain for providing access to the empirical data used in this study.

Conflict of interest

The authors declare no conflict of interest related to the content, data, or results of this study. All analyses were conducted independently, and no external parties influenced the interpretation or presentation of the findings.

Funding

The authors received no financial support for this article's research, authorship, and/or publication.

Ethical considerations

The authors declare compliance with all applicable ethical guidelines, including proper data handling, originality of content, and avoidance of duplicate submission.

References

Abbasnia, F., Zandieh, M., Bahrami, F., & Pourhejazy, P. (2025). A Decision Analysis Framework for the Identification and Performance Preservation of Strategic Products in the Supply Chain. *Logistics*, *9*(3), 89. https://doi.org/10.3390/logistics9030089

Abbaszadeh, A., Iranpoor, M., Mahnam, M., & Bahrami, F. (2025). Optimisation of assortment breadth and allocation of the selected product groups to the two-dimensional shelves. *International Journal of Systems Science: Operations & Logistics*, *12*(1), 2478426. https://doi.org/10.1080/23302674.2025.2478426

Agarwal, K., Jain, P., & Rajnayak, M. A. (2019). Comparative Analysis of Store Clustering Techniques in the Retail Industry. DATA, https://doi.org/10.5220/0007917500650073

Anderson, E. E., & Amato, H. N. (1974). A mathematical model for simultaneously determining the optimal brand-collection and display-area allocation. *Operations Research*, *22*(1), 13–21. https://doi.org/10.1287/opre.22.1.13

Aouad, A., Farias, V., Levi, R., & Segev, D. (2018). The approximability of assortment optimization under ranking preferences. *Operations Research*, *66*(6), 1661–1669. https://doi.org/10.1287/opre.2018.1754

Bahrami, F. (2025). Many-to-many location-routing problem with multiple paths, heterogeneous vehicles and time windows. *Industrial Management Journal*. https://doi.org/10.22059/imj.2025.402989.1008266

Baron, O., Berman, O., & Perry, D. (2011). Shelf space management when demand depends on the inventory level. *Production and operations management*, *20*(5), 714–726. https://doi.org/10.1111/j.1937-5956.2010.01201.x

Bianchi-Aguiar, T., Hübner, A., Carravilla, M. A., & Oliveira, J. F. (2021). Retail shelf space planning problems: A comprehensive review and classification framework. *European Journal of Operational Research*, *289*(1), 1–16. https://doi.org/10.1016/j.ejor.2020.06.018

Brandimarte, P., Craparotta, G., & Marocco, E. (2024). Inventory reallocation in a fashion retail network: A matheuristic approach. *European Journal of Operational Research*, *317*(2), 603–615. https://doi.org/10.1016/j.ejor.2024.04.016

Chen, X., Ma, W., Simchi-Levi, D., & Xin, L. (2024). Assortment planning for recommendations at checkout under inventory constraints. *Mathematics of Operations Research*, *49*(1), 297–325. https://doi.org/10.1287/moor.2023.1357

Çömez-Dolgan, N., Fescioglu-Unver, N., Cephe, E., & Şen, A. (2021). Capacitated strategic assortment planning under explicit demand substitution. *European Journal of Operational Research*, *294*(3), 1120-1138. https://doi.org/10.1016/j.ejor.2021.02.026

Curhan, R. C. (1972). The relationship between shelf space and unit sales in supermarkets. *Journal of Marketing Research*, *9*(4), 406–412. https://doi.org/10.2307/3149304

Czerniachowska, K., Michalak, K., & Hernes, M. (2023). Heuristics for the shelf space allocation problem. *Opsearch*, *60*(2), 835-869. https://doi.org/10.1007/s12597-023-00636-1

Czerniachowska, K., Sachpazidu-Wójcicka, K., Sulikowski, P., Hernes, M., & Rot, A. (2021). Genetic algorithm for the retailers' shelf space allocation profit maximization problem. *Applied Sciences*, *11*(14), 6401. https://doi.org/10.3390/app11146401

Désir, A., Goyal, V., Jiang, B., Xie, T., & Zhang, J. (2024). Robust assortment optimization under the Markov chain choice model. *Operations Research*, *72*(4), 1595–1614. https://doi.org/10.1287/opre.2022.2420

Dreze, X., Hoch, S. J., & Purk, M. E. (1994). Shelf management and space elasticity. *Journal of Retailing*, *70*(4), 301–326. https://doi.org/10.1016/0022-4359(94)90002-7

Eisend, M. (2014). Shelf space elasticity: A meta-analysis. *Journal of Retailing*, *90*(2), 168–181. https://doi.org/10.1016/j.jretai.2013.03.003

Farias, V. F., Jagabathula, S., & Shah, D. (2013). A nonparametric approach to modeling choice with limited data. *Management science*, *59*(2), 305–322. https://doi.org/10.1287/mnsc.1120.1610

Geismar, H. N., Dawande, M., Murthi, B., & Sriskandarajah, C. (2015). Maximizing revenue through two-dimensional shelf-space allocation. *Production and operations management*, *24*(7), 1148–1163. https://doi.org/10.1111/poms.12316

Haji, R., Moarefdoost, M., & Ebrahimi, B. (2009). Finding the cost of inventory in make to order supply chain under vendor managed inventory program. *Industrial Management Journal*, *1*(3). https://dor.isc.ac/dor/20.1001.1.20085885.1388.1.3.2.1

Hansen, J. M., Raut, S., & Swami, S. (2010). Retail shelf allocation: A comparative analysis of heuristic and meta-heuristic approaches. *Journal of Retailing*, *86*(1), 94–105. https://doi.org/10.1016/j.jretai.2010.01.004

Hansen, P., & Heinsbroek, H. (1979). Product selection and space allocation in supermarkets. *European Journal of Operational Research*, *3*(6), 474-484. https://doi.org/10.1016/0377-2217(79)90030-4

Heger, J., & Klein, R. (2024). Assortment optimization: A systematic literature review. *OR Spectrum*, *46*(4), 1099–1161. https://doi.org/10.1007/s00291-024-00752-4

Hense, J., & Hübner, A. (2022). Assortment optimization in omni-channel retailing. *European Journal of Operational Research*, *301*(1), 124–140. https://doi.org/10.1016/j.ejor.2021.09.045

Hübner, A. (2011). *Retail category management: decision support systems for assortment, shelf space, inventory and price planning* (Vol. 656). Springer Science & Business Media. https://doi.org/10.1007/978-3-642-22477-5_5

Hübner, A., & Kuhn, H. (2024). Decision support for managing assortments, shelf space, and replenishment in retail. *Flexible Services and Manufacturing Journal*, *36*(1), 1–35. https://doi.org/10.1007/s10696-023-09492-z

Hübner, A., Kuhn, H., & Kühn, S. (2016). An efficient algorithm for capacitated assortment planning with stochastic demand and substitution. *European Journal of Operational Research*, *250*(2), 505–520. https://doi.org/10.1016/j.ejor.2015.11.007

Hübner, A., & Schaal, K. (2017). An integrated assortment and shelf-space optimization model with demand substitution and space-elasticity effects. *European Journal of Operational Research*, *261*(1), 302–316. https://doi.org/10.1016/j.ejor.2017.01.039

Hübner, A., Schäfer, F., & Schaal, K. N. (2020). Maximizing profit via assortment and shelf-space optimization for two-dimensional shelves. *Production and operations management*, *29*(3), 547–570. https://doi.org/10.1111/poms.13111

Hübner, A. H., & Kuhn, H. (2012). Retail category management: State-of-the-art review of quantitative research and software applications in assortment and shelf space management. *Omega*, *40*(2), 199–209. https://doi.org/10.1016/j.omega.2011.05.008

Hwang, H., Choi, B., & Lee, M.-J. (2005). A model for shelf space allocation and inventory control considering location and inventory level effects on demand. *International Journal of Production Economics*, *97*(2), 185–195. https://doi.org/10.1016/j.ijpe.2004.07.003

Kök, A. G., & Fisher, M. L. (2007). Demand estimation and assortment optimization under substitution: Methodology and application. *Operations Research*, *55*(6), 1001–1021. https://doi.org/10.1287/opre.1070.0409

Kuhn, D., Shafiee, S., & Wiesemann, W. (2025). Distributionally robust optimization. *Acta Numerica*, *34*, 579–804. https://doi.org/10.1017/s0962492924000084

Leitner, M., Lodi, A., Roberti, R., & Sole, C. (2024). An exact method for (constrained) assortment optimization problems with product costs. *INFORMS Journal on Computing*, *36*(2), 479–494. https://doi.org/10.1287/ijoc.2022.0262

Lotfi, M., Rabbani, M., & Ghaderi, S. (2011). A weighted goal programming approach for replenishment planning and space allocation in a supermarket. *Journal of the Operational Research Society*, *62*(6), 1128–1137. https://doi.org/10.1057/jors.2010.52

Mohaghar, F., Jolai, F., & Heydari, J. (2020). Modeling a Four Echelon Omni-Channel Supply Chain for Seasonal Product under Stochastic Demand. *Industrial Management Journal*, *12*(2), 206–235. https://doi.org/10.22059/imj.2020.306753.1007759

Morrison, R. W., & De Jong, K. A. (2001). Measurement of population diversity. International conference on artificial evolution (evolution artificielle), https://doi.org/10.1007/3-540-46033-0_3

Murray, C. C., Talukdar, D., & Gosavi, A. (2010). Joint optimization of product price, display orientation and shelf-space allocation in retail category management. *Journal of Retailing*, *86*(2), 125–136. https://doi.org/10.1016/j.jretai.2010.02.008

Neri, F., Cotta, C., & Moscato, P. (2011). *Handbook of memetic algorithms* (Vol. 379). Springer. https://doi.org/10.1007/978-3-642-23247-3

OfoqKourosh. (2025). *Ofoq Kourosh Annual General Meeting*. https://okcs.com/posts/view/639

Pham, T., Mathmann, F., Septianto, F., & Chylinski, M. (2025). The impact of assortment size and population density on product evaluation. *Journal of Retailing*. https://doi.org/10.1016/j.jretai.2025.02.004

Russell, S. J., & Norvig, P. (2010). Artificial Intelligence-A Modern Approach, Third International Edition. http://aima.cs.berkeley.edu/

Sethuraman, R., Gázquez-Abad, J. C., & Martínez-López, F. J. (2022). The effect of retail assortment size on perceptions, choice, and sales: Review and research directions. *Journal of Retailing*, *98*(1), 24–45. https://doi.org/10.1016/j.jretai.2022.01.001

Shakir Hameed, A., Alrikabi, H. M. B., Abdul–Razaq, A. A., Ahmed, Z. H., Nasser, H. K., & Mutar, M. L. (2023). Appling the roulette wheel selection approach to address the issues of premature convergence and stagnation in the discrete differential evolution algorithm. *Applied Computational Intelligence and Soft Computing*, *2023*(1), 8892689. https://doi.org/10.1155/2023/8892689

Yu, Y., Wang, B., & Zheng, S. (2024). Data-driven product design and assortment optimization. *Transportation Research Part E: Logistics and Transportation Review*, *182*, 103413. https://doi.org/10.1016/j.tre.2024.103413

Ziari, M., & Sajadieh, M. S. (2025). Shelf space allocation in retailing: a literature review. *RAIRO-Operations Research*, *59*(5), 2721-2748. https://doi.org/10.1051/ro/2025037