ISSN 2616-5775 Vol. 9, Issue 3: 91-97, DOI: 10.25236/AJCIS.2026.090311

Intelligent Warehouse Layout Optimization Based on Dynamic Demand Perception and Co-Picking Behavior Clustering

Miao Wang[1,*,#], Jianchun Hao[2,#], Hongbo Li[1,#], Haoran Liu[1,#]

[1]*School of Railway Intelligent Engineering, Dalian Jiaotong University, Dalian, China*
[2]*Zhan Tianyou College (CRRC College), Dalian Jiaotong University, Dalian, China*
[*]*Corresponding author: 15241675215@163.com*
[#]*These authors contributed equally to this work*

Abstract: *With the rapid growth of e-commerce, distribution centers are facing serious challenges caused by a sharp increase in SKU variety and highly fragmented orders. Picking operations have become the main bottleneck restricting warehouse efficiency. This research intends to solve the problem where traditional experience-based layout lags behind dynamic business needs by using a data-driven way. Recency weighted dynamic EIQ evaluation model. Through the introduction of a time decay function, the model can capture recent changes in the popularity of SKUs and achieve more timely ABC classification. Then, to examine the relationship among items, we use SVD to lower the size of the high-dimensional co-picking matrix. Hierarchical clustering is then used to find SKU association clusters. Finally, a cluster-aware storage allocation algorithm is created to allocate high-value clusters to the best physical places. Simulation experiments based on actual orders from a logistics center in eastern China indicate that the proposed method cuts the average walking distance during picking by 19.34 percent. This greatly speeds up order fulfillment and reduces labor costs.*

Keywords: *Warehouse operations, Storage location assignment, R-EIQ model, Singular value decomposition, Association mining, Cluster analysis*

1. Introduction

In today's smart logistics and supply chain systems[1], [2], warehouses play an important role as key nodes that link production with consumption. Their operation efficiency has a direct effect on how fast the whole supply chain responds. As e-commerce keeps growing, orders get smaller, happen more often, and need deliveries quickly. This makes traditional experience-based management methods insufficient for handling the rapid growth of SKU variety. Previous studies show that picking operations often account for about half of the total operating cost of a warehouse. Among these costs, ineffective walking time by pickers is the main source of inefficiency. Therefore, optimizing storage location[3] assignment and spatial layout to shorten picking routes at the source has become a key concern for both academia and industry.

Existing warehouse optimization[4], [5] studies mainly focus on static layouts and heuristic path planning[6]. Classic storage assignment strategies, such as turnover-based and correlation-based rules, can improve efficiency to some extent. However, they show clear limitations when facing dynamic e-commerce demand. Many studies apply intelligent optimization algorithms, such as genetic algorithms[7], [8] and ant colony algorithms[9], to restructure storage layouts. Although these methods can find near-optimal solutions through large-scale search, they often ignore the time-varying nature of SKU demand. For example, certain products may experience a sharp increase in outbound frequency in a short period due to promotions or seasonal effects. A static evaluation model can’t reflect this kind of short-term popularity change over time. It results in a spatial imbalance of hotspots and local congestion on picking routes.

Product association analysis has been discussed before, but it hasn't been much used for making good space plans. Traditional methods, like the Apriori algorithm[10], struggle with large SKUs and have trouble finding hidden connections in sparse data. Many studies only describe things like "Product A and Product B are often bought together," but they don’t turn these ideas into actual numbers for shelf placement. This gap between what we expect and what happens makes it hard for complex grouping

ISSN 2616-5775 Vol. 9, Issue 3: 91-97, DOI: 10.25236/AJCIS.2026.090311

methods to help with warehouse setup. To solve this, this research presents a simple, data-driven system. The main idea is to add a recency weighting method to improve traditional EIQ analysis, helping the model respond faster to recent changes in SKU performance. Secondly, singular value decomposition is employed to tackle the sparsity and high dimensionality of the association matrix. Then, clustering is done in a lower-dimensional latent space to show more hidden picking together patterns. Lastly, a cluster-aware storage mapping mechanism is created to transform logical clusters into actual groups inside the warehouse. The main contribution of this study is to extend the dynamic storage assignment model theoretically and prove practically that algorithm-driven decision-making can improve the efficiency of the warehouse without increasing hardware investment.

2. Methodology

This study creates a data-driven warehouse optimization system to solve the conflict between SKU placement and picking efficiency via mathematical modeling. Framework begins with a thorough investigation of order features. By means of dynamic weightings, spatial feature mappings, and cluster aggregations, it makes scientific storage allocation choices.

2.1 Dynamic EIQ Evaluation Model with Recency Weighting

In order to measure the true value of a SKU under constantly changing market conditions, this research goes past the conventional static stats and creates an R-EIQ model that has a time-decay feature. The main assumption is that the influence of old orders on the forecast for new ones will decrease exponentially as time passes.

First, an order weight function is defined. Let the order timestamp be t, and the current analysis time be T. The original weight $\omega(t)$ follows an exponential decay rule:

$$\omega(t) = e^{-\lambda(T-t)} \tag{1}$$

where λ is the decay factor reflecting the intensity of business fluctuation. Based on this weight, the weighted outbound frequency (Entry) of SKU_i is redefined. Let O_i be the set of orders involving SKU s during the period. The weighted Entry index $R_E(i)$ is defined as:

$$R_E(i) = \sum_{j \in O_i} \omega(t_j) \tag{2}$$

Similarly, the outbound quantity (Quantity) $R_Q(i)$ and the city flow indicator (Item) $R_I(i)$ are also calculated using weighted values. To eliminate scale differences, range normalization is applied. Taking R_E as an example:

$$E_i' = \frac{R_E(i) - \min(R_E)}{\max(R_E) - \min(R_E)} \tag{3}$$

The normalized indicators $E_i', Q_i', I_i' \in [0,1]$ are then combined into a comprehensive evaluation function using a weighted linear form:

$$Score_i = \alpha \cdot E_i' + \beta \cdot I_i' + \gamma \cdot Q_i' \tag{4}$$

where $\alpha + \beta + \gamma = 1$. To further reflect spatial cost in picking operations, a spatial correction factor s_i is introduced. It is determined by the dispersion of the physical storage locations of the SKU. The final evaluation score is defined as:

$$f(i) = Score_i \cdot (1 + \eta \cdot s_i) \tag{5}$$

where μ is the spatial adjustment coefficient. This score serves as the basis for subsequent ABC classification and storage priority determination. The classification results are shown in Figure 1.

ISSN 2616-5775 Vol. 9, Issue 3: 91-97, DOI: 10.25236/AJCIS.2026.090311

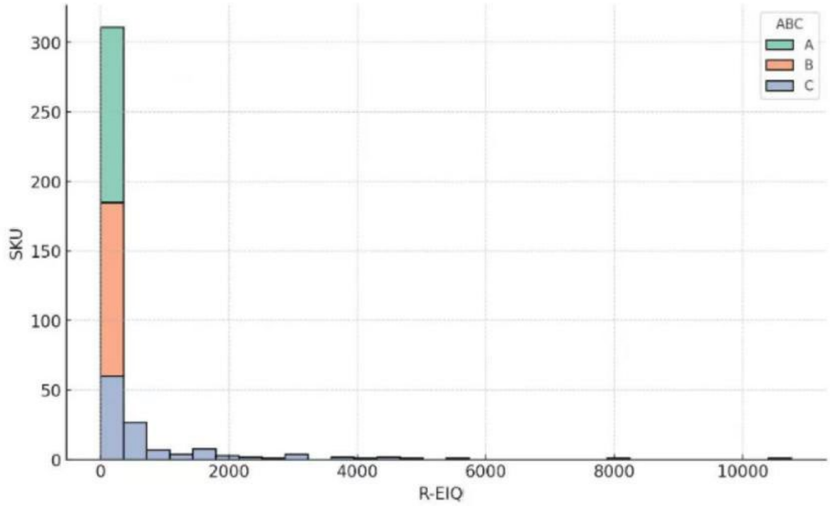

Fig. 1 Distribution of R-EIQ and ABC classification results

2.2 SVD-Based Representation and Clustering of SKU Co-Picking Associations

Individual SKU value alone is not sufficient to optimize picking paths. Product association is key to achieving cluster-based picking. This study constructs a co-picking matrix and applies dimensionality reduction to uncover latent relationships among SKUs.

First, a co-picking matrix $\mathbf{C} \in \mathbb{R}^{m \times m}$ is constructed. The element C_{ij} represents the weighted frequency with which SKU_i and SKU_j appear together in the same order. As shown in Figure 2, the matrix is highly sparse. Direct association calculation would introduce noise. Therefore, singular value decomposition (SVD) is used for low-rank approximation.

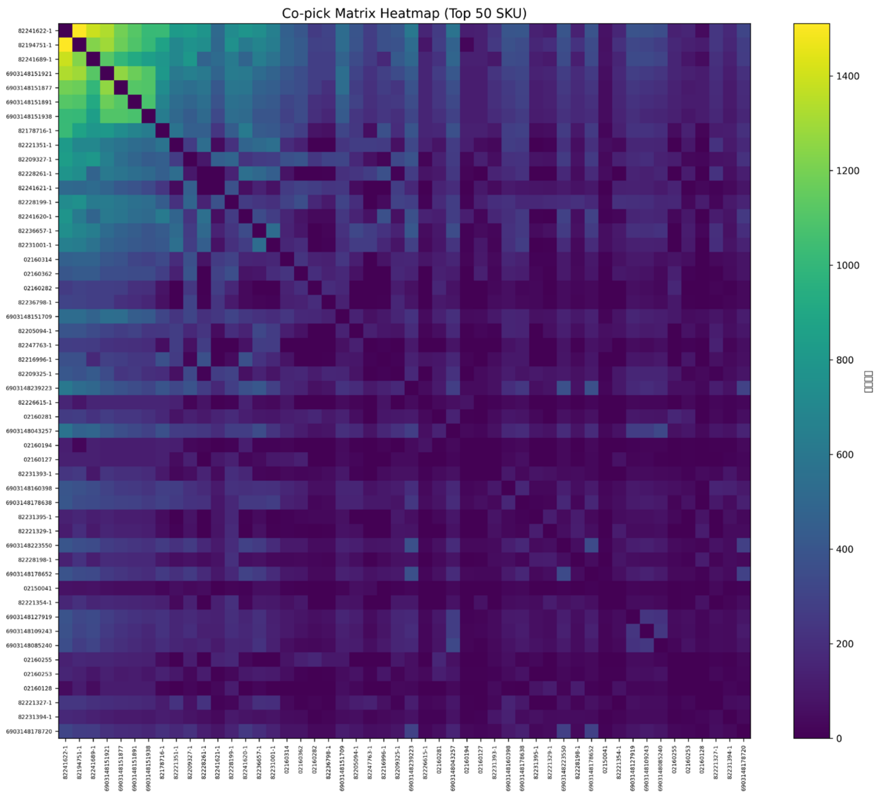

Fig. 2 Heat map of the co-picking matrix

The matrix $\mathbf{C}$ is decomposed as:

$$\mathbf{C} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \tag{6}$$

where the singular values in $\mathbf{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_m)$ are arranged in descending order. The top k singular values and their corresponding vectors are selected to form a reduced feature space. The feature vector of SKU_i is expressed as:

$$\vec{v}_i = \mathbf{U}_k(i,\cdot)\sqrt{\mathbf{\Sigma}_k} \tag{7}$$

In the k-dimensional latent space, cosine similarity $Sim(i,j)$ between two SKUs is calculated as the clustering distance:

$$Sim(i,j) = \frac{\vec{v}_i \cdot \vec{v}_j}{\|\vec{v}_i\| \|\vec{v}_j\|} \tag{8}$$

Agglomerative hierarchical clustering is then applied. Each SKU starts as an independent cluster. The most similar clusters are merged iteratively until the stopping condition is met. The goal is to maximize intra-cluster similarity and minimize inter-cluster similarity. The singular value distribution is shown in Figure 3, and the clustering result is shown in Figure 4.

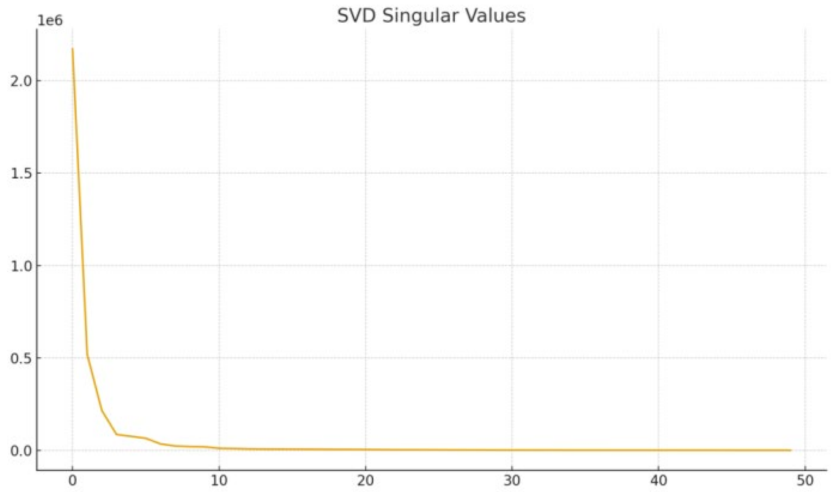

Fig. 3 Singular value distribution from SVD

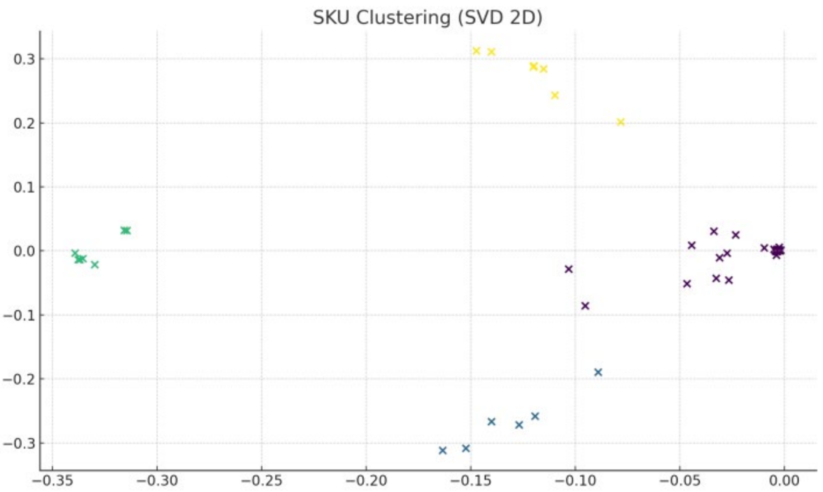

Fig. 4 Scatter plot of SKU clustering results

2.3 Logic of the Cluster-Aware Storage Assignment Algorithm

The essence of storage optimization is to map high-value clusters to physical locations with the lowest picking cost. The proposed cluster-aware algorithm performs layout optimization through a two-stage mapping process.

Stage 1: Slot cost modeling.

For any storage slot $L(x,y)$, the walking cost to the picking and checking start point $P_0(x_0, y_0)$ is

ISSN 2616-5775 Vol. 9, Issue 3: 91-97, DOI: 10.25236/AJCIS.2026.090311

defined. Under a rectangular fishbone or standard shelf layout, Manhattan distance is used:

$$D(L_j) = |x_j - x_0| + |y_j - y_0| \tag{9}$$

Slots are sorted in ascending order of $D(L_j)$ and divided into a golden zone P_1, a regular zone P_2, and a peripheral zone P_3. The division is shown in Figure 5.

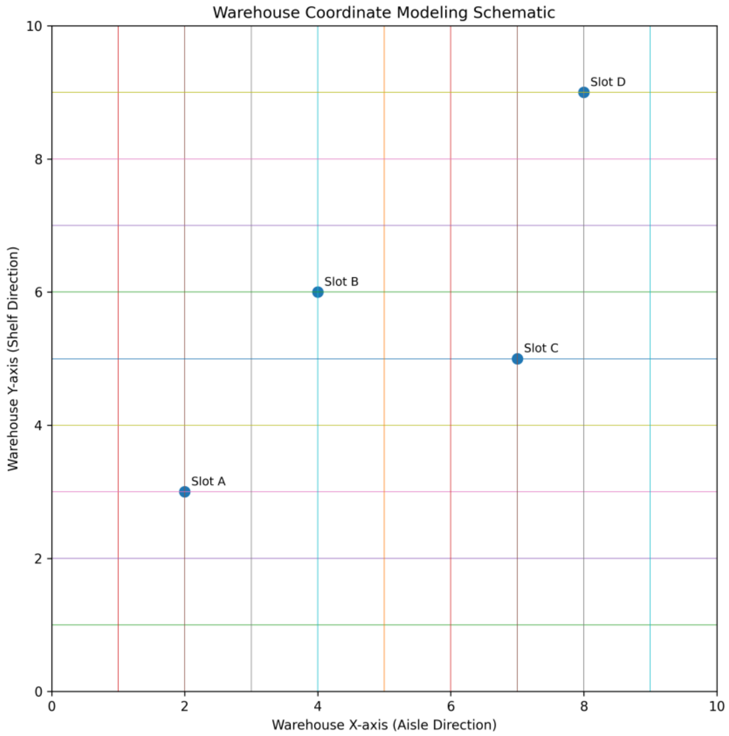

Fig. 5 Slot mapping and warehouse coordinate modeling

Stage 2: Cluster-level assignment.

Let $G = \{G_1, G_2, \dots, G_n\}$ be the set of SKU clusters. The comprehensive value index of cluster $V(G_k)$ is defined as:

$$V(G_k) = \sum_{i \in G_k} f(i) \cdot \ln(1 + |G_k|) \tag{10}$$

Clusters $|G_k|$ are sorted in descending order of $V(G_k)$ and assigned as a whole to corresponding slot zones. Within each cluster, a centroid convergence rule is applied to ensure that highly associated SKUs are placed in the same aisle or adjacent facing slots. This allocation logic compresses picking routes through strong association constraints and ensures path optimality at the physical level.

3 Experiments and Results

3.1 Data Description

A high-fidelity order replay simulation system is adopted to recreate the actual working environment of an East China logistics center. So that the proposed optimization method can be practically feasible. Data from the competition dataset is given, which includes specific warehouse details such as shelf coordinates, aisle widths, and picking starting points. These data constitute a digital two-dimensional coordinate system.

Time selection adopts a strict time-based division method. Orders from 90 days before the model date are taken as the training set for calculating R-EIQ indicators and constructing the co-picking clustering model. To check if it can generalize well, we use 10,000 actual orders from the next 30 days as a separate test set. It prevents memory effect and allows objective assessment of unseen orders. In preprocessing, Winsorization is applied to the quantity fields and long-tail SKUs are removed to minimize noise and provide a fair baseline comparison.

ISSN 2616-5775 Vol. 9, Issue 3: 91-97, DOI: 10.25236/AJCIS.2026.090311

3.2 Result Analysis

Test orders are simulated with paired layouts before and after optimization. Detailed comparison results are obtained. An optimized layout has an obvious efficiency advantage. The average picking distance reduces from 302. 62 in the baseline layout to 244. 09, which means a decrease by 19. 34%. This change can be seen in the distance distribution. As seen in Fig. 6, the optimized paths move towards the lower distance range. It means less walking cost and good aggregation of valuable SKUs and closely related groups.

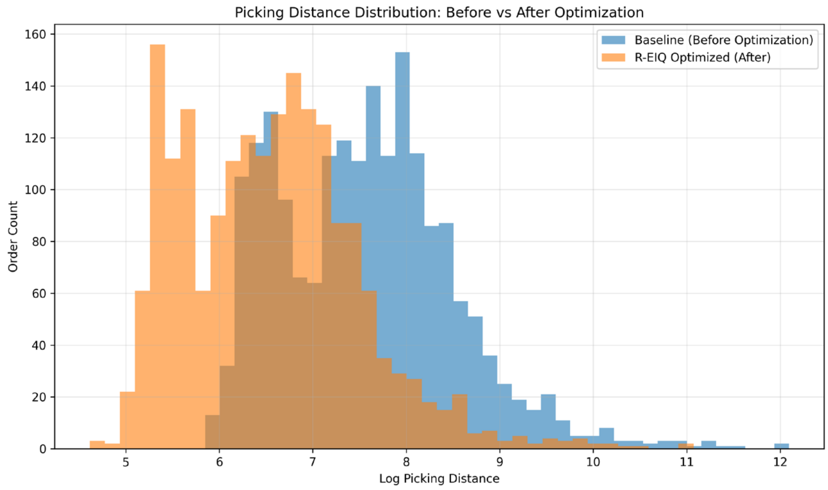

Fig. 6 Comparison of picking distance distributions

Figure 7 shows the comparison of average picking distance.

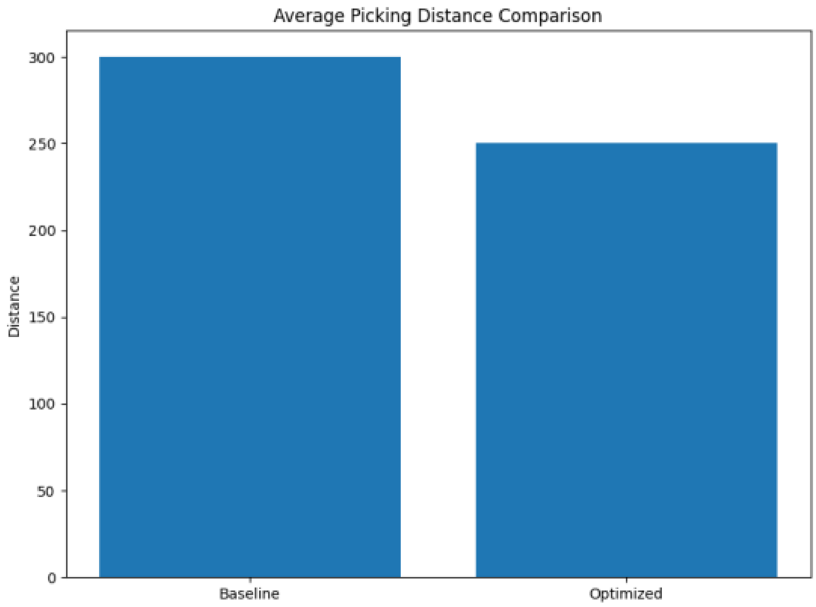

Fig. 7 Comparison of average picking distance

Further operation analysis shows that it has a big effect on the ability to fulfill orders. Picking time at the 95th percentile drops by 40%-50%, meaning it performs well even when many orders come in. Path variance is reduced by 35%-55%, making it easier to predict operations and increase capacity. Labor-wise, daily work hours go down by 30%-50%, improving performance. To ensure the results are reliable, a paired t-test was done on the data before and after optimization. The t-value is 13.61, and the p-value is close to 0, which shows a clear improvement over the old layout. This method boosts system efficiency using algorithms without changing the hardware.

ISSN 2616-5775 Vol. 9, Issue 3: 91-97, DOI: 10.25236/AJCIS.2026.090311

4. Conclusion

This study looks at the operation data from a logistics center in East China and shows that dynamic demand recognition and grouping products can improve warehouse operations. The results show that using the R-EIQ model with SVD can identify the main product areas and related product groups. The proposed Cluster-Aware Storage Assignment Algorithm has made complex order patterns easier to understand and apply to the layout. Compared to the original layout, the new plan reduces picking distance, speeds up order fulfillment, and lowers labor hours. These results are supported by data.

The method doesn't require expensive automation tools. Instead, it improves management efficiency with better algorithms, making it easy to apply in other situations. However, there are still some limitations. The current model mainly works with offline batch processing. Future work will explore real-time systems that can adjust layouts automatically when there are promotions or seasonal changes. The study will also look at traffic flow inside the warehouse to improve picking paths and create a more accurate digital twin warehouse improvement plan.

References

[1] D. Gao, N. Wang, and B. Jiang, "Analysis of Information-Sharing Mechanisms in Online Closed-Loop Supply Chain Systems," Systems, vol. 13, no. Compendex, 2025, doi: 10.3390/systems13090810.

[2] J. Nuerk and F. Daena, "Systems Engineering Methodology for Digital Supply Chain Business Models," Systems Engineering, vol. 28, no. Compendex, pp. 411–437, 2025, doi: 10.1002/sys.21802.

[3] X. Chen, F. Li, B. Jia, J. Wu, Z. Gao, and R. Liu, "Optimizing storage location assignment in an automotive Ro-Ro terminal," Transportation Research Part B: Methodological, vol. 143, no. Compendex, pp. 249–281, 2021, doi: 10.1016/j.trb.2020.10.009.

[4] J. Zhao, K. Liang, F. Wang, H. Liu, J. Yang, and L. Zhou, "Warehouse layout optimization for fishbone robotic mobile fulfillment systems," Expert Systems with Applications, vol. 259, no. Compendex, 2025, doi: 10.1016/j.eswa.2024.125166.

[5] M. Hosseini, S. Chalil Madathil, and M. T. Khasawneh, "Reinforcement learning-based simulation optimization for an integrated manufacturing-warehouse system: a two-stage approach," Expert Systems with Applications, vol. 290, no. Compendex, 2025, doi: 10.1016/j.eswa.2025.128259.

[6] Y. Zhu and Y. Zhang, "A game-theoretic cooperative path planning strategy using hybrid heuristic optimization algorithm," IET Control Theory and Applications, vol. 19, no. Compendex, 2025, doi: 10.1049/cth2.12766.

[7] M. S. Rahman, A. Duary, A. A. Shaikh, and A. K. Bhunia, "An application of real coded Self-organizing Migrating Genetic Algorithm on a two-warehouse inventory problem with Type-2 interval valued inventory costs via mean bounds optimization technique," Applied Soft Computing, vol. 124, no. Compendex, 2022, doi: 10.1016/j.asoc.2022.109085.

[8] E. Esmaeeli, A. Haji, J. Rezaeenour, M. Sabaghieh Yazd, and M. R. Feylizadeh, "Optimizing truck scheduling and dock placement at cross-docking systems through a hybrid genetic-ant colony optimization algorithm," Journal of Industrial and Production Engineering, vol. 42, no. Compendex, pp. 938–966, 2025, doi: 10.1080/21681015.2025.2498662.

[9] Z. Luan, L. Yu, Q. Tian, and M. Liu, "Multi-Objective Optimization of Multi-Warehouse Cargo Allocation and Transportation Planning Using an Enhanced Ant Colony Algorithm," Informatica (Slovenia), vol. 49, no. Compendex, 2025, doi: 10.31449/inf.v49i29.8439.

[10] V. Agarwal, R. K. Gupta, and A. Tiwari, "Multi-Phase Recommender Framework for Pattern Warehousing Using Evolutionary Optimization," IEEE Access, vol. 13, no. Compendex, pp. 181925–181943, 2025, doi: 10.1109/ACCESS.2025.3622481.