



Research article

A collaborative optimization model of efficiency frontier and resource in smart warehouses with stochastic disturbance

Guanghui Fan¹, Weihua Liu^{2,*}, Chao Fan³, Guodong Li⁴ and Hongxiao Zhang⁴

¹ School of Traffic and Transportation, Beijing Jiaotong University, No.3, Shangyuan Village, Haidian District, Beijing, 100044, China

² College of Management and Economics, Tianjin University, No.92, Weijin Road, Nankai District, Tianjin, 300072, China

³ Smart Management Business Division, Kunlun Digital Technology Co., Ltd., Block A19, CNPC Science and Technology Park, Changping District, Beijing 102206, China

⁴ JD Logistics, 8th Floor, Building 2, No.20 Kechuang 11th Street, Yizhuang Economic and Technological Development Zone, Beijing 101111, China

* **Correspondence:** Email: lwhliu@tju.edu.cn.

Abstract: As e-commerce fulfillment complexity increases, intelligent warehouse systems centered on automated guided vehicles (AGVs) face increasingly stringent demands for resource allocation efficiency. The coordinated configuration of AGVs and workstations directly determines system productivity and operational costs. However, existing research often relies on deterministic assumptions or single optimization methods, which struggle to capture the impact of random factors—such as order arrivals and processing times—on system performance, and lack comprehensive analysis of resource coordination and optimal investment. To address this gap, this study proposes a joint optimization approach integrating stochastic frontier analysis (SFA) with mixed-integer nonlinear programming (MINLP), establishing an intelligent warehouse resource allocation model focused on the integrated optimization of AGVs and workstations. First, SFA is employed to separate environmental randomness from technical inefficiency, identifying the efficiency frontier under random disturbances and delineating the effective boundary of resource allocation. Second, an MINLP model is constructed under the constraint of this efficiency frontier to jointly optimize discrete investments and matching relationships between AGVs and workstations. By coupling efficiency evaluation with combinatorial optimization, this method overcomes the limitations of single approaches in handling uncertainty and achieving global optimization. Third, a case study at a large e-

commerce company in China demonstrates that the model effectively identifies resource redundancy and imbalance: after optimization, the number of workstations and AGVs decreased by approximately 48% and 11%, respectively, achieving a reduction of resource inputs and a decline in weighted objective-function value, together with efficiency improvement. By introducing SFA into discrete resource optimization and deeply integrating it with MINLP, this research establishes a unified framework of efficiency frontier identification and resource optimization allocation, offering a novel methodology for designing intelligent warehouse systems under uncertain conditions.

Keywords: smart warehouse; stochastic frontier analysis; resource allocation; picking strategy; mixed-integer nonlinear programming

Mathematics Subject Classification: 90B06, 90B36, 90C11

1. Introduction

Automated storage plays an increasingly important role in modern e-commerce fulfillment systems. With the continued expansion of global e-commerce order volumes and rising consumer expectations for delivery speed, warehouse operations are accelerating their transformation from traditional “person-to-goods” to “goods-to-person”, automated, and intelligent modes. Worldwide, the large-scale application of automated guided vehicle (AGV), autonomous mobile robot (AMR), automated sorting system, and collaborative warehouse management system (WMS) and warehouse control system (WCS) control technologies has enabled automated operation of warehouse transportation, picking, sorting, and replenishment. In the intelligent-warehouse AGV-based “goods-to-person” mode, AGVs perform task scheduling, path planning, and shelf transport according to system commands, delivering target shelves to picking stations, where human pickers or automated equipment then complete item picking and sorting based on system information. Compared with traditional warehousing, this mode reduces worker travel distance, improves order processing efficiency, and enhances the accuracy and flexibility of warehouse operations. Overall, intelligent warehouse automation has become a key direction for upgrading global e-commerce fulfillment systems. Qu et al. (2025) further emphasized the important role of digitalization, intelligence, and collaborative optimization in enhancing resilience and operational performance in intelligent supply chain management [1].

In intelligent warehouse operations, the number of active AGVs and the number of open picking stations are critical factors affecting outbound picking efficiency and operating costs. Roodbergen and Vis (2009) showed that equipment capacity, system structure, and throughput performance are closely related in automated storage systems [2]. Deploying AGVs incurs electricity, maintenance, and wear costs, while each workstation requires staffing with associated ongoing labor costs. Over-investment in AGVs and workstations leads to idle equipment and wasted resources; under-investment limits system capacity, preventing the completion of picking tasks within the required time window, thereby causing order backlogs and delivery delays. There is a clear transport–processing coupled relationship between AGVs and workstations, and their collaborative optimization involves complex constraints and combinatorial decisions. Therefore, how to reasonably determine the numbers of AGVs and workstations to ensure fulfillment timeliness while reducing redundant investment is a critical issue for improving smart warehouse operational efficiency.

In smart warehouse operations, efficiency is constrained by AGV count and workstation count, while random disturbances—such as order arrival fluctuations, equipment status changes, operator efficiency variations, and system operating rules—affect the judgment of optimal configuration. Lamballais et al. (2017), in their performance evaluation study of mobile robot fulfillment systems, pointed out that in addition to robot quantity and workstation processing capacity, factors such as order structure and system operating rules also affect system throughput [3]. Aigner et al. (2023), investigating stochastic frontier production functions, indicated that in production systems with random disturbances and technical inefficiency, it is necessary to distinguish between random noise and efficiency losses in order to more accurately identify the true production frontier [4]. Therefore, when evaluating and optimizing AGV and workstation configurations, it is essential to strip out and correct for random disturbances to avoid misinterpreting stochastic noise as system efficiency differences, thereby more accurately identifying the model-derived optimal configuration.

Research has made some progress in smart warehouse operations optimization, but gaps remain. First, most literature focuses on system planning, path optimization, or single-resource allocation. For example, Weidinger et al. (2018) studied shelf storage assignment in KIVA-like warehouses, emphasizing that warehouse layout and operating rules have significant impacts on system efficiency, yet their focus remained on the effect of system rules on warehouse efficiency [5]. Nicoletti and Appolloni (2025) proposed the Warehouse 5.0 framework, further emphasizing that future-oriented warehouse systems need to establish coordination among human–robot collaboration, digital capabilities, intelligent decision-making, and sustainable operations, implying that research into smart-warehouse resource allocation needs to shift from single-equipment optimization to collaborative decision-making oriented toward system efficiency and uncertain environments [6]. Guerrazzi et al. (2024) studied picking and outbound loading efficiency in shuttle-based storage and retrieval systems (SBS/RS) and pointed out that due to the high complexity of smart warehouse systems, insufficient real-world validation leads to heavy reliance on simulation environments [7].

Second, although studies have addressed storage assignment, single-equipment scheduling, and order batching, research on the joint optimization of AGV and workstation counts remains insufficient. For example, Zhang et al. (2025) relied on heuristic algorithms such as ant colony and neural networks to optimize AGV path planning, but this approach can only improve AGVs in terms of operational efficiency [8]. In the case of Zhao and Zhang (2025), they focused on order allocation across multiple picking workstations, showing that order organization affects system efficiency, but did not conduct quantitative matching between AGV transport capacity and workstation processing capacity [9].

Other studies have relied on static or hypothetical assumptions and insufficiently accounted for order fluctuations, equipment status, and human factors. Mahmoudinazlou et al. (2025) noted that traditional path optimization methods tested on fixed order sets have limited adaptability to the complex, dynamically changing warehouse environment [10]. Such studies have not adequately characterized the efficiency frontier under complex disturbances.

To address the above problems, based on actual operational data from an e-commerce smart warehouse belonging to Enterprise A, this paper innovatively constructs a joint optimization model integrating stochastic frontier analysis (SFA) and mixed-integer nonlinear programming (MINLP). First, SFA is used to estimate the production function relationship between workstation count, AGV fleet size, and picking output and to measure the technical efficiency of the warehouse system. Then, an MINLP model is built to solve for the optimal numbers of AGVs and workstations that meet output requirements. According to the empirical results, the warehouse yields an estimated mean technical

efficiency of 0.685 and shows considerable deviations from the production frontier. This outcome points to resource mismatches alongside other unobserved sources of operational inefficiency; after joint optimization of AGV and workstation counts, the average number of open workstations decreases by 48%, and the average number of active AGVs decreases by 11%.

The contributions of this paper are summarized as follows: First, this study introduces stochastic frontier analysis (SFA) into the discrete resource allocation problem of smart warehouses. Unlike existing studies relying on deterministic assumptions or simulation-generated inputs, the proposed SFA model adopts real-world operational data to quantitatively decompose the composite error term into random disturbances and technical inefficiency and further identifies the stochastic production frontier under random operating conditions. This provides an empirically driven approach to characterize the input–output relationship of AGV-based goods-to-person warehousing systems under uncertainty. Second, this study constructs a unified two-stage SFA–MINLP joint optimization framework that tightly couples efficiency frontier estimation with mixed-integer resource optimization. Conventional SFA studies focus exclusively on efficiency measurement without producing discrete resource decisions, whereas traditional MINLP models neglect stochastic disturbances in their optimization objectives. The proposed framework converts efficiency evaluation results directly into integer configuration solutions for AGVs and picking workstations, bridging the gap between efficiency measurement and operational resource decisions. Third, validated by 658 sets of real hourly operational data from a large-scale e-commerce intelligent warehouse, the proposed model quantifies the degree of resource redundancy and outputs optimal integer resource schemes under different operational scenarios. The empirical results demonstrate that the average technical efficiency is merely 0.685, and joint optimization can reduce workstations by 48% and AGVs by 11% while maintaining fulfillment requirements. These findings offer quantitative decision support for collaborative AGV–workstation deployment under stochastic disturbance, which may provide decision support for warehouse resource planning.

2. Literature review

In smart warehouse systems, AGVs and picking workstations are two core resources that affect system operational efficiency and costs. Regarding smart-warehouse resource allocation, research has mainly focused on system design and equipment configuration, workstation capacity and task organization, collaborative system operation, and efficiency measurement with random disturbances.

2.1. Research on intelligent warehouse system resource allocation

From the perspectives of warehouse system design and automation equipment configuration, most studies have not directly discussed the quantitative ratio of AGVs to workstations, but instead explained the determinants of warehouse performance with respect to factors such as system structure and equipment capability, mobile robot operating rules, and collaboration between workstations and information systems.

At the level of system structure and equipment capability, studies have mostly focused on the impact of warehouse system structure and equipment capability on system performance. For example, Lee et al. (2018), from the perspective of IoT-based warehouse management systems, emphasized the linkage among information systems, equipment resources, and operational workflows [11]. Barros and

Nascimento (2021), in their review on mobile robot fulfillment systems, also noted that equipment capability, system structure, and operational strategies jointly determine the performance of automated warehouse systems [12]. Such research provides a foundation for understanding the role of AGV resources, but the emphasis has mostly been on general issues of system design or equipment control, with insufficient discussion on how to match AGV transport capacity with workstation processing capacity.

At the level of mobile robot operating rules, Koster et al. (2007), in their review of warehouse order picking design and control, pointed out that storage strategies, picking paths, order batching, zoning, and control rules all affect picking efficiency [13]. Lamballais et al. (2017), from a performance estimation perspective, stated that robot count, workstation configuration, and operating rules jointly affect system throughput [3]. Weidinger et al. (2018), in their study of shelf storage assignment in KIVA-like warehouses, pointed out that storage strategies affect robot transport efficiency [5]. Fragapane et al. (2021), from the perspective of mobile robot planning and control for internal logistics, noted that robot application must be considered together with task organization and site constraints [14]. Overall, research has revealed the influence of multiple factors on warehouse performance, but a detailed characterization of capacity matching and collaborative configuration between AGV quantity and workstation quantity is still lacking.

From the perspectives of workstation resource allocation and information system collaboration, research has focused on picking capacity and system service levels. Regarding system coordination and integrated operation, the coupling among multiple equipment types has been addressed. For example, Roy et al. (2012), in their performance analysis of autonomous vehicle warehouses, showed that there are design trade-offs among vehicle fleet size, system structure, and operational demand [15]. Valle and Beasley (2025) and Zhao and Zhang (2025) studied order allocation strategies across multiple picking workstations, concluding that there is a clear coupling among order assignment, robot transport, and workstation processing capacity [9,16]. Li et al. (2015), from the perspective of warehouse–retailer network design, further illustrated that the configuration relationship between warehouse nodes and downstream demand-points affects overall system cost and service efficiency [17]. Benzdia et al. (2019), Brunetti et al. (2024), and Han et al. (2026), from industry application and intelligent logistics node perspectives, indicated that the configuration of automated equipment needs to be matched with specific service scenarios and logistics node functions [18–20].

Although the above studies have begun to address overall system operational efficiency, most remained at the level of local optimization and did not adequately characterize the synergistic ratio between AGV and workstation counts. Therefore, this paper conducts joint optimization of AGV and workstation configurations for operational resource-allocation optimization.

2.2. Research on intelligent warehouse efficiency measurement and random disturbances

In addition to resource allocation issues, efficiency measurement and random disturbances in smart warehouse systems are also important research directions. Studies have addressed efficiency measurement methods, warehouse and robotic fulfillment system performance evaluation, stochastic disturbances, and complex system operations. However, insights into efficiency frontier identification under random disturbances have rarely been leveraged for subsequent resource allocation decisions. In particular, research on joint optimization of AGV and workstation counts is lacking.

In terms of efficiency measurement methods, Aigner et al. (2023) proposed a stochastic frontier production function model that decomposes the error term into random disturbance and technical

inefficiency, providing an effective tool for identifying production frontiers under uncertainty [4]. Unlike traditional deterministic efficiency analysis, SFA can distinguish random noise from technical inefficiency, making it more suitable for smart warehouse scenarios with equipment status fluctuations, order demand changes, and operator efficiency variations.

In terms of performance evaluation of smart warehouses and robotic fulfillment systems, the reviews of automated storage systems by Roodbergen and Vis (2009) and Gu et al. (2010) indicated that system capability assessment should not be based solely on single equipment quantities but should also consider factors such as storage structure, transport mechanisms, and operational workflows [2,21]. Lamballais et al. (2017) performed performance estimation for mobile robot fulfillment systems, showing that the number of mobile robots, workstation processing capacity, and operating rules affect system output [3]. These studies provide methodological inspiration for our use of input–output relationships to characterize smart-warehouse technical efficiency.

In terms of stochastic disturbances and complex system operations, Roy et al. (2012) and Boysen et al. (2019) both pointed out that order demand fluctuations, equipment operating status, system congestion, and operational organizational changes affect actual warehouse performance [15,22]. Azadeh et al. (2019), in their study on the development of automated warehouse systems, also indicated that because smart warehouse operations feature multi-equipment, multi-task, and multi-constraint interactions, configuring resources based solely on average demand or static capacity may fail to reflect real efficiency losses and resource redundancies in actual operations [23].

Although studies have analyzed automated warehouse and mobile robot fulfillment system performance from various perspectives, most have focused on system design, storage assignment, order allocation, or operating rule optimization—as in Weidinger et al. (2018) and Zhao and Zhang (2025)—and rarely converted efficiency frontier identification under random disturbances into joint optimization decisions for AGV and workstation counts [5,9]. To bridge the gap between efficiency measurement and resource decisions, this paper uses SFA to identify random disturbances and technical inefficiency and, on that basis, builds an MINLP model for discrete resource optimization.

2.3. Summary of research status

Research on smart-warehouse resource allocation has borne fruit, but several shortcomings remain. First, most studies have analyzed only individual links—order picking, automated storage, robot transport, or workstation allocation—and neglected system-wide collaborative configuration. Second, the literature pays attention mostly to local efficiency improvement, and researchers have yet to systematically investigate the quantitative ratio of robots to workstations. For example, although Boysen et al. (2017) and Zhao and Zhang (2025) included robot transport and workstation processing, their emphases were still on order processing or allocation strategy optimization [9,24]. Third, many relevant models have been based on deterministic or idealized assumptions, with insufficient consideration of stochastic factors. Overall, existing results lay a foundation for smart-warehouse resource optimization, but weaknesses remain, such as isolated optimization of single resources, insufficient characterization of collaborative relationships, and inadequate consideration of uncertainty.

Therefore, it is necessary to overcome these shortcomings and, on the basis of existing research, develop a smart-warehouse resource allocation model that integrates stochastic disturbance identification and discrete resource optimization. Combining SFA and MINLP, through the processing of real operational data, this paper first identifies the system efficiency frontier under stochastic

disturbances and then optimizes the AGV and workstation configurations under efficiency constraints. The result is a balanced trade-off among resource input, operational efficiency, and operating costs in the smart warehouse system.

3. Model construction

3.1. Problem description

This paper studies an automated-warehouse resource-allocation system composed of a smart warehouse environment, differentiated order demand users, and AGVs and workstations with different configurations. The warehouse operator, as the service provider, needs to meet order timeliness and processing capacity requirements. However, stochastic disturbances—such as natural conditions, order structure, and equipment status—affect the optimal allocation of core production factors, thereby constraining order processing efficiency and service levels. As illustrated in Figure 1, the first stage uses input-output sample data to fit a production function that includes stochastic disturbances and technical inefficiency components. It estimates output elasticities for workstations and AGVs, quantifies the system-level technical efficiency, and detects returns-to-scale properties. Consequently, an estimated stochastic production frontier (with random noise separated out) and corresponding input-output constraint relationships are established. The second stage, taking the production function parameters, factor elasticities, and efficiency frontier estimated by SFA as core constraints, combined with integer requirements for workstation and AGV counts, builds an MINLP model aimed at minimizing weighted resource costs, solves for the optimal integer input levels under different demand levels, and determines resource allocation plans that accommodate stochastic disturbances while guaranteeing output.

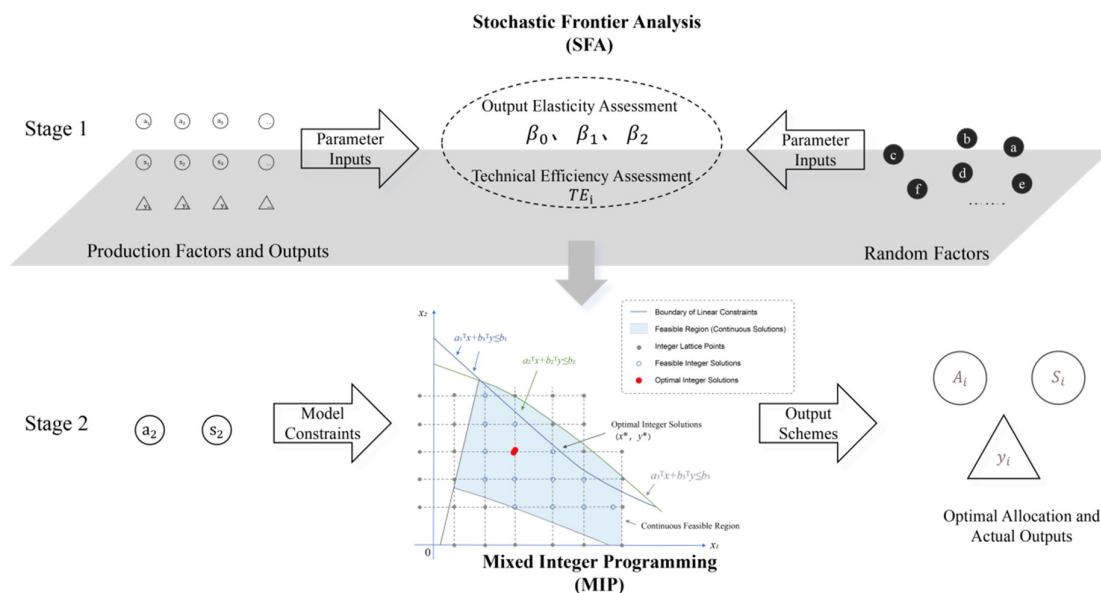


Figure 1. Schematic diagram of joint optimization framework.

In the joint optimization model constructed in this paper, it is assumed that all AGVs initially have 100% battery charge, excluding the situation where tasks cannot be performed due to insufficient power. Following common settings in existing studies of smart-warehouse resource allocation and actual warehouse operational characteristics, it is assumed that each AGV single-transport task involves visiting a workstation only once, without redundant repeated trips. Given that the research focus is on resource allocation optimization rather than real-time scheduling, to simplify the model solution, overtime penalty costs for workstations and AGV queuing waiting costs are not considered. Accordingly, this study only considers stochastic disturbances that can be characterized by SFA, while several real-time operational dynamics are beyond the analytical scope.

The basic model parameters are defined in Table 1.

Table 1. Definition of model symbols.

Symbol	Symbol definition
Set symbols:	
I	Set of all time periods
Parameter symbols:	
α_1	Share of comprehensive workstation cost in the weighted objective function
α_2	Share of comprehensive AGV cost in the weighted objective function
x_{in}	Actual input vector of n production factors corresponding to sample point i
y_i	Actual output of sample point i
β_n	Factor output elasticities under the stochastic production environment
$\hat{\beta}_n$	Factor output elasticities under the deterministic production environment
v_i	Stochastic disturbance term
u_i	Technical inefficiency term
TE_i	Technical efficiency value of sample point i
Decision variables:	
Z	Objective function value
$f(x_i; \beta)$	Production function corresponding to ideal maximum output
agv	Optimal deployment quantity of AGVs
$station$	Optimal activation quantity of STATIONS

The joint optimization logic of technical efficiency and optimal configuration for the smart warehouse is as follows:

First, in fitting the production function, the workstation output elasticity β_1 and AGV output elasticity β_2 directly correspond to the marginal contribution differences between labor and equipment inputs in warehouse operations. β_1 incorporates factors such as operator proficiency, shift efficiency, and error rates, reflecting the actual impact of labor input on order output; β_2 incorporates factors such as equipment failure rates, charging losses, and path congestion, reflecting the actual impact of equipment input on order output. The technical inefficiency term u_i further quantifies output losses caused by stochastic factors; by quantifying the inefficiency intensity, the model can effectively capture the level of uncertainty of the warehouse system.

Second, in resource allocation optimization, the model retains and extends the core parameters of classical MINLP. In addition to the unit input costs of AGVs and workstations at each time i , this paper

places special emphasis on the requirement that the value of the production function $f(x_i; \beta)$ must not be less than the required output y_i at that time.

3.2. Problem description

To characterize the input–output relationship of smart warehousing and separate stochastic factors from efficiency losses, the SFA method is adopted. With the picking quantity at time i as actual output y_i , and the number of open workstations and active AGVs as the input vector x_i , the basic SFA model can be expressed as

$$y_i = f(x_i; \beta) \cdot \exp \{v_i\} \cdot TE_i, \quad i=1,2,\dots,I \quad (1)$$

where $f(x_i; \beta)$ is the production function, representing the “ideal” order quantity produced with given inputs at time i ; the parameter β is the production elasticity, quantified under the combined effect of real factors, that truly reflects the marginal impacts of changes in the two core resources on output in the smart warehouse scenario; $v_i \sim N(0, \sigma_v^2)$ is the random disturbance term; $\exp \{v_i\}$ captures unobserved stochastic disturbances arising from factors such as weather fluctuations, differences in worker proficiency, varying equipment status, order structure, promotional events, and rack layout. Nevertheless, the individual effect of each factor cannot be separately identified in the current model; $f(x_i; \beta) \cdot \exp \{v_i\}$ is termed the stochastic production frontier. TE_i is the technical efficiency at time i ; to estimate the specific value of technical efficiency, let $TE_i = \exp \{-u_i\}$, where u_i is the technical inefficiency term, $u_i \geq 0$, and follows a half-normal distribution $u_i \sim N^+(0, \sigma_u^2)$, representing efficiency losses such as resource redundancy and poor operational coordination. Equation (1) indicates that at time i , given inputs x_i , the ideal output should be $f(x_i; \beta)$, but under the influence of various stochastic factors v_i and inefficient AGV or workstation configurations u_i , the actual output is y_i .

The relationship between actual output, ideal output, and the production function is illustrated schematically in Figure 2.

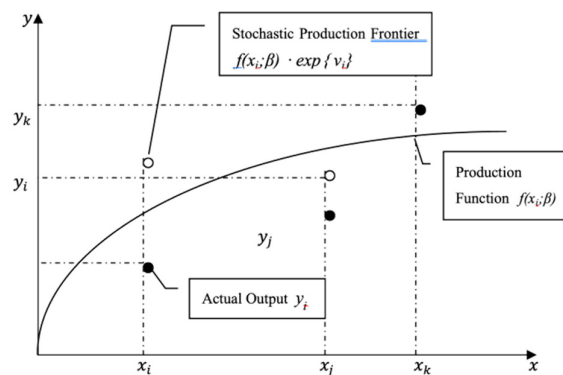


Figure 2. Schematic of production function and technical efficiency measurement in the SFA model.

In Figure 2, i , j , and k represent three possible cases in SFA, where $\bullet$ denotes actual output y , and $\circ$ denotes the stochastic production frontier $f(x; \beta) \cdot v$. When the favorable effect of the random disturbance is smaller than the adverse effect of technical inefficiency, the actual output lies below the ideal output; when both are adverse, both the stochastic frontier and actual output are below the ideal;

when the favorable effect of the random disturbance exceeds the adverse effect of technical inefficiency, both the stochastic frontier and actual output are above the ideal.

3.3. Problem description

The core aim of MINLP is to minimize or maximize an objective function subject to linear equality and inequality constraints, with some decision variables required to take integer values. MINLP is therefore well suited to the practical scenario of smart warehousing, where the AGV and workstation counts are integers. Fatthi et al. (2016), constructing an MINLP model for real-time door assignment and scheduling in cross-dock warehouses, demonstrated that the MINLP approach is suitable for discrete resource allocation and scheduling decisions in warehouse settings [25].

Based on the input–output relationship, factor output elasticities, and technical efficiency benchmark derived from the first-stage SFA model, this paper constructs a dedicated MINLP model for smart warehouse resource allocation, focusing on the optimal ratio of AGVs to workstations. To better align with actual warehouse operational needs and enhance practical applicability, combined with cost control objectives and based on the cost share structure of AGVs and workstations, a weighted objective function is constructed. Here, α_2 is the proportion of AGV costs (e.g., depreciation, energy, maintenance) in total resource costs, and α_1 is the proportion of workstation costs (e.g., labor, space, equipment depreciation) in total resource costs. By introducing cost weight coefficients, the model can simultaneously optimize resource input quantities and minimize the dimensionless weighted objective value, ensuring both order-picking demand satisfaction and reasonable resource-input control in warehouse operations. Combined with the input–output laws determined in the SFA stage, redundant calculation is avoided; finally, through model solving, the optimal numbers of AGVs and workstations are obtained, providing a practical quantitative solution for efficient smart-warehouse resource utilization.

After estimating the parameter β in the production function $f(x_i; \beta)$ via SFA, the specific functional relationship between inputs and output can be determined. The fitted value $f(x_i; \beta)$ represents the predicted maximum output achievable under given inputs x .

$$\min Z = \alpha_2 \frac{agv}{agv_{max}} + \alpha_1 \frac{station}{station_{max}} \quad (2)$$

$$s.t. \quad f(agv, station; \hat{\beta}) \geq y \quad (3)$$

$$1 \leq agvs \leq agv_{max} \quad (4)$$

$$1 \leq stations \leq station_{max} \quad (5)$$

$$agv, station \in Z^+ \quad (6)$$

In the above MINLP model, agv and $station$ are decision variables representing the numbers of active AGVs and open workstations. Considering the different units for AGVs and workstations, objective function (2) represents minimizing the sum of the normalized numbers of AGVs and workstations, where agv_{max} and $station_{max}$ denote the maximum available AGVs and workstations, respectively. Note that optimization outputs are sensitive to the setting of these practical upper-bound

parameters: constraint (3) ensures that the predicted output computed by the production function is not less than the future order output requirement y ; constraints (4) and (5) specify the allowable ranges for AGVs and workstations; and constraint (6) requires that the numbers of AGVs and workstations be positive integers. Notably, we employ frontier production-function parameters rather than sample-specific TE_i for constructing optimization constraints, which corresponds to our goal of identifying ideal resource allocation on the efficiency frontier.

4. Algorithm design

4.1. Parameter estimation by the SFA model

To estimate the parameter β and the technical efficiency TE_i , it is first necessary to select an appropriate production function to separate the random disturbance term v_i from the technical inefficiency term u_i . In this study, the Cobb-Douglas (C-D) production function and the Translog production function are constructed simultaneously, and the maximum likelihood estimation (MLE) method is employed to estimate the parameters of the warehousing system, decompose random disturbances, and measure technical efficiency.

Initial parameter values are specified, and two joint log-likelihood functions are constructed according to the distributional characteristics of the random disturbance term and the inefficiency term. The MLE algorithm is then used to iteratively maximize the log-likelihood functions and obtain the regression estimates of all parameters. After parameter estimation, t -tests, economic significance tests, and likelihood ratio tests are conducted sequentially to comprehensively evaluate the goodness of fit and rationality of the two production functions. The specific solution process is as follows.

Assume that the production function with N inputs is given by

$$f(x_i; \beta) = \exp(\beta_0) \prod_{n=1}^N x_n^{\beta_n} \quad (7)$$

Using the C-D specification for $f(x_i; \beta)$, taking the natural logarithm of both sides of Equation (1) yields the basic SFA model:

$$\ln y_i = \beta_0 + \sum_{n=1}^N \beta_n \ln x_{ni} + v_i - u_i \quad i=1, 2, \dots, I \quad (8)$$

Let $\varepsilon_i = v_i - u_i$ denote the composite error term, which consists of two components: the random disturbance term v_i and the technical inefficiency term u_i .

Since $E\varepsilon_i = Ev_i - Eu_i = -Eu_i < 0$, the ordinary least squares method cannot be directly applied to estimate β and ε . In SFA, the parameters are first estimated using the MLE method, including the parameter vector β and the distributional parameters of v_i and u_i . Subsequently, the conditional distribution is employed to separate u_i from ε_i , thereby enabling the estimation of technical efficiency as $TE_i = \exp\{-u_i\}$. Specifically, let $f(v)$ denote the probability density function of v , $g(u)$ denote the probability density function of u , $h(u, v)$ denote the joint probability density function of u and v , and $p(\varepsilon)$ denote the probability density function of ε .

$$h(u, v) = \frac{1}{\pi \sigma_u \sigma_v} \exp \left\{ -\frac{u^2}{2\sigma_u^2} - \frac{v^2}{2\sigma_v^2} \right\} \quad (9)$$

$$p(u, \varepsilon) = \frac{1 + \lambda^2}{\pi \lambda \sigma^2} \exp \left\{ -\frac{u^2(1 + \lambda^2)}{2\lambda^2 \sigma^2} - \frac{(\varepsilon + u)^2(1 + \lambda^2)}{2\sigma^2} \right\} \quad (10)$$

Letting $\sigma^2 = \sigma_u^2 + \sigma_v^2$, $\lambda = \frac{\sigma_u}{\sigma_v}$, the probability density function of ε can be expressed as

$$p(\varepsilon) = \frac{2}{\sigma} \cdot \varphi\left(\frac{\varepsilon}{\sigma}\right) \cdot \Phi\left(-\frac{\varepsilon\lambda}{\sigma}\right) \quad (11)$$

where $\varphi(\cdot)$ and $\Phi(\cdot)$ denote the probability density function and cumulative distribution function of the standard normal distribution, respectively.

Since $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_I$ are independent and identically distributed random variables, the joint likelihood function L is given by

$$L = p(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_I) = \prod_{i=1}^I p(\varepsilon_i) = \left(\frac{2}{\sigma}\right)^I \prod_{i=1}^I \left[\varphi\left(\frac{\varepsilon_i}{\sigma}\right) \Phi\left(-\frac{\varepsilon_i\lambda}{\sigma}\right)\right] \quad (12)$$

Taking the natural logarithm of both sides of Equation (12) yields the log-likelihood function

$$\ln L = c - I \ln \sigma + \sum_{i=1}^I \ln \Phi\left(-\frac{\varepsilon_i\lambda}{\sigma}\right) - \frac{1}{2\sigma^2} \sum_{i=1}^I \varepsilon_i^2 \quad (13)$$

where c is a constant, and $\varepsilon_i = \ln y_i - \beta_0 - \sum_{n=1}^N \beta_n \ln x_{ni}$.

The partial derivatives of Equation (13) are then taken sequentially with respect to the parameters $(\beta_0, \beta_1, \dots, \beta_n, \sigma_v^2, \sigma_u^2)$. Solving the resulting first-order conditions yields the maximum likelihood estimates of $(\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_n, \hat{\sigma}_v^2, \hat{\sigma}_u^2)$ and the residuals $\hat{\varepsilon}_i$ ($i = 1, 2, \dots, I$). Since the distributions of v , u , and ε are known, the technical inefficiency term u_i can be separated from ε_i . The estimate of technical efficiency is then obtained from the conditional expectation of $\exp\{-u_i\}$.

$$TE_i = E(\exp\{-u_i\} | \varepsilon_i) = \frac{\Phi(\mu_{*i}/\sigma_* - \sigma_*)}{\Phi(\mu_{*i}/\sigma_*)} \cdot \exp\left\{-\mu_{*i} + \frac{1}{2}\sigma_*^2\right\} \quad (14)$$

where $\mu_{*i} = -\varepsilon_i \hat{\sigma}_u^2 / \sigma^2$, $\sigma_*^2 = \hat{\sigma}_u^2 \hat{\sigma}_v^2 / \sigma^2$, and $\sigma^2 = \hat{\sigma}_v^2 + \hat{\sigma}_u^2$, and $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution.

4.2. Branch-and-bound algorithm

In this stage, an optimization model is developed for two core discrete operational resources of the intelligent warehouse, i.e., AGVs and picking workstations. With the minimization of weighted comprehensive warehouse operational cost as the objective, a mixed integer nonlinear programming (MINLP) model is formulated, subject to the nonlinear input-output relationship estimated via SFA and the upper bounds of equipment quantity capacity.

Traditional linear programming solvers are incapable of addressing nonlinear constraints, whereas generic heuristic algorithms exhibit limitations including weak global search ability and unstable optimization outputs. To satisfy the demand for batch computation using 658 sets of time-

series operational samples in this study, the branch-and-bound algorithm integrated with a nonlinear subproblem solution strategy is employed to cope with integer constraints. This solution framework inherently accommodates nonlinear constraints and integer variables. For discrete resource optimization in warehousing systems, it achieves rapid convergence and favorable computational stability, thereby fulfilling the engineering requirements of multi-period batch resource optimization for intelligent warehouses.

4.3. Algorithm flowchart

The workflow of the proposed SFA–MINLP-integrated algorithm is illustrated in Figure 3.

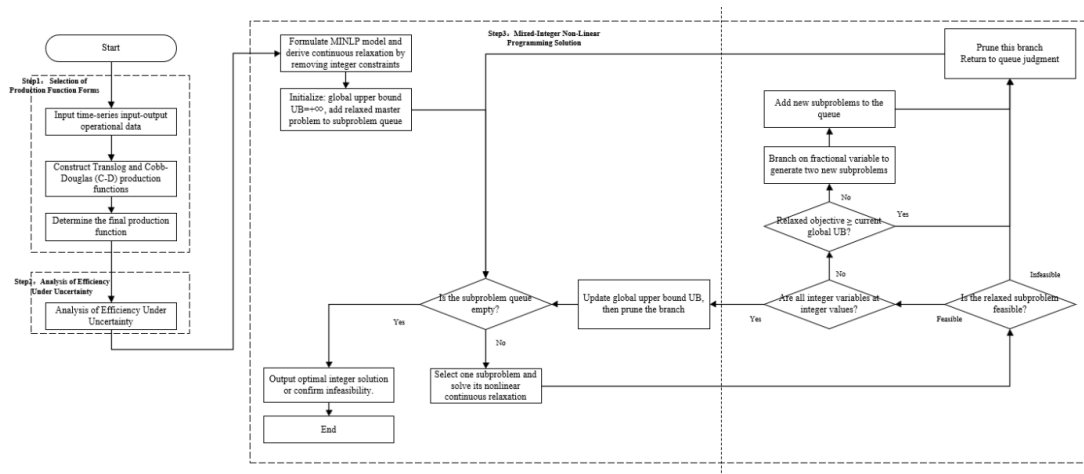


Figure 3. Integrated algorithm framework.

The proposed algorithm integrates the advantages of SFA and MINLP to address the complex resource allocation problem in intelligent warehousing under stochastic disturbances. The implementation consists of the following three steps:

Step 1: Production function selection: A dual-model fitting strategy based on SFA, together with statistical hypothesis testing, is employed to determine the appropriate production function. This approach mitigates the bias associated with subjective model selection and ensures that the estimated efficiency frontier accurately reflects the operational characteristics of intelligent warehousing.

Step 2: Efficiency analysis under uncertainty: The SFA model is employed to separate random noise from technical inefficiency, thereby estimating the output elasticities of AGVs and picking workstations. This allows accurate identification of the warehousing system's estimated stochastic production frontier, while accounting for stochastic disturbances including order fluctuations and equipment failures to mitigate their adverse impacts on resource-allocation decisions.

Step 3: Multi-resource coordinated optimization: An MINLP model is formulated within the GEKKO optimization framework to accommodate discrete decision variables and nonlinear constraints. The model simultaneously optimizes the allocation of multiple discrete warehousing resources, including AGVs and picking workstations, thereby achieving coordinated resource scheduling.

5. Empirical analysis

5.1. Introduction of empirical data

An empirical study is conducted using data from a 3C small-item intelligent warehouse operated by enterprise A, the largest e-commerce enterprise in China. The warehouse covers approximately 3000 m² and stores approximately 27,000 stock keeping units. On average, it processes 5000 customer orders and 7400 picked items per day. The warehouse is equipped with up to 80 AGVs and 7 picking workstations. Operational records indicate that the average AGV utilization rate is only 60%, increasing merely to 81% even during peak operating periods, while picking workstations frequently remain idle.

For empirical analysis, hourly picking time-series data of the intelligent warehouse, from August 1 to August 31, 2025, are selected. A total of 744 hourly observations could theoretically be obtained within this 31-day calendar cycle. Nevertheless, daily operation durations vary in practice. We have excluded invalid records of nighttime warehouse shutdowns and routine system maintenance during non-working hours and finally obtained 658 valid samples. The dataset covers picking quantity, the number of deployed AGVs, and activated workstations. Picking quantity is defined as the output indicator, whereas the quantities of AGVs and workstations are taken as core input factors. During the observation period, external disturbances are limited, and equipment scheduling strategies remain stable. The dataset covers both high-load and low-load operating conditions with satisfactory representativeness, which can objectively reflect the practical problems of misallocated resources and insufficient equipment utilization for AGVs and picking workstations in the intelligent warehouse.

The 658 valid samples consist of consecutive hourly time-series observations collected from a single warehouse, which implies potential temporal dependence and endogeneity between resource inputs and picking outputs. A Hausman-type endogeneity test conducted within the SFA framework confirms the presence of mild marginal endogeneity in the dataset. The null hypothesis for this test is: no systematic coefficient difference between two estimators. Such slight endogeneity only introduces limited estimation bias in the SFA specification and does not cause substantial deviations in the estimated production frontier and technical efficiency scores.

Since AGV and workstation counts vary over time, the hourly average number of active AGVs and operating workstations within each one-hour interval is employed to represent the corresponding equipment inputs. Descriptive statistics and scatter plots for the dataset are reported in Table 2 and Figure 4, respectively.

Table 2. Descriptive statistics of input and output variables.

Statistical index	Quantity	Station	Agv
Mean value	1013.6	3.5	41.9
Standard deviation	525.2	1.6	17.9
Minimum value	17.0	1.0	4.7
25% quantile	546.0	2.0	29.4
Median	1048.5	3.8	46.0
75% quantile	1414.0	5.0	55.6
Maximum value	2377.0	7.0	78.6

Note: Reported values are averaged from sample-level integer solutions. The same applies to subsequent tables.

As shown in Table 2, the warehouse operates an average of 41.9 AGVs and 3.5 picking workstations per hour, producing an average hourly picking volume of 1013.6 items. Both the input and output variables exhibit substantial variability. Across all observations, 50% of the samples involve between 29.4 and 55.6 AGVs, while the number of active workstations ranges from 3.8 to 5.0. During peak operating periods, all seven picking workstations are activated, with approximately 79 AGVs deployed simultaneously. The correlation coefficients among the variables and the corresponding scatterplots are reported in Table 3 and Figure 4, respectively.

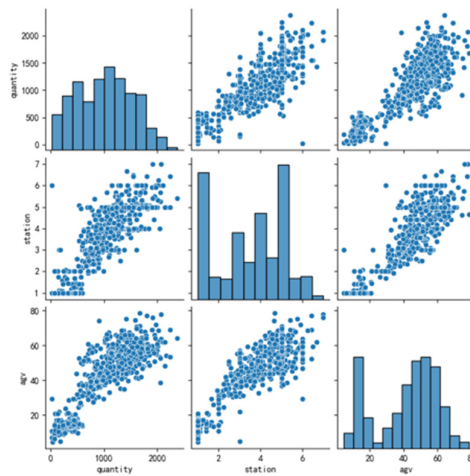


Figure 4. Scatter diagram of correlation among variables.

Table 3. Correlation coefficient matrix of variables.

	Quantity	Station	Agv
Quantity	1.00	0.86	0.85
Station	0.86	1.00	0.90
Agv	0.85	0.90	1.00

As illustrated in Figure 4 and Table 3, both workstation and AGV counts exhibit significant positive correlations with the picking output, indicating a strong complementary relationship between the two types of operational resources. However, considerable dispersion in the input–output observations is also evident, suggesting that warehouse operations are substantially influenced by stochastic factors, including fluctuations in order demand, equipment operating conditions, and workforce productivity. Consequently, simple correlation analysis is insufficient to quantify the respective effects of random disturbances and technical inefficiency, nor can it directly determine the optimal integer allocation of AGVs and picking workstations.

Therefore, it is necessary to first employ SFA to separate random noise from technical inefficiency, thereby estimating the underlying production frontier and measuring the output elasticities and technical efficiency of the input factors. These estimates subsequently provide the basis for determining the optimal allocation of AGVs and picking workstations that jointly minimizes operating costs while maintaining production efficiency.

5.2. Selection of production function form

The Translog and C-D production functions are two well-established functional forms in SFA and have been widely adopted in studies of warehousing logistics and intelligent equipment configuration. Compared with the Translog specification, the C-D production function is computationally more tractable and yields parameters with clearer economic interpretations, allowing the output elasticities of production factors to be estimated directly. These characteristics make it particularly suitable for analyzing the large-scale warehouse time-series dataset considered in this study.

Accordingly, the production function for each observation period is estimated using both the Translog and C-D specifications. After parameter estimation, *t*-tests are performed to evaluate the statistical significance of the estimated coefficients. The regression results are then compared to identify the production function that best represents the operational characteristics of the intelligent warehouse. Based on the selected model, the output elasticities of the two input factors, picking workstations and AGVs, are estimated accurately.

Using the operational dataset described above, the estimated parameters of the Translog and C-D production functions are reported in Table 4.

Table 4. Parameter estimation results of two production function models.

Translog production function estimation results					C-D production function estimation results				
Parameter	Coefficient	Standard deviation	<i>t</i>	<i>P</i>	Parameter	Coefficient	Standard deviation	<i>t</i>	<i>P</i>
β_0	2.495	0.453	5.503	0.000	β_0	4.485	0.187	23.971	0.000
β_1	0.009	0.288	0.033	0.974	β_1	0.339	0.059	5.792	0.000
β_2	1.808	0.300	6.019	0.000	β_2	0.633	0.068	9.275	0.000
β_3	0.344	0.119	2.884	0.004	λ	5.029	0.605	8.310	0.000
β_4	-0.157	0.061	-2.587	0.010	log-likelihood		222.567		
β_5	-0.085	0.148	-0.576	0.565					
λ	4.213	0.475	8.875	0.000					
log-likelihood			204.375						

It can be observed from Table 4 that several coefficient estimates of the Translog production function show counter-intuitive economic signs, including negative values for β_4 and β_5 . Moreover, parameters β_1 and β_5 fail to pass the *t*-test at the 95% statistical significance level ($p = 0.974$ and $p = 0.565$, respectively). Although the Translog model yields a log-likelihood value of 204.375, this value is lower than the log-likelihood value (222.567) obtained from the C-D specification. Taken together, the conflicting coefficient signs, insignificant parameter estimates, and inferior log-likelihood performance suggest that the Translog specification is not suitable for the dataset of this intelligent-warehouse case.

By contrast, for the C-D production function, all core coefficients β_0 , β_1 , and β_2 exhibit economically expected positive signs and pass the 95% significance *t*-test. The log-likelihood of the C-D model is higher than that of the Translog counterpart. Therefore, the C-D production function is selected as the final functional form for subsequent SFA parameter estimation.

Based on the estimation results of the C-D production function, the parameter $\lambda = \frac{\sigma_u}{\sigma_v}$ is estimated to be 5.029 and is statistically significant. As a core indicator for judging the applicability of the SFA model, this result reveals that the technical inefficiency term u_i dominates the composite

error term $\varepsilon_i = v_i - u_i$. Meanwhile, the stochastic disturbance term v_i exerts non-negligible effects, demonstrating that stochastic factors create tangible disturbances to warehouse output.

Analysis based on Table 5 shows that the average technical efficiency across all sample observations is only 0.685, which indicates considerable room for efficiency improvement in the warehouse system. In actual warehouse operations, stochastic factors such as weather fluctuations, differences in operator proficiency, and variations in equipment status lead to output fluctuations; these external disturbances correspond to the stochastic disturbance term v_i in the SFA model. By contrast, technical inefficiency arising from inappropriate resource allocation and unreasonable operational scheduling is captured by the technical inefficiency term u_i . Collectively, these results verify the practical rationality of incorporating stochastic disturbances in this study and confirm that the SFA framework can effectively separate random noise from technical inefficiency and accurately characterize the warehouse production frontier.

Table 5. Descriptive statistical results of technical efficiency.

Statistical index	Technical efficiency (TE)
Mean value	0.685
Standard deviation	0.174
Minimum value	0.014
25% quantile	0.595
Median	0.705
75% quantile	0.814
Maximum value	0.965

As reported in Table 4, the estimated output elasticity of workstations is 0.339, while that of AGVs is 0.633. Both elasticity estimates fall within a reasonable range, and their sum is 0.972, indicating that the intelligent warehousing system operates under approximately constant returns to scale. This finding implies that although increasing the numbers of picking workstations and AGVs simultaneously can improve picking output, the marginal increase in output gradually diminishes as resource inputs continue to expand. These findings provide strong justification for optimizing resource allocation at the current production level. Accordingly, an MINLP model is employed to determine the optimal workstation and AGV values, thereby avoiding excessive resource investment while simultaneously reducing operating costs and improving resource utilization.

The boxplot in Figure 5 provides further evidence of the distributional characteristics of technical efficiency. Specifically, it reveals a pronounced dispersion in efficiency values, accompanied by several low-efficiency outliers. Most observations are concentrated within the low-to-medium efficiency range, whereas highly efficient operating periods account for only a small proportion of the sample. In addition, several observations exhibit exceptionally low efficiency. These results indicate that warehouse operations are continuously affected by stochastic factors, including fluctuations in order demand, AGV start-up and shutdown losses, differences in workstation operator performance, and imbalances in task allocation. Consequently, the real-time deployment of AGVs and picking workstations does not closely match the underlying production frontier, resulting in widespread resource redundancy and suboptimal coordination between the two types of resources.

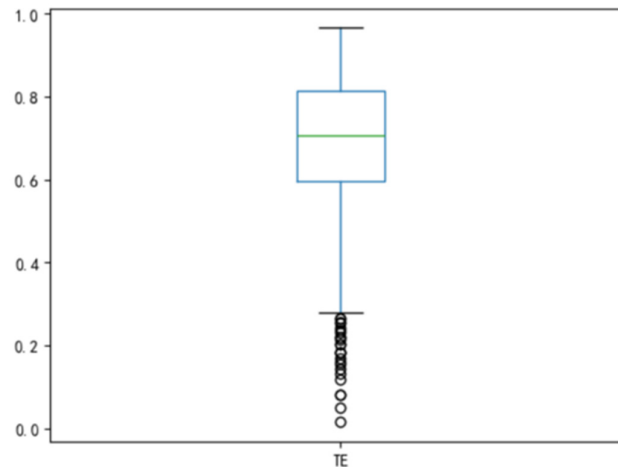


Figure 5. Box plot of technical efficiency distribution.

5.3. Joint optimization of the SFA–MINLP model

The average technical efficiency of the intelligent warehouse reaches merely 0.685, which reveals redundant resource allocation in system inputs.

According to the system configuration described previously, the intelligent warehouse can deploy a maximum of 80 AGVs and 7 picking workstations. Based on the results of the t-tests, the C-D production function is selected as the final production function. The parameter estimates obtained from the SFA model, namely β_0 , β_1 , and β_2 , are incorporated into the model constraints.

Considering the actual operating cost structure of the warehouse, the comprehensive picking-workstation cost, including labor costs, workstation occupancy costs, and equipment depreciation, accounts for approximately 60% of the total resource cost and is denoted by α_1 . The comprehensive cost of AGVs, including depreciation, energy consumption, and routine maintenance, accounts for approximately 40% of the total resource cost and is denoted by α_2 . Based on these settings, the MINLP model is formulated as follows:

$$\min Z = 0.4 \frac{agv}{80} + 0.6 \frac{station}{7} \quad (15)$$

$$s.t. \quad 4.485 + 0.339 \ln station + 0.633 \ln agv \geq \ln y \quad (16)$$

$$1 \leq agv \leq 80 \quad (17)$$

$$1 \leq station \leq 7 \quad (18)$$

$$agv, station \in Z^+ \quad (19)$$

In Equation (16), y denotes the hourly picking demand and is treated as a given parameter in this study. To evaluate the effectiveness of the proposed MINLP model in reducing resource inputs, the actual picking output of each observation is used as the value of y . The model is then solved with

the objective of minimizing the total operating cost associated with AGVs and picking workstations, thereby obtaining the optimal values of AGV and workstation counts for each observation. The optimized resource allocation is subsequently compared with the actual allocation to quantify the reduction in resource inputs.

The dataset contains 658 time series observations. The MINLP model is solved independently for each observation to obtain the corresponding optimal resource allocation, and the average values across all observations are reported as the optimal values of AGV and workstation counts. It should be noted that, according to the constraints defined in Equations (17) and (18), the optimization problem is formulated as an MINLP problem. The model is solved using the GEKKO optimization solver, with the initial values of both AGV and workstation counts set to 1.

Table 6 summarizes the optimized values of workstation and AGV counts, together with the percentage changes relative to the original resource allocation. As shown in Table 6, the average workstation count decreases to 1.5, representing an average reduction of 48% compared with the original allocation. The average AGV count decreases to 40.1, corresponding to an average reduction of 11%. Compared with AGV, the reduction in workstation is considerably greater. This result is consistent with the empirical findings presented above. Specifically, the output elasticity of AGV is higher than that of the workstation, indicating that AGVs contribute more substantially than picking workstations to warehouse output. Consequently, the optimization model tends to retain a larger AGV while reducing workstations to achieve the desired production level at a lower operating cost.

Table 6. Comparison of workstation and AGV quantity before and after optimization.

Statistical index	Pre-optimization scale		Post-optimization scale		Reduction amplitude	
	<i>Station</i>	<i>Agv</i>	<i>Station</i>	<i>Agv</i>	Station reduction rate	Agv reduction rate
Mean value	3.5	41.9	1.5	40.1	48%	11%
Standard deviation	1.6	17.9	0.7	22.5	26%	32%
Median	3.8	46.0	1.0	43.0	60%	9%

5.4. Model robustness test

To evaluate the robustness and generalizability of the proposed model under different operational workloads, warehouse operational scenarios are classified into three groups based on hourly picking volume; 200 observations are randomly selected for each group from all 658 samples to ensure balanced sample sizes, and the remaining 58 samples are reserved, excluded from robustness tests and prepared for out-of-sample validation in further extended research.

Small-scale scenario: 200 observations with fewer than 600 picked items per hour are selected, representing off-peak operations.

Medium-scale scenario: 200 observations with hourly picking volumes ranging from 600 to 1499 items are selected, representing normal daytime operations under regular workload conditions.

Large-scale scenario: 200 observations with 1500 or more picked items per hour are selected, representing peak outbound operations and high-intensity working periods during major promotional events.

Based on this classification, robustness tests are conducted separately for each operational scale. The corresponding empirical results are presented in Tables 7 and 8.

Table 7. Robustness analysis results of the SFA–MINLP joint optimization model under different load scales.

Load scale	Sample size	Solution success rate	Average configuration deviation rate
[0, 600)	200	1.00	0.021
[600, 1500)	200	1.00	0.033
[1500, 3000)	200	0.992	0.047

The results demonstrate that the proposed SFA–MINLP-integrated model can be solved reliably across all three operational scenarios. As reported in Table 7, the solution success rates are 100% for both the small-scale and medium-scale scenarios and 99.2% for the large-scale scenario, indicating excellent computational stability. To further evaluate the quality of the optimization results, the configuration deviation rate is defined as

$$\text{Configuration Deviation Rate} = \frac{|\text{Configuration-Reference MINLP Solution}|}{\text{Reference MINLP Solution}} \times 100\%.$$

Here, the *Reference MINLP solution* denotes the benchmark solution obtained by solving the identical MINLP model under complete information for each corresponding sample, which is adopted to compute the configuration deviation rate. It should be emphasized that this deviation metric evaluates the computational quality of the branch-and-bound solver relative to the reference solution of the same MINLP formulation. The average configuration deviation rates are 2.1%, 3.3%, and 4.7% for the small-scale, medium-scale, and large-scale scenarios, respectively, all remaining below 5%. These results indicate that the proposed model is capable of producing resource allocation solutions that remain close to the theoretical optimum under different operational workloads, including off-peak, normal, peak, and promotional periods. These results demonstrate that the proposed SFA–MINLP integrated model exhibits favorable computational stability across different historical workload groups. The SFA module separates random noise and technical inefficiency for observed historical samples; nevertheless, the stable optimized configurations within these retrospective sample groups cannot be interpreted as evidence of operational robustness under unseen future stochastic disturbances. Direct evaluation of operational robustness under independent or resampled disturbance scenarios remains a direction for future research.

Table 8. Optimization results of the SFA–MINLP joint model under different load scales.

Load scale	Pre-optimization resource scale		Post-optimization resource scale		Resource reduction amplitude	
	Station	Agv	Station	Agv	Station reduction rate	Agv reduction rate
[0, 600)	1.3	17.2	1.0	8.1	13.1%	52.5%
[600, 1500)	3.8	46.5	3.0	38.6	19.5%	16.6%
[1500, 3000)	5.1	56.5	6.7	25.9	−36.9%	51.6%

To further evaluate the optimization performance of the proposed model under different operational scales, the actual resource allocation before optimization is compared with the optimal allocation obtained by the proposed model for each operating scenario. The results are presented in

Table 8. In the small-scale scenario, the actual deployment of AGVs is substantially higher than required, indicating considerable resource redundancy. After optimization, the proposed model reduces the AGV count by 52.5% while still satisfying the picking demand during off-peak periods, thereby eliminating unnecessary resource waste. In the medium-scale scenario, the model determines an appropriate combination of AGVs and picking workstations according to the stable order workload. Compared with the actual allocation, the optimized configuration requires fewer resources while maintaining the required fulfillment efficiency, resulting in lower operating costs. In the large-scale scenario, picking operations approach the warehouse capacity limit, and the actual resource deployment is already at a relatively high level. Considering the operational characteristics of capacity-constrained conditions, the proposed model further optimizes the allocation of AGVs and picking workstations while ensuring the reliable completion of high-intensity outbound operations. Specifically, the workstation count is increased by 36.9%, whereas the AGV count is reduced by 51.6%, shifting the optimization strategy from increasing resource inputs to improving the allocation structure.

Overall, the proposed model can be solved reliably across different operating scenarios and operational scales, and the optimized configurations remain close to the reference MINLP solution of the underlying model. These results mainly demonstrate the computational stability and solution quality of the proposed branch-and-bound algorithm. Although SFA accounts for stochastic noise within observed historical data when constructing the efficiency frontier, this feature cannot guarantee that the optimized resource configurations will maintain satisfactory performance under unseen future stochastic disturbances. Therefore, comprehensive assessment of operational robustness under independent or resampled dynamic disturbance scenarios is reserved for future research.

5.5. Sensitivity analysis of weight coefficients

The weighting coefficients α_1 and α_2 represent the cost proportions of the two core resources in the intelligent warehouse, namely picking workstations and AGVs, respectively, with the constraint that $\alpha_1 + \alpha_2 = 1$. To investigate the sensitivity of the proposed model to different cost structures, a series of comparative experiments is conducted by systematically varying the two weighting coefficients. Specifically, α_1 is gradually decreased from 0.9 to 0.1, while α_2 is increased correspondingly from 0.1 to 0.9, covering a wide range of operating conditions, including workstation-oriented, balanced, and AGV-oriented optimization. For each parameter combination, the optimization model is solved under identical picking demand, resource availability, and operational constraints. The resulting objective values, optimal decision variables, and key operational performance indicators are then compared to evaluate the sensitivity of the proposed model to the weighting coefficients.

As reported in Table 9, the sensitivity analysis of α_1 and α_2 demonstrates that the optimization objective gradually shifts from minimizing AGV-related costs to minimizing workstation-related costs as α_1 increases and α_2 decreases. This change in the weighting coefficients leads to a clear substitution effect between the two types of production resources. Specifically, the optimal workstation count increases from 1.3 to 5.6, whereas the optimal AGV count decreases from 43.9 to 19.9, representing a maximum reduction of 54.8%. The opposite variation trends of the two decision variables are fully consistent with the expected behavior of a cost-oriented resource allocation model, indicating that the proposed model can effectively adjust the allocation of AGVs and picking workstations in response to changes in resource cost structures.

Table 9. Sensitivity analysis results of weight coefficients.

α_1	α_2	Optimal station	Optimal agv	Station variation amplitude	Agv variation amplitude	Objective function value
0.9	0.1	1.3	43.9	51.9%	3.0%	0.225
0.8	0.2	1.3	43.9	51.9%	3.0%	0.261
0.7	0.3	1.3	43.9	51.9%	3.0%	0.297
0.6	0.4	1.5	40.1	47.8%	10.4%	0.330
0.5	0.5	1.8	35.3	39.5%	19.7%	0.353
0.4	0.6	2.3	31.1	28.3%	28.2%	0.366
0.3	0.7	3.0	26.9	11.3%	36.7%	0.365
0.2	0.8	4.2	22.6	-25.5%	47.4%	0.345
0.1	0.9	5.6	19.9	-76.4%	54.8%	0.298

The sensitivity analysis indicates that the optimal AGV count is the most sensitive decision variable over the entire range of weighting coefficients, whereas the optimal workstation count exhibits high sensitivity only when $\alpha_1 < 0.7$ and remains relatively stable when $\alpha_1 > 0.7$. We explain that this stabilization phenomenon is mainly attributable to the lower-bound constraint of positive-integer decision variables. When $\alpha_1 > 0.7$, the model compresses workstation quantity as much as possible to reduce the weighted cost. Once the optimal workstation count reaches its feasible lower limit under the given demand level, further increasing α_1 can no longer reduce the number of workstations, so the optimal value shows little variation. The objective function displays an inverted U-shaped pattern, reaching its maximum value when $\alpha_1 = 0.4$ and $\alpha_2 = 0.6$. When the cost weight assigned to either resource becomes excessively large, the model substantially reduces the deployment of the corresponding high-cost resource, leading to a decline in the overall weighted operating cost. These findings suggest that although emphasizing the cost of a single resource can reduce its associated expenditure, excessive compression of one resource may result in an imbalanced resource allocation and increased implicit fulfillment costs. Therefore, a balanced weighting scheme is more appropriate for supporting long-term warehouse operations.

Accurate calibration of the weighting coefficients is essential to ensuring the effectiveness of the proposed optimization model and the reliability of its cost optimization results. In practical applications, particular attention should be paid to the estimation of the AGV cost weight. Because the optimal AGV allocation is considerably more sensitive to changes in the weighting coefficients than the workstation allocation is, the value of α_2 should be determined on the basis of the comprehensive operating cost of AGVs, including equipment depreciation, energy consumption, routine maintenance, and failure-related losses. This helps prevent inappropriate weighting from producing AGV allocation decisions that deviate substantially from actual operational requirements, thereby avoiding either resource redundancy or insufficient warehouse capacity.

5.6. Comparative analysis of algorithms

To further evaluate the performance of the proposed SFA–MINLP integrated optimization model, data envelopment analysis (DEA) is selected as the benchmark method for comparison. Quantitative comparisons are conducted under multiple demand fluctuation scenarios with picking output varying

by $\pm 20\%$ and $\pm 40\%$. The two approaches are evaluated from four perspectives, namely efficiency identification, resource allocation, adaptability to uncertainty, and practical applicability.

5.6.1. Theoretical overview of the DEA model

DEA is a nonparametric efficiency evaluation method that accommodates variable returns to scale without requiring a predefined production function. It involves constructing a relative production frontier through linear programming and evaluating the relative technical efficiency of warehouse decision-making units by maximizing picking output under given resource inputs. Although DEA offers the advantages of simple model formulation and does not require parameter estimation, it also has several inherent limitations. As a deterministic approach, DEA cannot distinguish random disturbances, such as demand fluctuations, equipment failures, and variations in worker productivity, from systematic technical inefficiency, making the estimated efficiency frontier susceptible to distortion by extreme observations.

Therefore, owing to the deterministic nature of the DEA model, it has been frequently adopted in recent journal literature for comparative analysis with the stochastic SFA model to verify the optimality of solutions under stochastic environments. For instance, Sakouvogui (2020) adopted both DEA and SFA for empirical comparison in the *Journal of Economic Studies* and analyzed their differences in measuring cost efficiency and resource allocation outcomes [26]. Sun et al. (2025) constructed a three-stage DEA-SFA model in PLOS ONE, with standalone DEA and standalone SFA set as control groups to examine their performance in estimating corporate resource allocation [27]. Rødseth et al. (2024) compared DEA, SFA, and StoNED models in the *Journal of Productivity Analysis* and carried out methodological tests focusing on allocative efficiency and optimal factor allocation [28].

Table 10. Comparative optimization results of the SFA–MINLP joint model and the DEA algorithm.

Output fluctuation amplitude	Algorithm category	Optimal station	Optimal agv	Station change rate	Agv change rate
40%	DEA model	5.7	58.9	54.1%	35.1%
20%		5.1	55.0	37.8%	26.1%
0		3.7	43.6	0	0
–20%		2.8	29.6	–24.3%	–32.1%
–40%		2.4	24.5	–35.1%	–43.8%
40%	SFA–MINLP model	2.1	50.8	40%	26.7%
20%		1.9	46.3	26.7%	15.5%
0		1.5	40.1	0	0
–20%		1.46	31.8	–2.7%	–20.7%
–40%		1.27	24.0	–15.3%	–40.1%

To further compare the resource allocation performance of the two approaches, demand fluctuation scenarios with picking output varying by $\pm 20\%$ and $\pm 40\%$ are generated based on the full dataset. The corresponding optimization results are summarized in Table 10.

As shown in Table 10, the DEA and SFA–MINLP models exhibit substantially different resource allocation patterns under varying demand conditions. Under the baseline scenario without demand fluctuations, the DEA model recommends 3.7 picking workstations and 43.6 AGVs. When demand increases by 20%, the optimal numbers increase to 5.1 workstations (37.8%) and 55.0 AGVs (26.1%).

When demand increases by 40%, the recommended configuration further rises to 5.7 workstations (54.1%) and 58.9 AGVs (35.1%). Conversely, both resources decrease proportionally as demand declines, indicating relatively large adjustments in resource allocation. In contrast, under the baseline scenario, the proposed SFA–MINLP model recommends 1.5 picking workstations and 40.1 AGVs. When demand increases by 20%, the optimal configuration changes to 1.9 workstations (26.7%) and 46.3 AGVs (15.5%). Under a 40% demand increase, the corresponding values become 2.1 workstations (40.1%) and 50.8 AGVs (26.7%). Resource inputs decrease accordingly as demand falls. Compared with the DEA model, the proposed SFA–MINLP model exhibits consistently smaller variations in resource allocation across all demand scenarios, resulting in smoother and more stable resource adjustment.

From a comparative perspective, these findings differ from prior warehouse efficiency studies in several important ways. Azadeh et al. (2019) and Lamballais et al. (2017) evaluated robotic fulfillment system performance primarily through analytical queueing models or simulation, which assume stationary demand and do not explicitly separate stochastic noise from technical inefficiency [3,23]. In contrast, the SFA–MINLP framework proposed in this study is driven by real operational data and explicitly accounts for demand fluctuations, equipment status variations, and worker productivity differences. Consequently, the resource allocation schemes derived from our model exhibit greater stability under demand uncertainty: across all $\pm 20\%$ and $\pm 40\%$ demand fluctuation scenarios, the SFA–MINLP model produces consistently smaller adjustments in both AGV and workstation counts compared with the DEA benchmark. This suggests that explicitly modeling stochastic disturbances can yield more stable resource decisions than deterministic evaluation approaches when operational data are subject to random noise.

5.6.2. Methodological comparison of COLS-MINLP and SFA-MINLP

To further examine the impact of stochastic disturbances on resource allocation decisions, this study additionally constructs deterministic benchmarks based on the ordinary least squares (OLS) and corrected ordinary least squares (COLS) frontiers. As a classical method for deterministic frontier analysis, COLS was originally developed for production frontier estimation and differs from the stochastic frontier analysis (SFA). Unlike SFA, the COLS method cannot decompose the total deviation from the production frontier into random noise and technical inefficiency, which constitutes the core distinction between the two models.

We first perform OLS regression on the log-linear Cobb-Douglas production function:

$$\ln \hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 \ln(\text{station}) + \hat{\beta}_2 \ln(\text{agv}) + \hat{\varepsilon}_i, \quad i = 1, 2, \dots, 658 \quad (20)$$

All 658 hourly observations are used for parameter estimation. After obtaining OLS fitted values $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 \ln(\text{station}) + \hat{\beta}_2 \ln(\text{agv})$, the log-scale residual for each sample is calculated as: $\hat{\varepsilon}_i = \ln(y_i) - \hat{Y}_i$.

We scan all 658 residual values and obtain the maximum residual $\hat{\varepsilon}_{\max} = 1.0474$. The COLS deterministic frontier intercept is obtained by shifting the OLS intercept upward by the magnitude of the maximum residual, while keeping the slope coefficients unchanged: $\beta_0^{\text{cols}} = \hat{\beta}_0 + \hat{\varepsilon}_{\max} = 2.7607 + 1.0474 = 3.8081$. Accordingly, the COLS deterministic log-form production frontier is written as $\ln y = 3.8081 + 0.0841 \ln(\text{station}) + 1.0712 \ln(\text{agv})$. This COLS frontier envelopes all observed sample points. It should be noted that COLS cannot disentangle stochastic noise and technical

inefficiency; all deviations between observations and the regression line are aggregated into residuals, which constitutes the major distinction from the stochastic SFA framework. The resulting COLS parameters are then embedded into the identical MINLP model for comparative resource-allocation optimization under $\alpha_1 = 0.6$ and $\alpha_2 = 0.4$.

Table 11. Comparative optimization results between original observed configuration, stochastic SFA-MINLP, and deterministic COLS-MINLP.

Indicator	Original observed configuration	Stochastic SFA-MINLP	Deterministic COLS-MINLP
Average number of workstations	3.49	1.5 (−48%)	1.00 (−71%)
Average number of AGVs	41.91	40.1 (−11%)	18.62 (−56%)
Weighted objective-function value	—	0.3298	0.1788

Note: The percentage values represent the resource-reduction rate relative to the original observed configuration.

As shown in Table 11, under the COLS parameter estimates ($\hat{\beta}_1 = 0.0841$, $\hat{\beta}_2 = 1.0712$), the output marginal contribution of workstations is extremely low: doubling the number of workstations only brings approximately 6% output growth. By contrast, AGVs exhibit a very high output elasticity. Consequently, increasing the workstation count can only marginally reduce the required quantity of AGVs to satisfy a given picking demand; the substitution benefit of adding workstations is very limited. Combined with the weighting setting $\alpha_1 = 0.6$ (higher cost weight for workstations), each additional workstation raises the weighted objective value substantially, while only achieving tiny savings on AGV resources. Therefore, the cost-minimizing solver pushes workstations down to the feasible lower bound of 1 and meets output requirements entirely by increasing the number of AGVs.

In comparison, the SFA-MINLP model obtains balanced interior solutions, because SFA disentangles stochastic noise and technical inefficiency and yields reasonable elasticities ($\hat{\beta}_1 = 0.3395$, $\hat{\beta}_2 = 0.6337$). In this stochastic-frontier setting, workstations and AGVs can effectively substitute for one another, generating trade-off-based resource allocation rather than extreme boundary solutions. This comparison demonstrates that deterministic frontier methods such as COLS are vulnerable to elasticity distortion under noisy operational data, which may lead to severely unbalanced resource-planning decisions for stochastic warehouse environments.

6. Conclusions and research prospects

6.1. Main research conclusions

To address the stochastic disturbances commonly encountered in intelligent e-commerce warehouses, including demand fluctuations, equipment failures, and variations in worker productivity, this study proposes an integrated SFA-MINLP optimization framework to overcome the limitations of existing studies in terms of theoretical foundation, optimization methodology, and practical applicability. By combining SFA with MINLP and validating the proposed approach using real operational data, this study establishes a new resource allocation framework for intelligent warehousing under stochastic operating conditions. The major findings of this study can be summarized from three perspectives, namely theory, methodology, and empirical validation.

First, this study introduces stochastic frontier production theory into the AGV-based goods-to-person warehousing system. The proposed SFA model separates random disturbances from technical inefficiency, enabling the estimation of the output elasticities of the two key production resources, namely picking workstations and AGVs, while identifying the system's efficiency frontier under stochastic operating conditions. This approach provides a clearer understanding of the input–output relationship within the coupled AGV–workstation system. From the perspective of production economics, the results explain why excessive investment in equipment may lead to diminishing marginal returns and resource redundancy. Consequently, the proposed framework addresses the lack of uncertainty-aware production theory in existing warehouse resource allocation studies and enriches the theoretical foundation for efficiency evaluation and resource allocation in robotic mobile fulfillment systems.

Second, this study addresses two major limitations of existing optimization approaches. Conventional SFA can evaluate technical efficiency but cannot provide operational resource allocation decisions, whereas standalone MINLP models generally neglect stochastic disturbances, resulting in limited robustness of the optimization outcomes. To overcome these limitations, this study develops an integrated SFA–MINLP optimization framework that combines the strengths of both methods. By incorporating the estimated efficiency frontier into the optimization process, the proposed model determines the optimal allocation of discrete warehouse resources under efficiency constraints. In doing so, it establishes a complete decision-making framework that links efficiency evaluation with resource allocation optimization, thereby providing a methodological foundation for intelligent warehouse resource planning under uncertainty.

Third, the proposed model is validated using real operational data collected from the intelligent warehouse of a large e-commerce company in China. The empirical results show that the average technical efficiency of the warehouse is only 0.685, indicating substantial resource misallocation and equipment underutilization. After applying the proposed SFA–MINLP model, while fully satisfying the required picking demand, the average number of active picking workstations is reduced by 48%, and the average number of deployed AGVs is reduced by 11%, which yields a reduction of resource inputs and a lower weighted objective-function value. These results demonstrate the effectiveness of the proposed model in identifying redundant resources and improving the allocation structure. Furthermore, the proposed model achieves an overall solution success rate of more than 99% across multiple operating scenarios, and the deviation of the optimized resource allocation from the theoretical optimum of the MINLP model remains below 5%, indicating favorable computational stability and solution quality. These findings, together with the SFA-based identification of the efficiency frontier under stochastic disturbances, suggest that the proposed approach can provide useful decision support for resource allocation in automated warehouses for e-commerce fulfillment.

6.2. Management insights

From a managerial perspective, the essence of resource allocation in intelligent warehousing lies not merely in determining the amount of resources to deploy but in achieving system-wide operational efficiency while satisfying order fulfillment requirements. The efficiency frontier analysis presented in this study indicates that warehouse performance is jointly influenced by resource allocation decisions, coordination among operational subsystems, and differences in operating scenarios. Accordingly, the managerial implications of this study can be discussed from three perspectives, namely resource utilization, system coordination, and scenario-specific resource allocation.

First, from the perspective of resource utilization, decisions regarding AGV deployment and workstation activation should not rely solely on historical operating experience or average order volumes. Instead, they should be guided by the production capacity implied by the estimated efficiency frontier. The empirical results show that the average technical efficiency of the sample warehouse is approximately 0.685, suggesting that operational inefficiency arises not only from insufficient resources but also from resource redundancy and mismatches in resource coordination under stochastic conditions, such as demand fluctuations, equipment status variations, and differences in worker productivity. Therefore, when determining the numbers of deployed AGVs and active picking workstations, warehouse managers should place greater emphasis on the underlying input–output efficiency relationship rather than relying exclusively on historical operating practices.

Second, from the perspective of system coordination, resource optimization in smart warehouses should shift from optimizing individual resources to the coordinated allocation of AGVs and picking workstations. Improving AGV operating efficiency alone does not necessarily increase overall warehouse throughput. If the processing capacity of picking workstations is not adjusted accordingly, bottlenecks such as waiting time, empty AGV travel, and workstation idleness may still occur. Likewise, simply increasing the number of picking workstations may not improve system performance if the available AGV capacity is insufficient to support them. Therefore, warehouse managers should jointly consider AGV transportation capacity, workstation processing capacity, and order demand within a unified decision-making framework, so that the allocation of the two types of resources can be dynamically adjusted according to operational requirements.

Third, from the perspective of scenario-specific resource allocation, resource planning should account for both the warehouse cost structure and the characteristics of different operating scenarios. The relative costs of AGV depreciation and maintenance, labor, and order processing intensity vary across warehouses, work shifts, and promotional periods. Consequently, applying a fixed resource allocation strategy may lead to inefficient resource utilization. Instead, managers should determine the weighting coefficients of different resources according to their actual cost structures and develop differentiated allocation strategies for off-peak, normal, and peak operating conditions. Such a demand-driven and cost-aware resource allocation strategy can improve operational stability while enhancing overall cost control.

Compared with existing literature on robotic mobile fulfillment systems, the managerial insights derived from this study are distinct in their empirical foundation and uncertainty awareness. Prior studies such as Boysen et al. (2019), Weidinger et al. (2018), and Zhao and Zhang (2025) have provided scheduling rules, storage assignment strategies, or order allocation policies primarily based on simulation experiments under assumed demand patterns [5,9,22]. While these studies offer valuable operational guidelines, their recommendations are often scenario-specific and may not generalize to warehouses with different cost structures or demand profiles. In contrast, the managerial insights of this study are derived from a data-driven stochastic efficiency frontier estimated from 658 real hourly observations. The finding that warehouse managers should shift from experience-based resource deployment to efficiency-frontier-guided allocation, and from single-resource optimization to coordinated AGV–workstation configuration, is grounded in the quantitative decomposition of random disturbance and technical inefficiency. Furthermore, the scenario-specific allocation strategy—differentiating off-peak, normal, and peak conditions with corresponding cost weight coefficients—provides a more nuanced and adaptable framework than the fixed strategies commonly proposed in the existing literature. These insights can inform warehouse

operators' resource-planning decisions and may be integrated into warehouse management systems for dynamic resource planning.

6.3. Future research prospects

This study investigates the allocation of AGVs and picking workstations in smart warehouses from the perspectives of efficiency frontier analysis and coordinated resource allocation. Nevertheless, several limitations remain and provide opportunities for future research. First, the SFA model developed in this study considers only two input variables, namely the numbers of AGVs and picking workstations. Other factors, such as order characteristics, warehouse layout, shelf distribution, worker proficiency, and the intensity of human–robot collaboration, are not explicitly incorporated into the analysis. If more fine-grained operational data become available, these variables could be included in the production function to obtain a more accurate representation of the production frontier under real operating conditions. Second, the proposed MINLP model focuses primarily on resource allocation at the planning horizon and does not fully capture dynamic operational constraints encountered during real-time warehouse operations. In practice, factors such as AGV battery levels, charging queues, equipment failures, unexpected surges in order demand, and traffic congestion may significantly affect the effective availability of warehouse resources within a given period. Future research could explicitly incorporate these dynamic factors into the optimization model, enabling resource allocation decisions that remain effective not only under static demand conditions but also in highly dynamic and uncertain operating environments. In addition, conducting out-of-sample or rolling-window tests to further examine the generalization ability and temporal stability of the model and optimized allocation schemes represents a valuable direction for follow-up research. Third, the empirical analysis is based on data from a single intelligent warehouse, and the findings therefore focus on the allocation of AGVs and picking workstations within an individual warehouse. Furthermore, mild endogeneity exists in the single-warehouse time-series dataset, and residual simultaneity bias cannot be fully eliminated due to data limitations. Future studies may conduct further econometric examinations using multi-warehouse datasets and extend the proposed framework to multi-warehouse systems, taking order allocation, complementary equipment capacities and cross-warehouse resource coordination into consideration. Such an extension would broaden the scope of resource optimization from a single warehouse to a network of interconnected warehouses, thereby providing more comprehensive decision support for large-scale intelligent warehousing operations.

Author contributions

CRedit: **Guanghui Fan:** Conceptualization, Methodology, Project administration, Resources, Writing – original draft; **Weihua Liu:** Conceptualization, Supervision, Writing – review & editing; **Fan Chao:** Investigation, Data curation, Validation, Writing – original draft; **Guodong Li:** Investigation, Validation, Writing – review & editing; **Hongxiao Zhang:** Data curation, Visualization, Writing – review & editing.

Use of Generative-AI tools declaration

No generative AI tools were used to conduct the research or draft the original manuscript. The research design, methodology, analysis, and conclusions were developed by the authors. The authors take full responsibility for the accuracy, originality, and integrity of the final manuscript.

Acknowledgments

The authors thank Enterprise A for providing the warehouse operational data used in this study and its staff for clarifying the warehouse processes and resource allocation practices. The authors also thank the editor and anonymous reviewers for their constructive comments and suggestions, which helped improve the manuscript. This research is also supported by Major Program of the National Social Science Foundation of China (Grant No. 24VRC060).

Conflict of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

1. S. Qu, X. Li, M. K. Lim, W. Liu, Z. Li, Y. Wang, Smart supply chain management: A systematic literature review and future research agenda, *Int. J. Logist. Res. Appl.*, (2025), 1–29. <https://doi.org/10.1080/13675567.2025.2569701>
2. K. J. Roodbergen, I. F. A. Vis, A survey of literature on automated storage and retrieval systems, *Eur. J. Oper. Res.*, **194** (2009), 343–362. <https://doi.org/10.1016/j.ejor.2008.01.038>
3. T. Lamballais, D. Roy, M. B. M. De Koster, Estimating performance in a robotic mobile fulfillment system, *Eur. J. Oper. Res.*, **256** (2017), 976–990. <https://doi.org/10.1016/j.ejor.2016.06.063>
4. D. Aigner, C. A. K. Lovell, P. Schmidt, Reprint of: Formulation and estimation of stochastic frontier production function models, *J. Econom.*, **234** (2023), 15–24. <https://doi.org/10.1016/j.jeconom.2023.01.023>
5. F. Weidinger, N. Boysen, D. Briskorn, Storage assignment with rack-moving mobile robots in KIVA warehouses, *Transp. Sci.*, **52** (2018), 1479–1495. <https://doi.org/10.1287/trsc.2018.0826>
6. B. Nicoletti, A. Appolloni, Warehouse 5.0 framework for a warehouse compatible with Industry 5.0, *Int. J. Logist. Res. Appl.*, **28** (2025), 1–25. <https://doi.org/10.1080/13675567.2025.2565224>
7. E. Guerrazzi, V. Mininno, D. Aloini, Balancing picking and outbound loading efficiency in an SBS/RS through a digital twin, *Flex. Serv. Manuf. J.*, (2024), 640–668. <https://doi.org/10.1007/s10696-024-09554-w>
8. B. Zhang, X. Liang, W. Song, Y. Chen, Multidimensional AGV path planning in 3D warehouses using ant colony optimization and advanced neural networks, *Commun. Comput. Inf. Sci.*, (2025), 180–191. https://doi.org/10.1007/978-981-96-9946-9_16
9. J. Zhao, C. Zhang, Order allocation strategy optimization in a goods-to-person robotic mobile fulfillment system with multiple picking stations, *Appl. Sci.*, **15** (2025), 9173. <https://doi.org/10.3390/app15169173>

10. S. Mahmoudinazlou, A. Sobhanan, H. Charkhgard, A. Eshragh, G. Dunn, Deep reinforcement learning for dynamic order picking in warehouse operations, *Comput. Oper. Res.*, **182** (2025), 107112. <https://doi.org/10.1016/j.cor.2025.107112>
11. C. K. M. Lee, Y. Lv, K. K. H. Ng, W. Ho, K. L. Choy, Design and application of Internet of Things-based warehouse management system for smart logistics, *Int. J. Prod. Res.*, **56** (2018), 2753–2768. <https://doi.org/10.1080/00207543.2017.1394592>
12. Í. R. da C. Barros, T. P. Nascimento, Robotic mobile fulfillment systems: A survey on recent developments and research opportunities, *Rob. Auton. Syst.*, **137** (2021), 103729. <https://doi.org/10.1016/j.robot.2021.103729>
13. R. de Koster, T. Le-Duc, K. J. Roodbergen, Design and control of warehouse order picking: A literature review, *Eur. J. Oper. Res.*, **182** (2007), 481–501. <https://doi.org/10.1016/j.ejor.2006.07.009>
14. G. Fragapane, R. de Koster, F. Sgarbossa, J. O. Strandhagen, Planning and control of autonomous mobile robots for intralogistics: Literature review and research agenda, *Eur. J. Oper. Res.*, **294** (2021), 405–426. <https://doi.org/10.1016/j.ejor.2021.01.019>
15. D. Roy, A. Krishnamurthy, S. S. Heragu, C. J. Malmborg, Performance analysis and design trade-offs in warehouses with autonomous vehicle technology, *IIE Trans.*, **44** (2012), 1045–1060. <https://doi.org/10.1080/0740817X.2012.665201>
16. C. A. Valle, J. E. Beasley, A two stage approach for order and rack allocation with order backlog in a mobile rack environment, *Ann. Oper. Res.*, **351** (2025), 167–201. <https://doi.org/10.1007/s10479-024-06058-6>
17. G. Li, J. Shao, D. Xu, W. Xu, The warehouse-retailer network design game, *J. Ind. Manag. Optim.*, **11** (2015), 291–305. <https://doi.org/10.3934/jimo.2015.11.291>
18. S. Benzidia, B. Ageron, O. Bentahar, J. Husson, Investigating automation and AGV in healthcare logistics: A case study based approach, *Int. J. Logist. Res. Appl.*, **22** (2019), 273–293. <https://doi.org/10.1080/13675567.2018.1518414>
19. M. Brunetti, M. Mes, E. Lalla-Ruiz, Smart logistics nodes: Concept and classification, *Int. J. Logist. Res. Appl.*, **27** (2024), 1984–2020. <https://doi.org/10.1080/13675567.2024.2327394>
20. P. Han, Y. Huo, W. Liu, E. Qi, H. Cai. The Development Strategy of Supply Chain Intelligent Technology Considering Technology Development Uncertainty. *Eur. J. Oper. Res.*, **328** (2026): 496–510. <https://doi.org/10.1016/j.ejor.2025.07.026>
21. J. Gu, M. Goetschalckx, L. F. McGinnis, Research on warehouse design and performance evaluation: A comprehensive review, *Eur. J. Oper. Res.*, **203** (2010), 539–549. <https://doi.org/10.1016/j.ejor.2009.07.031>
22. N. Boysen, R. de Koster, F. Weidinger, Warehousing in the e-commerce era: A survey, *Eur. J. Oper. Res.*, **277** (2019), 396–411. <https://doi.org/10.1016/j.ejor.2018.08.023>
23. K. Azadeh, R. de Koster, D. Roy, Robotized and automated warehouse systems: Review and recent developments, *Transp. Sci.*, **53** (2019), 917–945. <https://doi.org/10.1287/trsc.2018.0873>
24. N. Boysen, D. Briskorn, S. Emde, Parts-to-picker based order processing in a rack-moving mobile robots environment, *Eur. J. Oper. Res.*, **262** (2017), 550–562. <https://doi.org/10.1016/j.ejor.2017.03.053>
25. W. N. A. W. A. Fatthi, A. Shuib, R. M. Dom, A mixed integer programming model for solving real-time truck-to-door assignment and scheduling problem at cross docking warehouse, *J. Ind. Manag. Optim.*, **12** (2016), 431–447. <https://doi.org/10.3934/jimo.2016.12.431>

26. K. Sakouvogui, A comparative approach of stochastic frontier analysis and data envelopment analysis estimators: evidence from banking system, *J. Econ. Stud.*, **47** (2020), 1787–1810. <https://doi.org/10.1108/JES-01-2019-0051>
27. J. Sun, A. B. Rosli, A. Daud, Efficiency analysis of listed pharmaceutical companies in China: A method combining three-stage DEA with undesirable output, PCA, and tobit regression, *PLoS ONE*, **20** (2025). <https://doi.org/10.1371/journal.pone.0329767>
28. K. L. Rødseth, R. B. Holmen, T. Kuosmanen, H. Schøyen, Nonparametric estimation of allocative efficiency using indirect production theory: Application to container ports in Norway, *J. Prod. Anal.*, **62** (2024), 365–377. <https://doi.org/10.1007/s11123-024-00719-1>



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)