ISSN 1330-3651 (Print), ISSN 1848-6339 (Online)
Received: 25 May 2025; Accepted: 8 January 2026

https://doi.org/10.17559/TV-20250525002698
Original scientific paper

An Intelligent Purchasing Decision Model for Railway Wagon Components Using a Markov Decision Process

Shuo ZHANG, Chen LI, Hao HUA, Ping GAO, Wenzhuo WANG, Qiuli QIN*

Abstract: Overcoming the inefficiency in the procurement of railway wagon components is a crucial step in promoting the modernization of railway enterprises' governance capabilities. This study formulates components purchasing for railway wagon as a two-stage Markov Decision Process (MDP) that jointly optimizes planned purchase quantities and supplier selection. Inputs to the MDP (demand signals and inventory) are treated as pre-processed operational forecasts from the enterprise data platform; the paper therefore focuses on the design of state/action spaces, reward construction, and the value-iteration solution for producing actionable purchasing decision. Using real business records from a large railway equipment group, simulation results demonstrate that the MDP-based decision reduces total purchasing cost, shortens purchasing lead time, and improves supplier quality metrics compared with legacy manual strategies.

Keywords: Markov decision process; purchasing decision; railway wagon condition-based maintenance; smart purchase

1 INTRODUCTION

With the rapid advancement of the economy, inter-regional economic and trade exchanges have grown increasingly frequent, and the ever-increasing economic mobility has brought new requirements and challenges to the national freight transportation system. Railway enterprises need to continuously enhance their market competitiveness as well as operation and management capabilities to promote the modernization of enterprise governance capabilities. As a crucial link in production and operation, material purchase directly influences the operation quality and industry competitiveness of enterprises and plays a non-negligible role in enterprise cost control and efficiency improvement [1, 2]. Currently, railway freight car maintenance still mainly relies on planned preventive maintenance, supplemented by condition-based maintenance [3], but is gradually transforming to a condition-based maintenance (CBM) model. CBM has a positive effect on saving resource costs and promoting a sustainable and circular economy [4, 5], but its randomness is pronounced, making it challenging to control the material demand for maintenance work and subsequently presenting greater challenges to purchase management. Railway wagon's procurement bottlenecks directly impact lifecycle cost and operational reliability.

Railway wagon components purchase under condition-based maintenance is often hampered by highly subjective decision-making and poor planning accuracy. Motivated by the need to enhance objectivity and precision in this critical process, the main objectives of this study are to: (1) develop a robust purchasing decision-making model based on a Markov Decision Process (MDP) that integrates inventory levels, historical demand, and supplier performance; (2) identify and quantify the key factors influencing procurement outcomes; and (3) validate the effectiveness of the proposed model using actual business data from a railroad equipment company. By leveraging enterprise production and operation status alongside inventory, historical purchasing, and supplier datasets, our approach generates optimal purchase decisions, produces actionable procurement plans and management recommendations, and demonstrably enhances both planning accuracy and decision objectivity in a live business environment.

2 LITERATURE REVIEW

In recent years, traditional purchase decision-making, relying heavily on decision-makers' experience and subjective preferences, has proven inadequate for the increasing complexity of supply chains. Consequently, a growing body of research has focused on developing intelligent, data-driven models that leverage advanced algorithms and information technologies. For example, Xu et al. investigated a dual-sourcing inventory problem faced by loss-averse retailers, formulating a balanced objective that maximizes both expected profit and minimizes expected loss to derive optimal ordering policies [6]. Lei addressed the multi-objective challenge of military equipment procurement by constructing a fuzzy assignment model under multiple constraints and solving it via a genetic-algorithm-based approach [7]. Additionally, recent years have seen a rapid growth in applying Markov decision processes and reinforcement learning to supply-chain and procurement problems, offering data-driven strategies for inventory control, ordering and supplier choice under uncertainty. Comprehensive surveys highlight the rising use of RL methods in logistics and SCM and document the progress from MDP-based inventory control to deep-RL frameworks for end-to-end supply-chain decisions [8]. In parallel, several application studies demonstrate how contextual or deep RL can be adapted to procurement-like problems, for example, contextual MDP approaches for supply-chain control and deep-RL ordering mechanisms for multi-echelon systems [9]. There are also studies that extend the decision-making objectives beyond immediate cost and delivery metrics [10]. These have laid a good foundation for the diversified development of procurement decisions.

Within the railway sector-where procurement decisions must align with rigorous safety, reliability and cost requirements, scholars have begun to explore digital and intelligent procurement methods tailored to industry specifics. Zuzana et al. proposed that railway enterprises should make full use of the Internet of Things technology to achieve intelligent material management by collecting, monitoring, and evaluating data in real time [11]. Shih et al. explored the expected cost minimization of railway purchase plans through deterministic and stochastic optimization models of the scenario generation process by conducting sensitivity analysis on budgets, storage capacities, and expiration periods [12]. The concept of "smart procurement" has since emerged, emphasizing not only increased informatization of procurement processes but also standardized management practices supported by robust governance frameworks. This approach leverages IoT, big data analytics, and artificial intelligence to enhance decision precision, intelligence, and sustainability [13]. For instance, Oihab et al. explored how AI can reconfigure the purchase function, redefine the roles of purchase personnel, supplier relationship management policies, and interdepartmental collaboration through several case studies [14]. Liu et al. further demonstrated, developed a multi-period dynamic procurement model that combines variational mode decomposition with long short-term memory networks to forecast market prices, incorporating production scheduling and market volatility [15].

At present, various industries have begun to explore appropriate intelligent procurement models and are constantly seeking development [16-18]. However, based on the industry and business characteristics of the research questions, the digital and intelligent transformation methods of procurement work vary significantly. Compared with other industries, in-depth analyses of specific industries such as railway enterprises have been relatively rare. Although some studies have proposed application models and system architectures for the above issues, there are still significant limitations in terms of performance and practicality. The level of intelligence in the railway material field is relatively low, indicating a substantial opportunity and need for more targeted, high-fidelity research in this critical sector.

3 RESEARCH METHODOLOGY

This section mainly focuses on the construction of the Markov Decision Model, including the analysis of the purchasing process for railway wagons components, the analysis of relevant influencing factors, as well as the design and assumptions of the model. The purpose is to clearly explain the principles and foundation of the model construction and to provide a basis for the subsequent instance verification of the model.

3.1 Analysis of Business Process and Influencing Factors

3.1.1 Overview of Railway Wagons Components Procurement Workflow

At present, the railway enterprises of China have two kinds of processes: centralized purchase and emergency purchase. The centralized purchase is the main purchase method for most of the required materials. Centralized purchase management of materials means that the unified purchase department of an enterprise satisfies equipment materials and other production materials in the production process by obtaining marketing resources [19]. The procurement process for railway-wagon components organization is a recurring, multi-actor workflow that spans demand identification, sourcing, ordering, receipt inspection and inventory management. The specific business process is specifically described in Fig. 1 below.

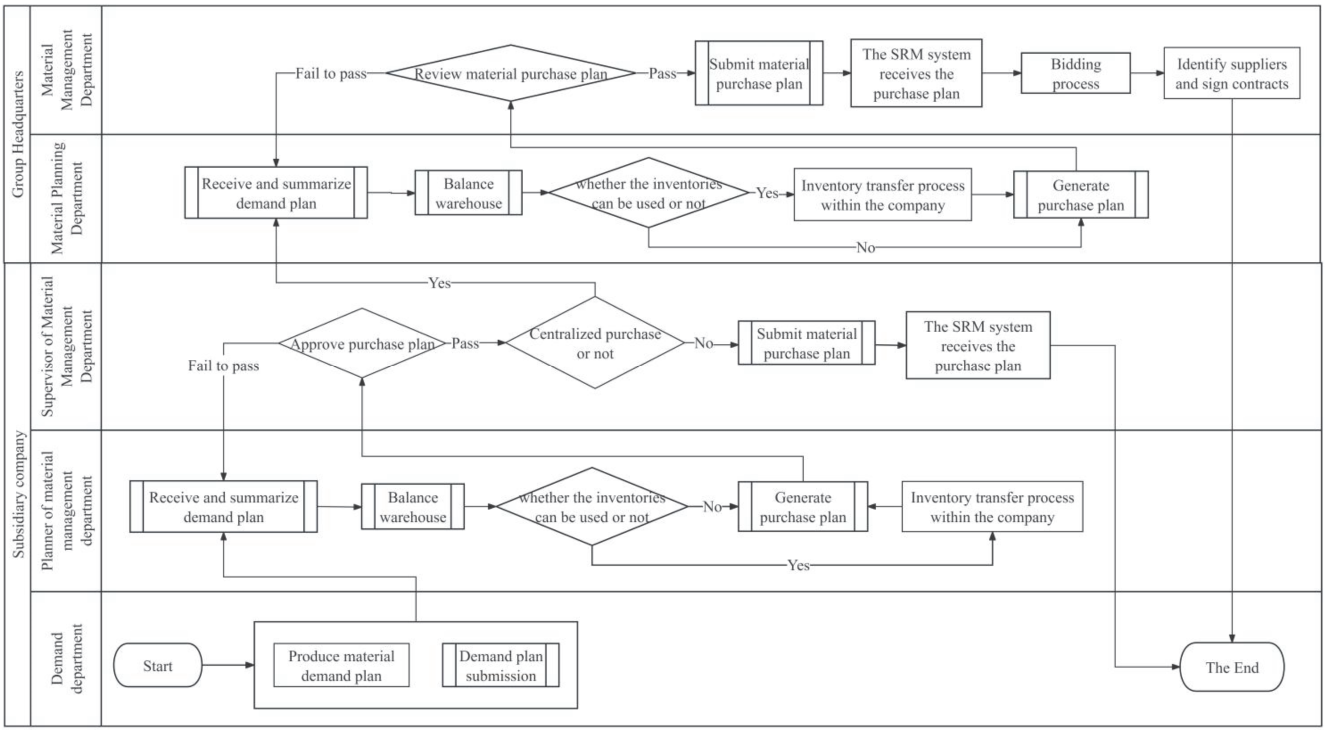

Figure 1 Purchase business process flow chart

(1) The purchase business process starts with the purchase requirements of the subsidiary's demand department. The subsidiary prepares the material demand plan based on the actual needs of the team and department

and reports it to the planner of the material management department.

(2) After receiving the demand plan, the planner of the material management department balances the inventories among the warehouses, decides whether the inventories can be used or not, generates the purchase plan according to the specific demand, and reports it to the supervisor.

(3) After the material management department is in charge of approving the purchase plan, if there is no emergency, it will be reported to the headquarters material planning department for centralized purchase. If the emergency purchase process is followed, it must be submitted directly to the purchase plan and entered into the Supplier Relationship Management (SRM).

(4) The subsidiary shall report the demand plan to the headquarters material planning department. Then the group shall uniformly balance the inventories and generate the total material purchase plan and report it to the material management department.

(5) After the headquarters material management department reviewed the approval, submitting the material purchase plan, the SRM system accepts the purchase plan and makes subsequent purchase arrangements.

(6) According to the purchase plan, the bidding process will be carried out. The bid evaluation team will evaluate and select suppliers, sign contracts with the suppliers, and include relevant data into the system.

3.1.2 Limitations of Current Practice

The current procurement process follows a broadly standardized workflow for requisition, approval and ordering. However, its core links lack adequate informatization and automation, producing several operational weaknesses:

(1) Siloed information and delayed reconciliation. Demand forecasts, inventory records and supplier-performance metrics are maintained in disparate systems. This fragmentation reduces the timeliness and accuracy of inputs available to purchasing decisions.

(2) Subjective, heuristic-driven decisions. Procurement staff commonly rely on experience-based rules of thumb (for example, holding excessive buffer stock "to be safe"), which introduce a conservative bias that systematically inflates holding costs and conceals opportunities to reduce inventory without materially increasing stockout risk.

(3) Under-utilization of historical dynamics. Although historical transaction records exist, they are rarely exploited to be used. Decisions therefore under-exploit information that could reduce uncertainty.

With the in-depth implementation of the condition-based maintenance, enterprises' demand for materials is becoming more and more complex and more flexible. These limitations create a clear opportunity for a structured decision model that (i) unifies state information, (ii) explicitly represents stochastic transitions in supplier outcomes, and (iii) operationalizes trade-offs between cost, delivery and quality.

3.1.3 Analysis of Influencing Factors of Purchase Decision-Making

Purchase decision-making is a complex and crucial process, involving comprehensive considerations of multiple factors, including but not limited to the links between purchasing quantity decision-making and supplier decision-making. The traditional decision-making method relying on the knowledge and preferences of decision-makers can no longer meet the needs of purchasing work. Using intelligent algorithms and information technology to construct decision-making models and solve problems is the focus of the current trend of related research [20]. In order to solve the problem of purchasing railway wagon components and formulate a scientific and reasonable intelligent solution, we need to conduct an in-depth analysis of the related influencing factors that affect the formulation of purchasing quantity plans and supplier decisions, so as to provide a basis for the subsequent design of the smart purchasing model.

(1) Market Factors

From the market perspective, market demand is one of the important factors affecting the formulation of the components purchasing quantity plan. The railway wagons in this study are all coal transport gondola cars. Local coal prices and production quantities, as market factors, will affect the formulation of the purchasing plan. In addition, in the purchasing process, the corresponding market price level of the components is also an important factor affecting the final purchasing decision. The purchasing plan will be reasonably adjusted based on demand and price level during the generation of the purchasing decision.

(2) Inventory Factors

Inventory level is one of the important factors determining the purchasing quantity of components. An excessively high inventory level will increase the enterprise’s warehousing costs and cause waste of funds, while a too low inventory level may lead to delays in implementing maintenance. The enterprise needs to comprehensively consider the actual consumption demand and inventory status to formulate an appropriate purchasing quantity of components. Therefore, the model in this paper will output the purchasing decision based on the combined status of inventory and demand.

(3) Supplier Factors

The reasonable selection of suppliers is very important for the formulation of purchasing decisions, but the selection of suppliers is complex. The enterprise needs to conduct a comprehensive evaluation of suppliers to ensure the smooth implementation of the purchasing plan and avoid the risk of interruption or delay of the maintenance plan due to unstable supply [21, 22]. Combined with the actual data situation, this study will mainly consider the factors of suppliers in terms of supply price, supply cycle, and product quality to examine the comprehensive ability of suppliers and make decisions.

3.2 Purchasing Decision Model Based on Markov Decision Process

3.2.1 Model Design

The existing procurement process is unable to leverage historical data in real time, resulting in procurement strategies that cannot dynamically adapt to changes in inventory levels and demand status. Moreover, supplier selection based on manual evaluation is often driven by individual experience and preference, overlooking a systematic trade-off among quality, delivery performance, and cost. To address these limitations, this study adopts a reinforcement learning framework in which the procurement decision is formulated as a dynamic, sequential decision-making problem and modeled using a Markov Decision Process (MDP). An MDP characterizes the interaction between the agent (the purchase decision-maker) and the environment (market conditions, inventory, and demand) via three components, state, action, and reward, and through iterative updating of the value function and policy, can discover long-term optimal strategies even in settings with sparse data and large state spaces. By incorporating multi-source data, the MDP can achieve both real-time responsiveness and multi-objective optimization. The overall MDP workflow is illustrated in Fig. 2. Compared with traditional methods, reinforcement learning offers intrinsic advantages in high-uncertainty, data-scarce environments by exploiting a trial-and-error feedback mechanism to uncover data-driven optimal policies [23, 24].

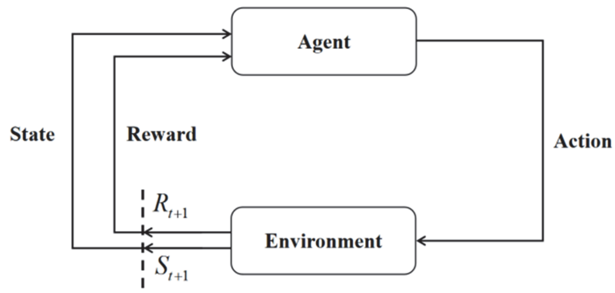

Figure 2 Markov Decision Process workflow

Grounded in the actual business context and process analysis, the procurement decision problem is decomposed into two interrelated subproblems: (i) a planned-purchase-quantity decision that determines the order quantity class for the coming period; and (ii) a supplier-selection decision that chooses the supplier type for the planned order.

By solving each stage's MDP via value iteration and then integrating the resulting policies, the model yields a combined purchase decision, specifying both the optimal order quantity for each component (consumption demand and inventory) and the preferred supplier type based on that quantity and market conditions. The detailed research route is depicted in Fig. 3.

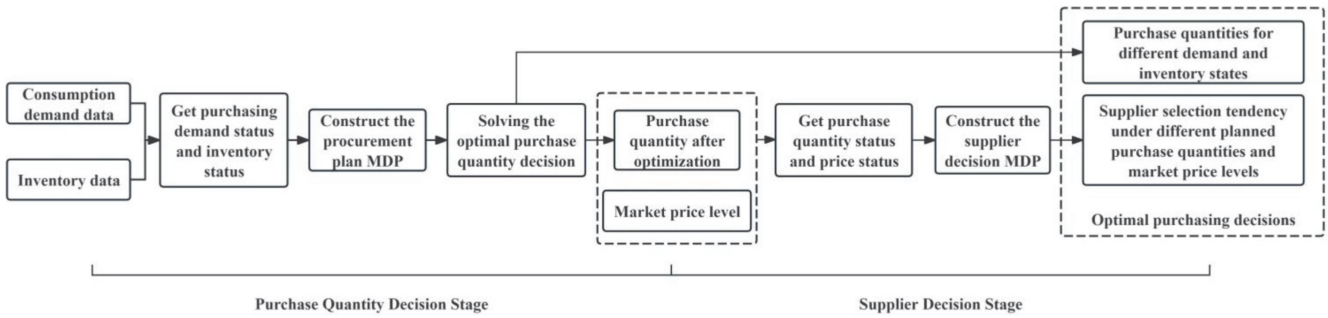

Figure 3 Purchase decision-making research route

3.2.2 Construction of a Purchasing Decision Model Based on MDP

In this study, let the length of the decision cycle of the whole purchasing plan development and implementation process be T, and the time node at which a specific decision needs to be t, $t \in \{1, 2, ..., T\}$.

(1) Stage 1: Purchase quantity decision

Demand data and current inventory are treated as exogenous, pre-processed inputs fed into Stage 1. For clarity, the essential Stage 1 design is summarized here:

a. State: $s_t = (d_t, i_t)$, where $d_t \in \{0, 1, 2, 3\}$ is discretized demand level (low/medium/high/peak) and $i_t \in \{0, 1, 2\}$ denotes inventory level (low/medium/high). The discretization thresholds follow the quartiles of demand and inventory of the components in the past 12 months. And it is worth noting that the medium inventory state corresponds to Q_1 to Q_3 of the total inventory.

b. Action: $a_t \in \{1, 2, 3, 4\}$ corresponding to planned purchase quantity {no purchase, small quantity, medium quantity, large quantity}.

c. Reward: The immediate reward balances purchase cost, expected holding cost, and shortage penalty:

$$r_t\left(s_t, a_t\right) = -\text{purchasecost}\left(a_t\right) - \\ -\text{holdingcost}\left(s_t, a_t\right) - \text{shortagecost}\left(s_t, a_t\right) \quad (1)$$

All cost components are converted to monetary units.

d. Transition: $P\left(s_{t+1} | s_t, a_t\right)$ is the probability of transferring from state s_t to state s_{t+1} when decision takes action a_t.

e. Objective: Maximize expected discounted cumulative reward over the horizon; Stage 1 is solved by value iteration to obtain the planned-quantity decision $\pi^{(1)}(s)$.

(2) Stage 2: Supplier decision

The supplier-stage objective is to maximize expected cumulative reward from the point of supplier selection onward, under the assumption that planned quantity q is fixed for the current decision epoch. Operationally, Stage 1 produces q and passes it to Stage 2; Stage 2 returns the chosen supplier b^* and, in deployment, the realized outcomes feed back into the enterprise records used for periodic re-estimation of supplier profiles. The specific structure of Stage 2 is as follows:

a. State: $\varphi = (q, p)$, where $q \in \{1, 2, 3\}$ is the planned-purchase-quantity level (small quantity, medium quantity, large quantity) decided in Stage 1 and $p \in \{1, 2, 3\}$ represents current market price level (price decline, price stability, price increase). The market price state is divided according to the change and 5% comparison between the average price of each candidate supplier and the last purchase transaction price of the component.

b. Action: $b_t \in \{1, 2, 3\}$ corresponding to selecting a supplier type that prioritizes the named attribute {price advantage, delivery advantage, quality advantage}. On the basis of data analysis and combined with the opinions of professionals inside the enterprise, the definition of each type of supplier is given.

c. Reward: The supplier-stage reward integrates price, expected lead time and arrival-quality conformity into a scalar via AHP-derived weights. Concretely, domain experts provided pairwise comparisons using the standard Saaty 1-9 scale, and individual judgment matrices were aggregated by weighted averaging to form a composite pairwise matrix. Let price(b), leadtime(b) and quality_score(b) denote the expected metric values for supplier type b. The priority vector $w = \{w_{\text{price}}, w_{\text{delivery}}, w_{\text{quality}}\}$ is obtained from the principal eigenvector of the aggregated matrix and normalized to sum to one. For each aggregated matrix we compute $\lambda_{\max}$, the consistency index (CI) and the consistency ratio (CR = CI/RI) and accept the matrix when CR ≤ 0.10. The reward is:

$$r_t(\varphi_t, b_t) = -w_{\text{price}} \cdot \text{price}(b_t) - w_{\text{delivery}} \cdot \text{leadtime}(b_t) + \\ + w_{\text{quality}} \cdot \text{quality\_score}(b_t) \tag{2}$$

d. Transition: Supplier-stage transitions model stochastic delivery and quality outcomes conditional on the chosen supplier type.

e. Solution & integration: Each stage is solved by value iteration. The integrated optimal decision is obtained by applying Stage 1 to select q^* and then executing Stage 2 decision $\pi^{(2)}(\varphi)$ conditional on that q^*.

Based on this, the MDP is used to learn from a sample of outcomes observed in the environment to derive the optimal strategy, and the goal of the decision is to maximize the expected return. Finally, the model will obtain the optimal value function and the optimal decision.

4 RESULTS AND DISCUSSION

This section mainly focuses on the practical application experiments of the model, covering the process from data preparation to result analysis. It continues from the previous section's discussion on the design and construction of the model and presents the outcomes of this study.

4.1 Data Preparation and Simulation Realization

This study is based on the shift consumption data, material management data, wagon track data, supplier management data, and purchase contract details in the HCCBM system, HMIS system, ERP system, and SRM system of a group of railway equipment companies in China. According to the actual needs of the business process, with the help of the MySQL database and Navicat for MySQL database management and development tools, the data integration work is completed, and the data standardization is realized by stipulating a unified data format, standard, and structure. The feature structure of the extracted instance data related to the construction and implementation of the smart purchase model is shown in Tab. 1.

Table 1 Feature extraction

Data usage	Field Name	Description	Data Type
Purchasing decision supplier information	SUPPLIER_NO	Supplier code	INT
	SUPPLIER_NAME	Supplier name	VARCHAR(50)
	COMPONENT_NO	Component code	INT
	PRICE_AVG	Historical average purchase price	FLOAT(20, 2)
	DELIVERY_DAYS	Average delivery days	FLOAT(20, 2)
	ON_TIME_RATE	Timely arrival rate	FLOAT(20, 2)
	PASS_RATE	Pass rate	FLOAT(20 ,2)
	QUANTITY	Purchase quantity	FLOAT(20, 2)
	UNIT_PRICE	Purchase unit price	FLOAT(20, 2)
	AMOUNT	Purchase amount	FLOAT(50, 2)
	TYPE	Supplier type	INT

Due to the limited period of historical data, the decision period is set as half a year, and January 2023 is selected as the initial time. The demand for the corresponding components at the decision time is from enterprise's intelligent prediction platform, and the inventory level of the period is obtained by using the difference between the arrival quantity and the monthly consumption of the components in the corresponding period of historical purchase. Using the historical data before January 2023, based on a total of 24 times of components consumption and purchase related historical data from January 2021 to December 2022, taking the traction bar component as an application example, the state transition matrix corresponding to each action taken in the two stages of purchase quantity decision and supplier selection decision is obtained.

Based on various conditions, the two are compared between the two kinds of action, on the basis of the AHP judgment matrix, calculated reward function of the Markov decision process. Taking the state of $q = 3$ in the supplier decision stage as an example, the weights extracted by the AHP method, namely the reward value, are shown in Tab. 2.

Table 2 Example of reward value extraction

$q = 3$	Reward Value
$p = 1, b = 1$	0.0223
$p = 1, b = 2$	0.0789
$p = 1, b = 3$	0.1944
$p = 2, b = 1$	0.0225
$p = 2, b = 2$	0.1617
$p = 2, b = 3$	0.1454
$p = 3, b = 1$	0.0654
$p = 3, b = 2$	0.1360
$p = 3, b = 3$	0.1734

From the decision-maker preference results shown in the table above, when procurement volume is high and market prices are rising ($q = 3$, $p = 3$), the reward values for selecting price-advantage, delivery-advantage, and quality-advantage suppliers are 0.0654, 0.1360, and 0.1734, respectively. Under conditions of large purchase quantities, the impact of quality stability on the continuity of maintenance schedules is markedly greater than that of short-term cost savings or delivery speed, this finding corresponds closely with the real-world pain point that "quality issues in bulk component production can cause downstream maintenance to stall" thereby providing a data-driven basis for prioritizing criteria in supplier selection decisions.

4.2 Evaluation of Experimental Results

The purchasing decision optimization of purchasing quantity and supplier selection from January 2021 to June 2023 is carried out separately for each key component using Markov decision process value iteration method. Using Python to define the state of the Markov decision-making process for each component, set the discount factor $\gamma = 0.9$, solve the optimal decision-making for different states in the decision-making cycle, and finally the optimal decision under the definition of the same state is obtained by the comprehensive voting of the solution results of the purchase process for different components. In summary, the optimal purchasing decision under each state in two stages can be obtained, as shown in Fig. 4 and Fig. 5. For Fig. 4, the horizontal axis represents the demand status ($d = 0$: low demand; $d = 1$: medium-low demand; $d = 2$: medium-high demand; $d = 3$: high demand), the vertical axis represents the inventory status ($i = 0$: low inventory; $i = 1$: medium inventory; $i = 2$: high inventory), and the content in the cells represents the optimal purchase quantity decision. For Fig. 5, the horizontal axis represents the planned purchase quantity status ($q = 1$: small quantity; $q = 2$: medium quantity; $q = 3$: large quantity), the vertical axis represents the market price status ($p = 1$: price decline; $p = 2$: price stability; $p = 3$: price increase), and the content in the cells is the optimal supplier type.

In the optimal decision for purchasing quantity, when the inventory level is higher than average and the demand level is low, the reinforcement learning results show a no-purchase action strategy. In the case of high-demand state, the optimal decision of the smart purchasing model adjusts the purchasing quantity action to medium and large quantity purchases according to the inventory state, which increases the flexibility of making purchasing plans. In the case of the same inventory level state, the optimal purchasing decision also adjusts the purchasing quantity decision according to the shift in the demand state to increase the accuracy of the purchasing plan.

i=2	a_1	a_2	a_2	a_3
i=1	a_1	a_2	a_2	a_3
i=0	a_2	a_3	a_4	a_4
	d=0	d=1	d=2	d=3

a_1: no purchase
a_2: small quantity
a_3: medium quantity
a_4: large quantity

Figure 4 Optimal purchase quantity decision

p=3	b_1	b_1	b_1/b_3
p=2	b_2	b_2/b_3	b_3
p=1	b_2	b_3	b_3
	q=1	q=2	q=3

b_1: price advantage
b_2: delivery advantage
b_3: quality advantage

Figure 5 Optimal supplier decision

In the supplier's optimal decision, in the state of small purchase quantity, the propensity to purchase price and delivery will be higher, and the delivery time will have higher requirements, but the emphasis on product quality is relatively low. However, as the purchase quantity becomes larger, the purchase decision will be more inclined to high-quality products, not to pursue the speed of delivery, and the sensitivity to price is relatively lower. In the case of a large purchase, if there are more quality problems, the subsequent maintenance plan may be affected, and the processing efficiency is low. Therefore, in the decision-making of the intelligent purchase model, under the condition of large purchase quantity, no matter how the price trend changes, the supplier with superior quality will have a clear decision-making tendency.

In order to present and analyze the effectiveness and benefits of the model more clearly, we compared the key indicators of the original decision and the optimal decision of the traction bar, which is one of the key components, over its entire historical cycle. Tab. 3 summarizes the primary comparative outcomes for the traction bar. The gains arise from (i) more accurate order sizing that reduces unnecessary holding costs while avoiding stockouts, and (ii) an adaptive supplier-selection rule that trades off price, delivery and quality according to order size and market signals.

The results confirm that the intelligent purchasing decision model based on MDP can optimize the decision-making process and achieve better decision-making outcomes. The model can bring benefits in terms of cost savings, efficiency improvement, and quality management. This validates the effectiveness and feasibility of the intelligent purchasing model for key components of railway wagons.

Table 3 Decision effectiveness comparison

Indicator	Manual (recorded)	Intelligent purchasing decision model	Absolute change	Relative change
Planned purchase quantity (PC)	5100.00	4610.00	−490.00	−9.6%
Average inventory (PC)	176.89	163.53	−13.36	−7.6%
Total procurement cost (10k CNY)	2176.68	1942.19	−234.49	−10.8%
Average delivery time / days	35.16	30.31	−4.85	−13.8%
Arrival conformity rate / %	85.79	88.67	+2.88 pp	+2.9 pp

5 CONCLUSION

This paper presents a two-stage Markov Decision Process (MDP) framework tailored to railway-wagon components procurement that jointly optimizes planned purchase quantities and supplier selection. The main methodological contribution is an operationally interpretable decomposition: a quantity-stage MDP that outputs planned-quantity classes and a supplier-stage MDP that conditions on those classes and trades off price, lead time, and arrival quality through AHP-elicited weights. The approach preserves interpretability and deployability by keeping state and action representations compact and by estimating transition behavior empirically from transaction records. We validate the framework by replaying the intelligent model on historical records from a large railway equipment enterprise and comparing aggregated outcomes to recorded manual decisions. The results demonstrate that the proposed two-stage design can deliver measurable operational benefits while remaining transparent to procurement practitioners.

While these results are encouraging, several avenues remain for future research. First, the definition in this paper is not unique, and later on the basis of richer business data, we can consider further improving the problem abstraction method and definition of the model. Second, extending supplier evaluation criteria such as risk assessment and model evaluation metrics such as sustainability to facilitate the generation of more comprehensive decisions. Third, integrating additional influencing factors, such as macroeconomic fluctuations or advanced forecasting techniques, would further enhance model robustness. Finally, extending this framework to other freight scenarios (e.g., spare parts logistics, intermodal containers or bulk commodities) offers promising opportunities to generalize our approach across the railway supply chain.

6 REFERENCES

[1] Kalantari, M., Taghaddos, H., & Heydari, M. (2024). BIM framework for efficient material procurement planning. *Automation in Construction*, *168*, 105803. https://doi.org/10.1016/j.autcon.2024.105803

[2] Liu, Q. & Wang, X. (2015). Material Requirement Planning Application Example Analysis-Shuohuang Railway Company Case Study. *LISS 2014: Proceedings of 4th International Conference on Logistics, Informatics and Service Science*, 253-258. https://doi.org/10.1007/978-3-662-43871-8_39

[3] Li, Q., Xu, Y. et al. (2022). Research on intelligent operation and maintenance technology of railway freight wagons. *Intelligent Rail Transit*, *59*(05), 38-41.

[4] Ingemarsdotter, E., Kambanou, M. L., Jamsin, E., Sakao, T., & Balkenende, R. (2021). Challenges and solutions in condition-based maintenance implementation-A multiple case study. *Journal of Cleaner Production*, *296*, 126420. https://doi.org/10.1016/j.jclepro.2021.126420

[5] Sarp, S., Kuzlu, M., Jovanovic, V., Polat, Z., & Guler, O. (2024). Digitalization of railway transportation through AI-powered services: digital twin trains. *European Transport Research Review*, *16*(1), 58. https://doi.org/10.1186/s12544-024-00679-5

[6] Xinsheng, X., Ping, Ji., & Felix T. S. C. (2023). Retailers' optimal ordering policies for a dual-sourcing procurement. *Industrial Management & Data Systems*, *123*(3), 1052-1072. https://doi.org/10.1108/IMDS-07-2022-0458

[7] Lei, S. Y. & Liu, J. X. (2020). Equipment Procurement Optimization Based on Genetic Algorithm. *China Management Science*, *28*(10), 194-200.

[8] Yan, Y., Chow, A. H., Ho, C. P., Kuo, Y. H., Wu, Q., & Ying, C. (2022). Reinforcement learning for logistics and supply chain management: Methodologies, state of the art, and future opportunities. *Transportation Research Part E: Logistics and Transportation Review*, *162*, 102712. https://doi.org/10.1016/j.tre.2022.102712

[9] Batsis, A. & Samothrakis, S. (2024). Contextual reinforcement learning for supply chain management. *Expert Systems with Applications*, *249*(Part A), 123541. https://doi.org/10.1016/j.eswa.2024.123541

[10] Rosyidi, C. N. & Pratama, D. M.(2024). An optimization model of supplier selection and order allocation with transportation mode alternatives under carbon cap and trade policy. *Cogent Engineering*, *11*(1). https://doi.org/10.1080/23311916.2024.2321757

[11] Gerhátová, Z., Zitrický, V., & Klapita, V. (2021). Industry 4.0 implementation options in railway transport. *Transportation Research Procedia*, *53*, 23-30. https://doi.org/10.1016/j.trpro.2021.02.003

[12] Shih, H. C., Yeh, C. H., & Lai, Y. C. (2022). Optimization of Multi-Period Rail Procurement Plan. *Transportation Research Record*, *2676*(4), 324-333. https://doi.org/10.1177/03611981211058676

[13] Qi, D., Lu, Y. et al. (2020). Research on the wisdom procurement model of construction enterprises under "Internet +". *Construction Economy*, *41*(03), 38-41.

[14] Allal-Chérif, O., Simón-Moya, V., & Ballester, A. C. C. (2021). Intelligent purchasing: How artificial intelligence can redefine the purchasing function. *Journal of Business Research*, *124*, 69-76. https://doi.org/10.1016/j.jbusres.2020.11.050

[15] Liu, Y., Yang, C., Huang, K., Gui, W., & Hu, S. (2022). A systematic procurement supply chain optimization technique based on industrial internet of things and application. *IEEE Internet of Things Journal*, *10*(8), 7272-7292. https://doi.org/10.1109/JIOT.2022.3228736

[16] Mircea, M., Stoica, M., & Ghilic-Micu, B. (2022). Analysis of the impact of blockchain and internet of things (BIoT) on public procurement. *IEEE Access*, *10*, 63353-63374. https://doi.org/10.1109/ACCESS.2022.3182656

[17] Häselbarth, S., Winkels, O., & Strunz, K. (2023). Blockchain-based market procurement of reactive power. *IEEE Access*, *11*, 36106-36119. https://doi.org/10.1109/ACCESS.2023.3263669

[18] Li, Z., Chen, D., Yin, L., Yang, N., & Liu, W. (2023). Research and Application of intelligent purchasing and transportation model for coal-fired power plant. *Journal of Physics: Conference Series*, *2422*(1), 012004. https://doi.org/10.1088/1742-6596/2422/1/012004

[19] Ding, F. X., Liu, S. F., & Li, X. W. (2023). An innovative framework for sustainable and centralized material procurement management based on a full-domain set theory. *Advances in Production Engineering & Management*, *18*(1). https://doi.org/10.14743/apem2023.1.453

[20] Xu, X., Ji, P., & Chan, F. T. (2023). Retailers' optimal ordering policies for a dual-sourcing procurement. *Industrial Management & Data Systems*, *123*(3), 1052-1072. https://doi.org/10.1108/IMDS-07-2022-0458

[21] Wang, G. Q., Fan, X. W., Wang, X. N., Chen, H. B., Chen, W., & Xiong, X. J. (2014). A Theory of Hierarchy-Gray for Purchase Decision Support Systems. *Applied Mechanics and Materials*, *624*, 681-686. https://doi.org/10.4028/www.scientific.net/AMM.624.681

[22] Chen, S. C., Lai, M. C., Chu, C. H., Chen, H. M., & Nafei, A. (2024). Enhanced Supplier Evaluation in Digital Transformation: A BWM-Neutrosophic TOPSIS Approach for Decision-Making Under Uncertainty. *Studies in*

Informatics and Control, *33*(4).
https://doi.org/10.24846/v33i4y202409

[23] Ramesh, S. S., Sessa, P. G., Hu, Y., Krause, A., & Bogunovic, I. (2024). Distributionally robust model-based reinforcement learning with large state spaces. *International Conference on Artificial Intelligence and Statistics*, 100-108).

[24] Nguyen, T. T., Nguyen, C. M., Huynh-The, T., Pham, Q.-V., Nguyen, Q. V. H., Razzak, I., & Reddi, V. J. (2023). Solving complex sequential decision-making problems by deep reinforcement learning with heuristic rules. *International Conference on Computational Science*, 298-305).
https://doi.org/10.1007/978-3-031-36021-3_30

Contact information:

Shuo ZHANG
School of Economics and Management,
Beijing Jiaotong University, Haidian, 100044, China
E-mail: betsy_zs@163.com

Chen LI
Energy Railway Equipment CO., LTD, China
E-mail: 11590077@ceic.com

Hao HUA
Energy Railway Equipment CO., LTD, China
E-mail: 11655733@ceic.com

Ping GAO
Energy Railway Equipment CO., LTD, China
E-mail: 11590255@ceic.com

Wenzhuo WANG
China Merchants Bank Co., Limited Jinan Branch, 250010, China
E-mail: wwzada@163.com

Qiuli QIN, Professor
(Corresponding author)
School of Economics and Management,
Beijing Jiaotong University, Haidian, 100044, China
E-mail: qlqin@bjtu.edu.cn