LEARNING-GUIDED SEARCH FOR RELOCATION PROBLEM IN MULTI-DEEP AVS/RS SYSTEMS

Jakob Marolt[1], Valerija Kotnik[1], Mateja Lešnik[1]
[1] Fakulteta za logistiko, Univerza v Mariboru, Mariborska ulica 7, Slovenija
e-mail: jakob.marolt@um.si mateja.lesnik@um.si valerija.kotnik@um.si

Izlaganje sa znanstvenog skupa
https://www.doi.org/10.58952/zr20261501149
UDK / UDC 658.78:004.85:519.857

Abstract
Efficient relocation handling is essential in multi-deep automated vehicle storage and retrieval systems (AVS/RS), where requested stock keeping units are frequently blocked by other items stored in front of them. As warehouse depth and occupancy increase, the resulting relocation problem becomes highly combinatorial and difficult to solve with exact optimization methods. This paper proposes a learning-guided search framework for the restricted relocation problem in multi-deep AVS/RS systems. The approach combines Reverse Renumbering Relocation Construction (R3C), supervised value learning, and beam search. First, R3C generates synthetic warehouse states with known relocation counts by construction, enabling large-scale supervised learning without repeated exact optimization. Second, a DeepSets-style ValueNet is trained to estimate the remaining number of relocations from a given system state. Third, the learned value model is embedded into a beam search procedure that prioritizes promising relocation sequences. The experimental analysis focuses on a medium-sized 5x40 warehouse and shows that enlarging the synthetic training dataset from 45,000 to 135,000 instances improves validation performance from RMSE 0.712 and MAE 0.535 to RMSE 0.589 and MAE 0.475. This improvement also reduces the average search gap to the R3C reference from 14.78 to 7.62 relocations under the same beam-search setting. The results further show that the proposed framework performs very well at lower occupancy levels, while dense warehouse states remain significantly more challenging. The study demonstrates that supervised learning can provide effective search guidance for relocation decision-making in complex AVS/RS environments.

Keywords: *AVS/RS, relocation problem, supervised learning, beam search, warehouse optimization*
JEL classification: C45, C61, C63, M11

INTRODUCTION

Automated vehicle storage and retrieval systems (AVS/RS) are widely used in high-density warehousing because they improve space utilization, operational consistency, and handling efficiency. In multi-deep layouts, however, the requested stock keeping unit (SKU) is often blocked by other units stored in front of it. These blocking units must be relocated before retrieval can continue, creating additional time and operating effort. Minimizing such non-productive relocations is therefore an important decision problem in dense automated storage systems.

This relocation task is closely related to the block relocation problem and the container relocation problem. In all of these settings, the goal is to determine a sequence of moves that minimizes the number of relocations required to access the next requested item. The difficulty of the problem grows rapidly with system size, storage depth, and occupancy, because each relocation affects future accessibility and generates a large combinatorial search space.

Exact methods provide high-quality benchmarks, but their practical use becomes limited as instances become larger and denser. This creates strong motivation for approaches that preserve good solution quality while reducing computational burden. The framework proposed in this paper follows this idea through three steps. First, a constructive generator creates synthetic warehouse states with known relocation counts. Second, a neural value model is trained on these labels. Third, the learned value estimates are used inside a beam search procedure to guide the exploration of relocation sequences.

The contribution of the paper is threefold. First, it introduces a scalable synthetic data-generation procedure for relocation states in multi-deep AVS/RS. Second, it develops a supervised value-prediction model that learns the relocation structure of these states. Third, it shows that better value prediction translates into better search performance, especially when the model is trained on a richer synthetic dataset. The empirical analysis is centered on the medium-sized 5x40 warehouse, where both learning and search behavior can be examined clearly.

1. RELATED WORK

Relocation decisions in dense storage systems have been studied extensively in the block relocation problem (BRP) and container relocation problem (CRP) literature. Exact methods form a major part of this research stream. Tang, Zhao, and Liu (2015) presented an exact approach for the blocks relocation problem, while Tanaka and Takii (2018) developed an exact algorithm for the unrestricted block relocation problem. Bacci, Mattia, and Ventura (2020) proposed a branch-and-cut algorithm for the restricted problem, and Tanaka et al. (2022) introduced a new exact approach based on improved integer programming formulations. Jin and Tanaka (2023) later strengthened exact search further with new lower bounds and dominance rules for the unrestricted container relocation problem. These studies show both the maturity of exact relocation research and the difficulty of scaling exact methods to larger instances.

Heuristic and bounded-search procedures provide a more scalable alternative. Bacci, Mattia, and Ventura (2019) proposed a bounded beam search algorithm for the restricted block relocation problem and showed that limited-width search can remain effective when guided by suitable evaluation rules. Machine learning has become increasingly relevant in this context. Yu, de Koster, Guo, and Zhu (2020) proposed machine-learning-driven algorithms for the container relocation problem, while Ye, Ye, and Zheng (2023) showed that learning-based guidance can improve decision quality in blocks relocation. More recent studies continue this direction by combining reinforcement learning with relocation heuristics and search.

In the AVS/RS domain, previous work has focused more strongly on system design, relocation strategies, and operational performance than on supervised learning for relocation control.

Marolt, Kosanić, and Lerher (2021) evaluated relocation and storage assignment strategies in a multiple-deep AVS/RS with undetermined retrieval sequence, while Marolt, Šinko, and Lerher (2023) studied depth-first storage and relocation strategies in multiple-deep AVS/RS. Marolt, Rosi, and Lerher (2022) explored reinforcement learning for relocation assignment in a multiple-deep storage environment. The present paper builds on this line of work, but addresses a different gap: how to generate large labeled datasets for relocation learning and use supervised prediction as a practical search heuristic.

2. PROBLEM DESCRIPTION

This paper considers the relocation problem in a single bay of a multi-deep AVS/RS system. The bay is represented as a matrix with depth T and number of columns C. Each storage position contains either a positive integer or zero. Positive integers represent stock keeping units, while zero denotes an empty slot. Lower label values indicate earlier retrieval, so the next required SKU is always the smallest label currently present in the bay.

Because the system is accessible only from the front side of a column, the next required SKU may be blocked by one or more items located in front of it. In that case, the blocking items must be relocated to other non-full columns before retrieval can continue. Each relocation is a non-productive move, and the objective is to minimize the total number of relocations needed to empty the bay according to the given retrieval order.

The present study considers the restricted version of the relocation problem. Under this rule, only the SKU that directly blocks the currently required item may be moved. This assumption matches the classical restricted block relocation setting and reflects a practically meaningful operational policy. Although three warehouse geometries were studied in the broader experimental pipeline, the detailed analysis in this paper focuses on the medium 5x40 system. To analyze the effect of warehouse density, occupancies of 50%, 60%, 70%, 80%, and 90% are considered.

3. R3C SYNTHETIC INSTANCE GENERATION

A central challenge in applying supervised learning to the relocation problem is the need for large numbers of labeled training examples. In principle, labels can be obtained by solving each instance with an exact optimization procedure. In practice, this quickly becomes computationally prohibitive as warehouse size and occupancy increase. To address this issue, this paper uses a constructive synthetic data generation procedure called Reverse Renumbering Relocation Construction (R3C), which produces relocation instances together with known relocation counts.

The idea of R3C is to construct relocation difficulty explicitly rather than waiting for it to appear randomly. Starting from an empty warehouse, the procedure first samples a sequence of relocation cycles. Each cycle has size $k_i \geq 2$, and each such cycle contributes exactly $k_i - 1$ relocations to the final instance. By combining multiple cycles, the generator creates states with controlled relocation structure. The total relocation label assigned to the generated state is therefore $K_{R3C} = \sum_{i=1}^{m}(k_i - 1)$, where m denotes the number of embedded cycles and k_i denotes the size of cycle i.

After the cycle structure is selected, the warehouse is filled using reverse renumbering. Newly inserted high-priority SKUs are assigned small labels, while previously inserted labels are shifted upward. In this way, blocking configurations are embedded deliberately into the bay. The generated state is therefore not random in a purely combinatorial sense; instead, it is constructed so that the number of required relocations is known by design.In the final step, the remaining empty slots are filled with large-label background items. These background items increase the occupancy of the bay without changing the relocation count imposed by the constructed cycle structure.

This allows the generator to produce instances at different occupancy levels while preserving the known relocation label. In the present study, occupancy levels of 50%, 60%, 70%, 80%, and 90% are considered for benchmark analysis, while the core training datasets were generated around the main paper regimes of 50%, 70%, and 90%. The main advantage of R3C is that it provides a scalable source of labeled data for supervised learning. Instead of solving every instance optimally with an exact method, labels are obtained directly from the constructive process. This makes it possible to generate large datasets efficiently, while still controlling important structural properties such as warehouse size, occupancy, number of cycles, and cycle sizes. As a result, R3C serves as the foundation of the proposed learning-guided search framework.

4. LEARNING-GUIDED SEARCH FRAMEWORK

The proposed framework combines supervised value prediction with heuristic search. The main idea is to train a neural model that estimates the remaining number of relocations from a given warehouse state, and then use this estimate to guide the exploration of candidate relocation sequences. In this way, learning provides fast state evaluation, while search remains responsible for constructing feasible move sequences.

4.1. VALUENET FOR RELOCATION ESTIMATION

The supervised model used in this study is a DeepSets-style value network, referred to as ValueNet. The model receives a representation of the current warehouse state and predicts the remaining number of relocations needed to clear the bay under the prescribed retrieval order.
This architecture is suitable for the present problem for two reasons. First, the relocation burden depends strongly on the internal structure of each column, including occupied depth, blocking patterns, and remaining free capacity. Second, the model should focus on structural relationships rather than memorizing absolute column identities. A shared column encoder with invariant pooling captures this idea naturally. The training target is the R3C relocation label K_{R3C}. Thus, the model does not learn from expensive exact optimization outputs, but from the synthetic labels created by the constructive generator. The trained ValueNet can evaluate a state very quickly, making it suitable for repeated use inside a search procedure.

4.2. BEAM SEARCH GUIDED BY VALUENET

To convert fast value estimates into decision support, the trained ValueNet is embedded into a beam search procedure. At each decision point, the current warehouse state is expanded according to the restricted relocation rule: only the topmost SKU blocking the currently required item may be relocated, and it may be moved only to a non-full destination column. Each feasible relocation produces a new candidate state. Instead of exploring all reachable states, beam search keeps only a fixed number of the most promising candidates at each stage. In this paper, the beam width is denoted by k, and the main experiments use $k = 10$. For every newly generated candidate state, ValueNet predicts the remaining number of relocations. This prediction is combined with the relocations already performed, so that the search favors states with low estimated total relocation cost. The role of ValueNet is therefore heuristic rather than exact. It does not directly produce the relocation sequence. Rather, it scores intermediate states so that the search procedure can allocate its limited budget to the most promising branches. Compared with uninformed search, this substantially reduces wasted exploration. Compared with exact optimization, it remains computationally light enough for larger and denser warehouse configurations. The resulting framework is a learning-guided search method with two complementary components. R3C provides large-scale labeled data, ValueNet learns to approximate remaining relocation cost, and beam search uses those predictions to navigate the combinatorial search space.

The effectiveness of this combination is evaluated in the experimental section by comparing the achieved number of relocations against the known R3C relocation count.

5. EXPERIMENTAL SETUP

The empirical study evaluates both predictive quality and search quality. The main analysis focuses on the medium-sized 5x40 warehouse, which contains 200 storage positions and provides a suitable balance between structural complexity and interpretability.
Two training configurations are considered for the medium model. The original dataset contains 45,000 R3C-labeled instances. To test whether additional synthetic data improves downstream search performance, the dataset was later expanded by a factor of three, resulting in 135,000 labeled instances. In both cases, the learning target was the R3C relocation count K_{R3C}.
Model quality is assessed by validation RMSE and MAE. Search quality is assessed by the average number of relocations found by beam search, the average gap to the R3C reference, and the exact-match rate. The main benchmark consists of 100 freshly generated R3C instances for the medium model. Occupancy levels of 50%, 60%, 70%, 80%, and 90% are used to evaluate how warehouse density affects search performance.

6. RESULTS AND DISCUSSION

6.1. LEARNING PERFORMANCE OF THE MEDIUM MODEL

The first important result of the study is that the supervised model learned the synthetic relocation signal very effectively on the medium-sized warehouse. The original medium model, trained on 45,000 labeled instances, achieved a best validation RMSE of 0.712 and MAE of 0.535. After increasing the dataset size to 135,000 instances, the best validation performance improved to RMSE 0.589 and MAE 0.475. Thus, enlarging the synthetic dataset led to a measurable increase in predictive accuracy.

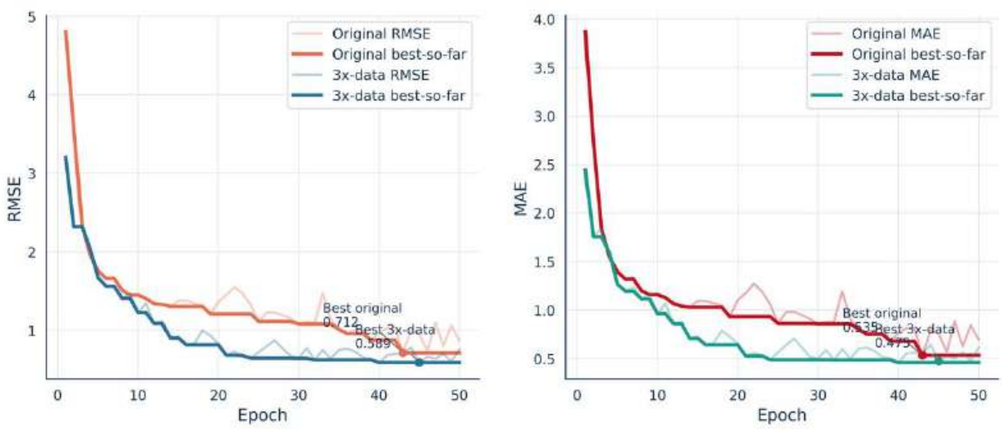

Figure 1: Validation performance of the medium ValueNet model during training

This result is important because the search procedure depends on fine distinctions between candidate states. It is not sufficient for the value model to be approximately correct on average; it must also rank alternative relocation choices in a way that supports good pruning decisions. The observed reduction in validation error therefore suggests that the larger dataset did not merely improve regression quality in a general sense, but also increased the usefulness of the model as a search heuristic.

At the same time, the experiments showed that simply prolonging training on the original medium dataset did not produce comparable gains. Additional epochs on the same dataset tended to give little benefit or even slightly worse validation behavior. By contrast, increasing the amount of synthetic training data improved both validation quality and downstream search performance. This indicates that, for the present framework, data diversity was more important than longer repeated training on the same examples.

6.2. BEAM SEARCH PERFORMANCE RELATIVE TO THE R3C REFERENCE

The main benchmark result is that the improved medium model substantially enhanced beam-search performance. Using the original medium ValueNet and beam width $k = 10$, the average gap between beam search and the R3C reference was 14.78 relocations. After retraining the model on the expanded 135,000-instance dataset, this average gap dropped to 7.62 relocations under the same beam-search setting. In other words, the average deviation from the R3C reference was reduced by almost half.

The average R3C relocation count in the final medium benchmark was 29.81, while the average ValueNet root prediction was 30.07. The corresponding average absolute prediction error at the root was approximately 0.50 relocations. The average beam-search solution required 37.43 relocations, resulting in the above average gap of 7.62. The exact-match rate, meaning the proportion of cases in which beam search exactly reproduced the R3C reference relocation count, was 0.29.

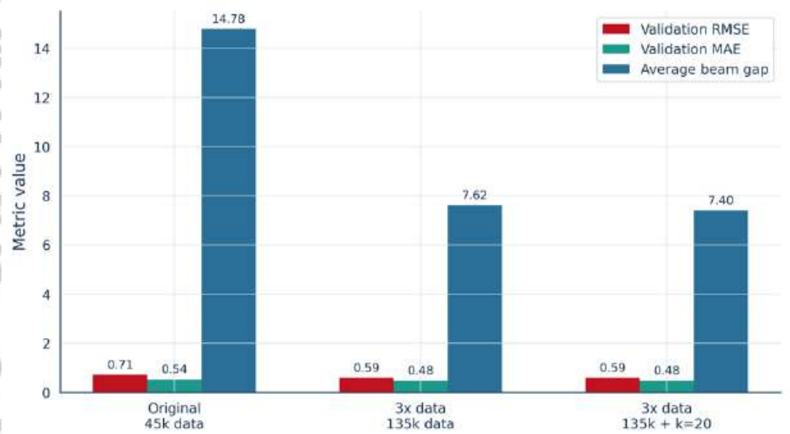

Figure 2: Effect of improved learning on beam-search performance

These results show that the value model predicts the overall relocation burden very accurately, but that search remains more difficult than pure prediction. This difference is important. A small prediction error at the root does not automatically imply perfect action selection during search. Beam search must repeatedly choose among many similar child states, and in dense instances even a small ranking error can cause the correct branch to be discarded early.

To test whether the search was still strongly limited by beam width, an additional benchmark with $k = 20$ was performed. The wider beam improved the average gap only slightly, from 7.62 to 7.40, while roughly doubling runtime and the number of expanded nodes. The exact-match rate remained unchanged at 0.29. This suggests that widening the beam alone does not fundamentally solve the problem. The larger performance gain came from improving the value model through additional training data rather than simply increasing search effort.

6.3. OCCUPANCY-SPECIFIC BEHAVIOR

The occupancy-based analysis provides one of the clearest findings of the paper. When the medium warehouse is only moderately filled, the proposed method performs very well. At 50% occupancy, the average R3C reference value was 25.44 relocations, while beam search achieved an average of 25.68 relocations, corresponding to an average gap of only 0.24. The exact-match rate in this regime was 0.85, meaning that in 85% of the tested cases the beam search recovered the R3C relocation count exactly.

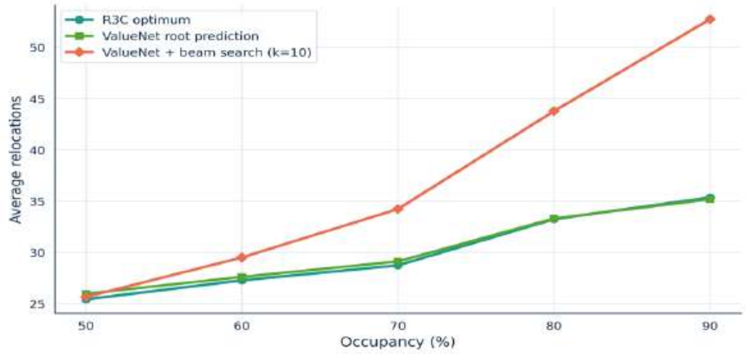

Figure 3: Average relocations by occupancy for the medium warehouse

At 60% occupancy, performance was still reasonably strong. The average R3C relocation count was 27.30, and beam search achieved an average of 29.52 relocations, corresponding to an average gap of 2.22. The exact-match rate decreased to 0.14, showing that the problem had already become noticeably harder, but the overall search quality remained acceptable.
At 70% occupancy, the difficulty increased substantially. The average beam-search result rose to 34.24 relocations compared with an average R3C reference of 28.76, giving an average gap of 5.48. At 80% occupancy, the average gap increased further to 10.52, and at 90% occupancy it reached 17.36. In both the 70%, 80%, and 90% regimes, the exact-match rate dropped to zero in the evaluated benchmark sets.
An interesting observation is that ValueNet prediction accuracy remained relatively stable across occupancy levels. For example, the average absolute root prediction error was 0.59 at 50%, 0.56 at 60%, 0.39 at 70%, 0.32 at 80%, and 0.51 at 90%. These values remain small compared with the corresponding search gaps. This means that the main deterioration at high occupancy is not caused by a complete breakdown of value prediction. Rather, it arises because dense bays create much more sensitive combinatorial search decisions. In such cases, even a value model with low average prediction error may still fail to preserve the correct branch inside a narrow beam.

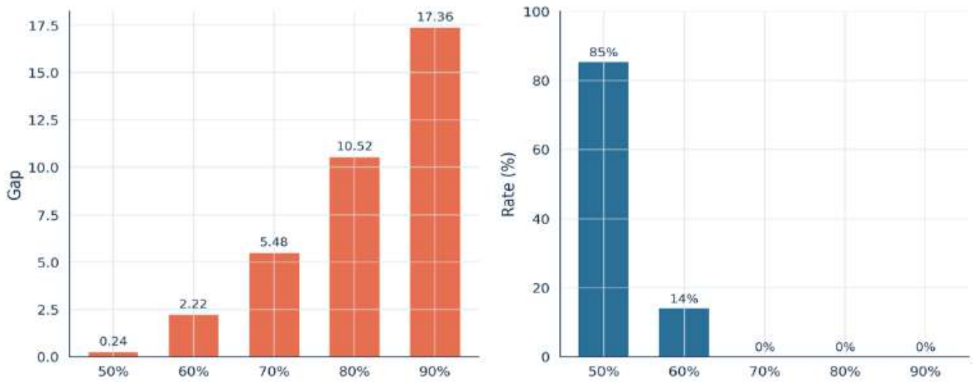

Figure 4: Search gap and exact-match rate across occupancy levels

This distinction is important for interpreting the proposed framework. The supervised model successfully captures the global relocation burden of the state, but dense instances remain challenging because search quality depends on repeated local ranking decisions. Therefore, the occupancy analysis reveals two different regimes. In lower-density systems, ValueNet-guided beam search is close to optimal and often exactly matches the R3C reference. In high-density systems, the predictive component remains informative, but the search component becomes the dominant bottleneck.

DISCUSSION

This paper proposed a learning-guided search framework for the restricted relocation problem in multi-deep AVS/RS systems. The framework combines Reverse Renumbering Relocation Construction (R3C), supervised value prediction, and beam search. The main methodological contribution is the creation of a scalable labeled-data pipeline that does not rely on repeated exact optimization for every training instance.

The empirical analysis on the medium-sized 5x40 warehouse showed that enlarging the synthetic dataset from 45,000 to 135,000 instances improved ValueNet validation performance and reduced the average beam-search gap to the R3C reference from 14.78 to 7.62 relocations. The occupancy study further showed that the method performs very well at 50% occupancy, remains useful at 60%, and becomes progressively less accurate in denser bays. These findings indicate that supervised learning can provide effective search guidance for relocation decisions, but that dense systems still pose significant challenges for limited-width search.

Future work should therefore focus on stronger neural representations, action-oriented models, and improved search strategies for high-occupancy instances. Even in its current form, however, the framework demonstrates that supervised learning can serve as a practical and scalable decision-support tool for relocation problems in multi-deep AVS/RS systems.

REFERENCES

[1] Bacci, T., Mattia, S., and Ventura, P. (2019), "The bounded beam search algorithm for the block relocation problem", *Computers & Operations Research*, 103, 252-264.
[2] Bacci, T., Mattia, S., and Ventura, P. (2020), "A branch-and-cut algorithm for the restricted Block Relocation Problem", *European Journal of Operational Research*, 287(2), 452-459.
[3] Jin, B., and Tanaka, S. (2023), "An exact algorithm for the unrestricted container relocation problem with new lower bounds and dominance rules", *European Journal of Operational Research*, 304(2), 494-514.
[4] Marolt, J., Kosanić, N., and Lerher, T. (2021), "Relocation and storage assignment strategy evaluation in a multiple-deep tier captive automated vehicle storage and retrieval system with undetermined retrieval sequence", *The International Journal of Advanced Manufacturing Technology*, 118, 3635-3658.
[5] Marolt, J., Rosi, B., and Lerher, T. (2022), "Reinforcement learning relocation assignment in the multiple-deep storage system", in *Proceedings of the XXIV International Conference MHCL'22*, Belgrade, 2022.
[6] Marolt, J., Šinko, S., and Lerher, T. (2023), "Model of a multiple-deep automated vehicles storage and retrieval system following the combination of Depth-First storage and Depth-First relocation strategies", *International Journal of Production Research*, 61(15), 4991-5008.
[7] Tang, L., Zhao, X., and Liu, J. (2015), "An exact approach for the Blocks Relocation Problem", *Expert Systems with Applications*, 42(17-18), 6408-6422.
[8] Tanaka, S., and Takii, K. (2018), "An exact algorithm for the unrestricted block relocation problem", *Computers & Operations Research*, 95, 12-31.
[9] Tanaka, S., Takii, K., Tierney, K., and Voß, S. (2022), "An exact approach to the restricted block relocation problem based on a new integer programming formulation", *European Journal of Operational Research*, 296(2), 485-503.
[10] Ye, R., Ye, R., and Zheng, S. (2023), "Machine learning guides the solution of blocks relocation problem in container terminals", *Transportation Research Record*, 2677(3), 721-737.
[11] Yu, Y., de Koster, R., Guo, X., and Zhu, W. (2020), "Machine learning-driven algorithms for the container relocation problem", *Transportation Research Part B: Methodological*, 139, 102-131.
[12] Zhou, Y., Kang, Y., Li, P., and Feng, X. (2026), "Container relocation problem: a systematic review and bibliometric analysis", *Journal of Shipping and Trade*, 11, Article 11.