

Based on Deep Learning Methods End-To-End Multi-Cycle Replenishment Strategy Research

Ziheng Wang*

Shanghai University, Shanghai, China

*Corresponding Author: Ziheng Wang

ABSTRACT

Based on the historical sales data and inventory data of a large retailer, this paper aims to output optimal replenishment decisions from a data-driven perspective, in order to reduce inventory-related costs. The main research body of this paper is the establishment of an end-to-end multi-cycle replenishment model, and a mathematical model of multi-cycle replenishment is constructed after a series of preprocessing of historical datasets. At the same time, an end-to-end multi-cycle replenishment deep learning model is built based on the mathematical model of multi-cycle replenishment, and the end-to-end label construction is carried out by using the preprocessed data set, so as to train the end-to-end multi-cycle replenishment deep learning model, and then compare it with other models.

KEYWORDS

Data-Driven, Deep Learning, Inventory Management, Replenishment Decisions, Cost Reduction.

1. INTRODUCTION

This paper focuses on an end-to-end multi-cycle replenishment strategy based on deep learning. Specifically, by using deep neural network technology, multiple information such as sales, inventory and purchase are input into the model to automatically learn and identify the sales patterns and market trends of different products, and finally output the optimal replenishment suggestions for each product in multiple cycles. The paper also pre-processed the relevant historical sales data, and established a multi-cycle replenishment model that considers the cost of lost sales, the cost of inventory carrying and the cost of fixed subscriptions. By applying deep learning methods to the pre-processed dataset, the order quantity and order time for the lowest cost of lost sales and inventory carrying costs are solved. The results of this paper show that an end-to-end multi-cycle replenishment strategy based on deep learning can effectively improve the efficiency and accuracy of inventory management and reduce inventory costs.

2. LITERATURE REVIEW

Inventory management based on deep learning is a popular research area in modern supply chain management. It includes methods such as demand forecasting, initial inventory decisions, and inventory decisions. It has a number of advantages over traditional methods, such as the ability to better forecast demand, more accurately determine initial inventory, and make replenishment decisions more optimally. Its main applications in inventory management include demand forecasting as well as inventory optimization.

Inventory demand forecasting is one of the important issues in inventory management. Its main purpose is to predict sales over a period of time in the future based on historical sales data and other factors. Cheng and Sun¹ proposed a hybrid model of Gray correlation analysis and time series neural networks. The model adopts the grey correlation analysis method, selects the main influencing factors as the input data of the time series neural network, and the experimental results show that the accuracy of the mixed model is significantly higher than that of the general prediction method and the BP neural network. Punia² studied the application of deep learning-based predictive analytics systems in demand forecasting. Drif et al.³ developed a Mutual Interaction Graph Attention Network "MIGAN", a new algorithm based on Massive Bipartite Graph (BGNN), and experimental studies on data showed that MIGAN has advantages over traditional algorithms in terms of prediction accuracy and efficiency. Kilimci⁴ developed an intelligent demand forecasting system that combines time series analysis techniques, support vector regression algorithms, and deep learning models to analyze historical data to improve demand forecasting. Giri⁵ proposed an intelligent forecasting system built using machine learning and deep learning techniques that combines the image feature attributes of clothing with its sales data to predict future demand. Mediavilla et al.⁶ summarized the current status of AI applications in demand forecasting in supply chain management, analyzed the advantages and disadvantages of deep learning methods, and pointed out the directions and challenges of future research. Mezzogori et al.⁷ proposed a deep learning method using entity embedding to predict the demand for highly differentiated products. In this method, the enterprise and product information is converted into vector representations, and the multi-layer perceptron model is used to train the prediction model. Experimental results show that the proposed method has better prediction effect than other deep learning methods.

Inventory optimization refers to maximizing profits and minimizing inventory carrying costs under an inventory management strategy. Deng Chuning et al.⁸ proposed a deep inventory management method using deep learning based on long short-term memory theory, which can make intelligent decisions on inventory management by predicting customer needs, so as to minimize costs and maximize profits. Boute et al.⁹ discussed the application of deep reinforcement learning in the field of inventory management and proposed a roadmap for inventory control. Zhou¹⁰ built a neural network model to achieve collaborative adjustment of inventory control, pricing decisions, and replenishment strategies. Wang¹¹ proposed a new Q-learning algorithm and experimentally verified it for different inventory management scenarios. The results show that the use of deep reinforcement learning method can effectively optimize inventory management, improve efficiency and achieve cost savings. Madeka et al.¹² proposed a method to optimize inventory by using the prediction advantages of deep convolutional neural networks and the time series forecasting technology of LSTM. This method combines current and historical data to improve management accuracy and reduce inventory costs. Dittrich et al.¹³ proposed an optimization method for inventory management based on deep Q-learning algorithms and applied it to the supply chain of linear process chains. By training a deep Q-learning neural network to predict inventory and purchasing decisions, and conducting experiments to verify its effect, the inventory cost was reduced and the supply chain efficiency was improved. Moor¹⁴ proposed a reward shaping method based on deep reinforcement learning techniques in perishable inventory management, which mainly includes reward shaping based on variable scale and reward shaping based on progressive scale. Experimental results show that the proposed method can significantly improve the accuracy and performance of deep reinforcement learning algorithm in perishable inventory management.

3. MODEL

This article considers an inventory management problem with a planning cycle of T , and on day 0, the first problem faced by decision-makers is how to determine the order quantity for that day. This decision will have a direct impact on subsequent operational efficiency and costs. If you order too much, you will have a high cost of ownership, and conversely, if you order too little, you may run out

of stock and lose the cost of sales. This article assumes that inventory that exceeds the demand of the current day can be carried over to the next day, but at the end of the planning cycle, any remaining inventory will no longer have value, i.e., its residual value will be zero. At the same time, it is assumed that the unmet demand of the day will immediately turn into a stockout.

In the model, the fixed cost of order K , the cost per unit of carrying h and the cost per unit of lost sales p are the key cost parameters, which together make up the total cost of the order. Uncertain daily demand D_t and fluctuating replenishment lead times L_t form random factors in the decision-making process. At the same time, this paper assumes that the contextual information (historical demand, replenishment lead time, supplier reliability and seasonal variation, etc.) on day t is x_t , and its sample space is X . In addition, this paper assumes that on $T+L_t$ day, the decision-maker will only get the actual information about the replenishment lead time L_t on day T . In order to describe the research background and problems more clearly, this paper introduces a definition function $(\cdot)^+ = \max(\cdot, 0)$, thus emphasizing that this article is concerned with non-negative values.

On day t , the corresponding inventory level I_t changes with the previous day's inventory level I_{t-1} , the number of orders ordered L_t days ago q_i and the demand D_t of the day, as shown in Equation 3-1:

$$I_t = \left(I_{t-1} + \sum_{i \in \{j: j+L_j=t\}} q_i - D_t \right)^+ \quad \forall t = 0, \dots, T \quad (3-1)$$

At the same time, the out-of-stock quantity on day t z_t also changes with the demand D_t of the day, the number of orders ordered L_t days ago q_i and the inventory level I_{t-1} of the previous day (T-1), which is updated as shown in Equation 3-2:

$$z_t = \left(D_t - \sum_{i \in \{j: j+L_j=t\}} q_i - I_{t-1} \right)^+ \quad \forall t = 0, \dots, T \quad (3-2)$$

The goal of this paper is to minimize the expected total cost by planning the replenishment quantity and replenishment time in a limited cycle, and the total cost is composed of the fixed order cost Ka_t , the holding cost hI_t and the out-of-stock cost, which are pz_t equations 3-3:

$$\begin{aligned} \min_{q_t, I_t, y_t, z_t} \mathbb{E} \left[\sum_{t=0}^T (Ky_t + hI_t + pz_t) | x_t \right] \\ = \min_{q_t, I_t, y_t, z_t} \mathbb{E} \left[\sum_{t=0}^T \left(Ky_t \right. \right. \\ \left. \left. + h \left(I_{t-1} + \sum_{i \in \{j: j+L_j=t\}} q_i - D_t \right)^+ \right. \right. \\ \left. \left. + p \left(D_t - \sum_{i \in \{j: j+L_j=t\}} q_i - I_{t-1} \right)^+ \right) | x_t \right] \end{aligned} \quad (3-3)$$

In addition to considering the total cost, the replenishment decision needs to meet the following seven constraints:

1. y_t indicates whether the order is made on the day, if the order is made, the y_t is 1, and the y_t is 0.

2. y_t is a binary variable
3. M is a large enough number that the order quantity cannot exceed its limit.
4. The out-of-stock of the day must be less than the demand of the day
5. The initial inventory is 0
6. The ordered quantity, the inventory level, and the out-of-stock quantity are all non-negative
7. Order quantities, inventory levels, and out-of-stock quantities are non-negative integers

As shown in Equations 3-4 to 3-10:

$$y_t = \begin{cases} 1, & q_t > 0 \\ 0, & q_t = 0 \end{cases} \quad \forall t = 0, \dots, T \quad (3-4)$$

$$y_t \in \{0,1\} \quad \forall t = 0, \dots, T \quad (3-5)$$

$$q_t \leq My_t \quad \forall t = 0, \dots, T \quad (3-6)$$

$$z_t \leq D_t \quad \forall t = 0, \dots, T \quad (3-7)$$

$$I_0 = 0 \quad (3-8)$$

$$q_t \geq 0, I_t \geq 0, z_t \geq 0 \quad \forall t = 0, \dots, T \quad (3-9)$$

$$q_t, I_t, z_t \in \mathbb{Z}^+ \quad \forall t = 0, \dots, T \quad (3-10)$$

In summary, the mathematical model of end-to-end multi-cycle replenishment constructed in this paper is as follows:

$$\min_{q_t, I_t, y_t, z_t} \mathbb{E} \left[\sum_{t=0}^T (Ky_t + hI_t + pz_t) | x_t \right] \quad (3-12)$$

s.t.

$$I_t = \left(I_{t-1} + \sum_{i \in \{j: j+L_j=t\}} q_i - D_t \right)^+ \quad \forall t = 0, \dots, T$$

$$z_t = \left(D_t - \sum_{i \in \{j: j+L_j=t\}} q_i - I_{t-1} \right)^+ \quad \forall t = 0, \dots, T$$

$$y_t = \begin{cases} 1, & q_t > 0 \\ 0, & q_t = 0 \end{cases} \quad \forall t = 0, \dots, T$$

$$y_t \in \{0,1\} \quad \forall t = 0, \dots, T$$

$$q_t \leq My_t \quad \forall t = 0, \dots, T$$

$$z_t \leq D_t \quad \forall t = 0, \dots, T$$

$$I_0 = 0$$

$$q_t \geq 0, I_t \geq 0, z_t \geq 0 \quad \forall t = 0, \dots, T$$

$$q_t, I_t, z_t \in \mathbb{Z}^+ \quad \forall t = 0, \dots, T$$

4. LABELING

A suitable training dataset is crucial for training deep learning neural networks. In the training dataset, each sample should contain observations of features and related objects or labels, and the goal of a

deep learning neural network is to learn a mapping from the input features to the output labels so that new inputs can be accurately predicted and classified. This is because the total data set preprocessed above contains only observations of historical demand and replenishment lead times, and there is no labeling information for optimal replenishment decisions. Therefore, this paper conducts a tagging process to calculate the historical optimal replenishment decision.

Because both historical demand and replenishment lead times are observed in historical data, the mathematical model of multi-cycle replenishment described above can be reduced to a deterministic batch problem. In this problem, the historical demand and replenishment lead times range from day $_H$ to day $-T$, and the random variables D_t and L_t are replaced with the d_t and l_t values to achieve the values.

In practice, retailers can spend a lot of time gathering historical demand and historical replenishment lead times, a process that can last months or years, so this article assumes that T_H is a large enough number.

In this paper, the Wagner-Whitin algorithm¹⁵ is used to solve the problem, which is a dynamic programming algorithm, which is mainly used to solve the problem of dynamic bulk ordering with fixed subscription cost and holding cost.

In order to facilitate expression, $q^*(D;L)$ is set as the multi-cycle solution to the multi-cycle replenishment problem with the demand vector D and the replenishment lead time vector L .

In reality, major retailers such as Freshippo, Fat Donglai, Sam's, etc., will update their replenishment plans on a daily basis, for example, some products have recently increased sales due to the popularity of online promotion, and need to replenish a larger number of stocks, or some categories need to stop or reduce orders due to seasonal changes due to declining sales. Considering the practical factors, in the end-to-end multi-cycle replenishment problem studied in this paper, it is only necessary to let the deep learning neural network learn the replenishment decision $q^*(D;L)$ of the day, which can be split into two parts, namely whether the replenishment order should be made on day t and the order quantity of the replenishment order. Thus, according to time t , the order sequence can be broken down into two decisions, i.e., the next order time o_t^* and the next order quantity w_t^* , which are defined as follows:

$$o_t^* = \inf \{i \in \mathbb{Z}: i \geq 0, q_{i+t}^* \neq 0\} \quad (4-1)$$

$$w_t^* = q_{o_t^*+t}^* \quad (4-2)$$

The goal of the end-to-end multi-cycle replenishment model is to learn a mapping from the input feature vector to the output o_t^* and w_t^* . In this article, the mapping from the replenishment vector $q^*(D;L)$ to $(o_t^*; w_t^*)$ is expressed as $\pi_t: \mathbb{R}^{(T_H)} \rightarrow \mathbb{R}^2$, which takes a T_H -dimensional real vector as input and then maps it to a two-dimensional real vector. By using neural networks to fit this complex relationship, the end-to-end multi-cycle replenishment model is able to accurately predict the optimal time and quantity for order decisions, enabling automated inventory management.

In this paper, the LSTM long short-term memory neural network is selected as the encoder. In the end-to-end multi-cycle replenishment problem, the input of the long short-term memory neural network is the optimal sequential sequence obtained from the end-to-end label construction or historical data described above. They are:

$$\{X_i\}_{i=-\tau+t}^{-1+t}; \{D_i\}_{i=-\tau+t}^{-1+t}; \{S_i\}_{i=-\tau+t}^{-1+t}; \{A_i\}_{i=-\tau+t}^{-1+t}; \{I_i\}_{i=-\tau+t}^{-1+t}; \quad (4-3)$$

Among them, X_t are the characteristics obtained before time t , such as product type, historical sales volume, etc., D_t is the demand observed before time t , S_t , A_t and I_t are the order time, order arrival time and historical inventory before time t , respectively, and τ represents the time span of

historical data considered before the current time t , that is, the time span from day $-\tau + t$ to day $-1 + t$.

At the same time, the outputs of the long short-term memory neural network are represented as Out 1, Out 2, Out 3 and Out 4. The meanings and corresponding labels are shown in Table 4.1 below:

Table 4.1 All outputs of long short-term memory network and corresponding labels

outputs	meanings	corresponding labels
<i>Out 1</i>	Replenishment lead time forecasts	l_t
<i>Out 2</i>	The next order arrival time	$r_t^*, r_t^* = o_t^* + l_t$
<i>Out 3</i>	Next order time	o_t^*
<i>Out 4</i>	Quantity for the next order	w_t^*

The goal of the end-to-end deep learning model training studied in this paper is to predict the order decision at the current time by learning relevant historical data, which can be expressed as fitting a function f , so that the input of the function f is the historical characteristics, demand, order time, order arrival time and historical inventory from day $-\tau+t$ to day $-1+t$, and the output is close to the order time and quantity of the current time t , that is:

$$f(\{X_i\}_{i=-\tau+t}^{-1+t}, \{D_i\}_{i=-\tau+t}^{-1+t}, \{S_i\}_{i=-\tau+t}^{-1+t}, \{A_i\}_{i=-\tau+t}^{-1+t}, \{I_i\}_{i=-\tau+t}^{-1+t}) \rightarrow (o_t^*; w_t^*) \quad (4-4)$$

The function of the training objective is:

$$\min_{\theta_1, \theta_2, \theta_4} : \sum_{t \in T_I} \{ \lambda_1 \mathcal{L}_{\theta_1}(Out1, l_t) + \lambda_2 \mathcal{L}_{\theta_2}(Out2, r_t^*) + \lambda_3 \mathcal{L}_{\theta_1, \theta_2}(Out3, o_t^*) + \mathcal{L}_{\theta_4}(Out4, w_t^*) \} \quad (4-5)$$

Among them, T_I is the total training time; θ represents the network weight that needs to be trained; λ_1 ; λ_2 and λ_3 are constants that can be used to adjust the step size of each iteration in gradient descent.

5. FIELD EXPERIMENT

The dataset in this article consists of transaction and in-store display data for 5 categories and 41 subcategories of 28,757 SKUs provided by a large retailer in China from January 2019 to October 2019, including daily sales, inventory, purchase costs, and face numbers. The total sales revenue exceeded 92 million yuan. A total of 174155 replenishments were made during the data collection period.

The experimental environment for training the deep neural network model in this paper is as follows: processor AMD Ryzen 9 6900HS with Radeon Graphics, memory 16GB, graphics card NVIDIA GeForce GTX3060ti, Windows 11, Python version 3.9.7, Python code written in Jupyter notebook.

In the training process, the learning rate is set to 0.05, and in the training objective function, the weight parameters are set to $\lambda_1 = \lambda_2 = 0.1$, $\lambda_3 = 1$, and K is set to 40, h is set to 1, and p is set to 8. The specific training curve is shown in Figure 5.1, where the abscissa represents the number of training sessions and the ordinate represents the average total cost value. This indicates that the value of the objective function does not change at all.

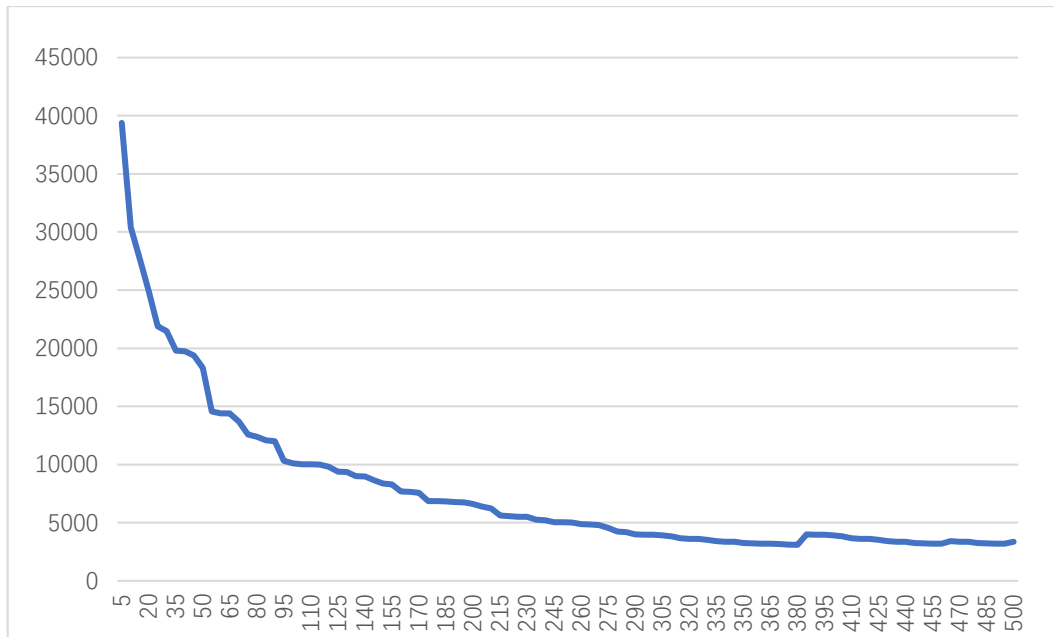


Figure 5.1 Training Graph

When the end-to-end deep learning model is trained to 500 times, the training results have tended to be stable and the early stop condition is satisfied.

After the parameters are set, the dataset is split according to different SKU_id, and different algorithms will predict the optimal replenishment decision for each sub-table in the next two months.

In this paper, 600 SKUs are randomly selected for 20 days of sales, order, and inventory data, and the inventory changes of these SKUs are tracked through simulation. For example, let's say a SKU is a mobile phone. First of all, starting with the first order, assuming that this order was placed on the 3rd day, we will check every day from the order date of the first order to the arrival date of the last order. If an order arrives during this period, such as on the 5th day, the corresponding amount of mobile phone inventory will be increased. If it doesn't arrive, we don't make any changes. Secondly, it will check if there are any sales orders during this time. For example, if someone buys 5 phones on day 6, they will reduce their phone inventory accordingly. If there is no sales order, no changes are made. This process will be ongoing, and each day the phone inventory levels will be adjusted based on order and sales data to maintain the accuracy of the inventory and thus calculate the total cost of inventory for that SKU.

In the basic setup, this paper assumes a carrying cost of 1 per SKU per day, a fixed order cost of 40 per order, and a unit out-of-stock cost of 8, i.e., h is set to 1, K is set to 40, and p is set to 8. The following is a comparison of the average cost results of each model.

Table 5.1 Comparison of average cost results by model

Algorithm Name	Fixed Cost	Cost Holding, , \	Cost Out of Stock	Total Cost
End-to-end deep learning models	271.36	2095.05	1327.78	3694.19
AdaBoost	344.54	2659.87	1559.02	4563.44
RF	277.35	4070.70	1774.51	6122.56
GBRT	604.19	2855.59	1372.38	4832.16
LinearRegression	869.01	4689.29	1535.59	7093.89

From the table, we can clearly see that the total cost of the end-to-end deep learning model is 3694.19, and the fixed cost, holding cost, and out-of-stock cost are 271.36, 2095.05, and 1327.78, respectively. This combined performance advantage makes end-to-end deep learning models a better choice than other comparison models, especially when it is necessary to find a balance between fixed costs, carrying costs, and out-of-stock costs, and can more effectively adapt and process complex data patterns, resulting in more efficient inventory management and replenishment decisions in real-world applications.

6. CONCLUSIONS

This paper focuses on an end-to-end multi-cycle replenishment strategy based on deep learning. Specifically, by using deep neural network technology, multiple information such as sales, inventory and purchase are input into the model to automatically learn and identify the sales patterns and market trends of different products, and finally output the optimal replenishment suggestions for each product in multiple cycles. The paper also pre-processed the relevant historical sales data, and established a multi-cycle replenishment model that considers the cost of lost sales, the cost of inventory carrying and the cost of fixed subscriptions. By applying deep learning methods to the pre-processed dataset, the order quantity and order time for the lowest cost of lost sales and inventory carrying costs are solved. The results of this paper show that an end-to-end multi-cycle replenishment strategy based on deep learning can effectively improve the efficiency and accuracy of inventory management and reduce inventory costs.

REFERENCES

- [1] Cheng, F.; Sun, R. Inventory Demand Forecast Based on Gray Correlation Analysis and Time Series Neural Network Hybrid Model. In 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD); IEEE: Guilin, 2017; pp 2491–2496. <https://doi.org/10.1109/FSKD.2017.8393167>.
- [2] Punia, S.; Shankar, S. Predictive Analytics for Demand Forecasting: A Deep Learning-Based Decision Support System. *Knowledge-Based Systems* 2022, 258, 109956. <https://doi.org/10.1016/j.knosys.2022.109956>.
- [3] Drif, A.; Cherifi, H. MIGAN: Mutual-Interaction Graph Attention Network for Collaborative Filtering. *Entropy* 2022, 24 (8), 1084. <https://doi.org/10.3390/e24081084>.
- [4] Kilimci, Z. H.; Akyuz, A. O.; Uysal, M.; Akyokus, S.; Uysal, M. O.; Atak Bulbul, B.; Ekmis, M. A. An Improved Demand Forecasting Model Using Deep Learning Approach and Proposed Decision Integration Strategy for Supply Chain. *Complexity* 2019, 2019, 1–15. <https://doi.org/10.1155/2019/9067367>.
- [5] Giri, C.; Chen, Y. Deep Learning for Demand Forecasting in the Fashion and Apparel Retail Industry. *Forecasting* 2022, 4 (2), 565–581. <https://doi.org/10.3390/forecast4020031>.
- [6] Mezzogori, D.; Zammori, F. An Entity Embeddings Deep Learning Approach for Demand Forecast of Highly Differentiated Products. *Procedia Manufacturing* 2019, 39, 1793–1800. <https://doi.org/10.1016/j.promfg.2020.01.260>.
- [7] Mohamadi, N.; Niaki, S. T. A.; Taher, M.; Shavandi, A. An Application of Deep Reinforcement Learning and Vendor-Managed Inventory in Perishable Supply Chain Management. *Engineering Applications of Artificial Intelligence* 2024, 127, 107403. <https://doi.org/10.1016/j.engappai.2023.107403>.
- [8] Deng, C.; Liu, Y. A Deep Learning-Based Inventory Management and Demand Prediction Optimization Method for Anomaly Detection. *Wireless Communications and Mobile Computing* 2021, 2021, 1–14. <https://doi.org/10.1155/2021/9969357>.
- [9] Boute, R. N.; Gijbrecchts, J.; Van Jaarsveld, W.; Vanvuchelen, N. Deep Reinforcement Learning for Inventory Control: A Roadmap. *European Journal of Operational Research* 2022, 298 (2), 401–412. <https://doi.org/10.1016/j.ejor.2021.07.016>.
- [10] Zhou, Y.; Guo, K.; Yu, C.; Zhang, Z. Optimization of Multi-Echelon Spare Parts Inventory Systems Using Multi-Agent Deep Reinforcement Learning. *Applied Mathematical Modelling* 2024, 125, 827–844. <https://doi.org/10.1016/j.apm.2023.10.039>.
- [11] Wang, Q.; Peng, Y.; Yang, Y. Solving Inventory Management Problems through Deep Reinforcement Learning. *J. Syst. Sci. Syst. Eng.* 2022, 31 (6), 677–689. <https://doi.org/10.1007/s11518-022-5544-6>.

- [12] Mediavilla, M. A.; Dietrich, F.; Palm, D. Review and Analysis of Artificial Intelligence Methods for Demand Forecasting in Supply Chain Management. *Procedia CIRP* 2022, 107, 1126–1131. <https://doi.org/10.1016/j.procir.2022.05.119>.
- [13] Dittrich, M.-A.; Fohlmeister, S. A Deep Q-Learning-Based Optimization of the Inventory Control in a Linear Process Chain. *Prod. Eng. Res. Devel.* 2021, 15 (1), 35–43. <https://doi.org/10.1007/s11740-020-01000-8>.
- [14] De Moor, B. J.; Gijbrecchts, J.; Boute, R. N. Reward Shaping to Improve the Performance of Deep Reinforcement Learning in Perishable Inventory Management. *European Journal of Operational Research* 2022, 301 (2), 535–545. <https://doi.org/10.1016/j.ejor.2021.10.045>.
- [15] Sadeghi, H.; Sadjadi, S.; Aryanezhad, M. An Improved Wagner-Whitin Algorithm. *38th International Conference on Computers and Industrial Engineering* 2008 2008, 3, 3023–3029.