Inventory Optimization Using Data Science Technologies for Supply Chain 4.0

Emna Bouazizi
University of Jeddah
College of Computing and Information Technology
Department of Information Systems
Jeddah, Saudi Arabia
aaltaher@uj.edu.sa

Ayman Khedr
University of Jeddah
College of Computing and Information Technology
Department of Information Systems
Jeddah, Saudi Arabia
aeelsayed@uj.edu.sa

Sherin Elfaioumy
Faculty of Computing and Information Sciences
Egypt University of Informatics, Egypt
Shereen.elfayoumi@eui.edu.eg

Mohamed Belal
Department of Computer Science
Faculty of Computers and Artificial Intelligence
Helwan University, Egypt
Belal@fci.helwan.edu.eg

Abstract: *In the context of supply chain 4.0, Data Science (DS) can improve operations and inventory management by using statistical and machine learning techniques. Additionally, big data can reveal valuable insights for predictive and prescriptive analytics, which can aid in enhancing competitiveness in today's business environment. Accurate demand forecasting in supply chain inventory is a fundamental step to improve inventory management, address ordering uncertainties, and minimize costs while meeting customer demands. This requires the efficient use of demand forecasting tools, assessing predictive data analysis techniques. Challenges can arise from inefficient ordering models, selecting accurate forecasting models, and considering issues like over-and underestimation leading to food waste and profit margin impacts, particularly for products with short shelf life. Other challenges include enhancing inventory optimization efficiency, adapting to dynamic demands, and formulating optimal inventory decisions. In this work, we introduce a supply chain 4.0 inventory management approach, where we combined DS techniques, predictive analytics, and big data approach to enhance inventory control. A prediction model is also introduced in order to forecast incoming and outgoing inventory. This model is based on a detailed dataset that takes into account the districts and seasons data. The model's performance was evaluated based on performance measurements, such as Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). Significantly, the results show that the random forest algorithm was the best in predicting OUT inventory quantities with an average error equal to 0.011, while the linear regression algorithm produced the best performance in predicting IN inventory quantities with an average error of 0.03.*

Keywords: *Supply chain 4.0, inventory management, regression algorithm, optimization.*

Received April 12, 2024; accepted September 29, 2024
https://doi.org/10.34028/iajit/21/6/4

1. Introduction

Inventory management is considered a core component in supply chain management. It includes several tasks such as inventory tracking, space optimization, deadstock tracking, and regulations satisfaction. To accomplish these tasks efficiently, advanced technologies and operations, such as demand prediction, planning, stocking and safety stock management are necessary.

In supply chain 4.0, inventory management has advanced and employed advanced technologies and data analysis, which results in enhancing operations and decision-making. These technologies include stock sensing, Radio Frequency IDentification (RFID), IoT, blockchain, 3D printing, autonomous vehicles, and digital twins.

In addition, supply chain 4.0 empowers digitalization, connectivity, and real-time capabilities, besides, it can enhance efficiency and sustainability in the whole process of supply chain management [39]. The integration of these emerging technologies empowers the ability of the supply chains. In particular, it can optimize inventory control, predict requirements accurately, and respond rapidly to varying demands [9].

Using data analysis in inventory management can include inventory prediction, automated replenishment, dynamic stock safety, and supplier collaboration analytics, which can employ both quantitative and qualitative methods [37]. Big data integration in inventory management enables real-time updates, perpetual inventory management at various levels, and adaptive strategies to encompass diverse inventory sources. This leads to improved logistics efficiency for carriers, reduced shrinkage for manufacturers, and enhanced accuracy for retailers, ultimately streamlining supply chain operations and ensuring customer satisfaction [37]. Supply chain 4.0 leverages cutting-edge technologies to transform and optimize inventory management, fostering responsiveness, cost-effectiveness, and adaptability across the supply chain [9].

Data Science (DS) integrates statistics and machine learning for data-driven insights. It also includes vast and diverse data, which can be used to reveal valuable insights. On the other hand, predictive analytics forecasts future outcomes for proactive decisions, prescriptive analytics optimizes performance with actionable insights. Furthermore, diagnostic analytics identifies root causes for problem-solving, and finally descriptive analytics explains past performance and trends using historical data. These terms collectively empower organizations to harness data's potential for improved efficiency and competitiveness in the modern business landscape [14].

Leveraging DS and Big Data Analytics (BDA) in supply chain 4.0 inventory management is essential to enhance performance and competitiveness. Inventories data sources can include huge amounts of logistics data and customer patterns. Based on these data sources, BDA can extract valuable insights predictions, and knowledge. Thus, DS can be applied to making informed decisions, reducing costs, and improving sustainability. In addition, organizations stay responsive to customer demands and market trends. Advanced tools and software, along with digital supply chain twins, empower professionals to optimize operations for greater efficiency. To conclude, DS and BDA are pivotal in transforming supply chain management into a data-driven and strategically advantageous endeavor [14].

The purpose of the paper is to introduce a supply chain 4.0 inventory management approach using DS techniques, predictive analytics, and big data to enhance inventory control. The main research problem is improving accurate demand forecasting in supply chains to address ordering uncertainties and minimize costs while meeting customer demands. The objectives include efficiently using demand forecasting tools, assessing predictive data analytics techniques, selecting accurate forecasting models, and addressing issues like over- and underestimation leading to food waste and profit margin impacts. Additionally, challenges include enhancing inventory optimization efficiency, adapting to dynamic demands, and formulating optimal inventory decisions. The inventory status is well-prepared to store the products correctly. We did not factor logistics into our analysis because our focus is on tracking the quantities in and out, allowing us to predict future stock levels and ensure seasonal demand is met without any shortages.

The paper is structured as follows: section 2 introduces inventory management and optimization within supply chain 4.0. In section 3, there's a review of literature and related works. Section 4 presents the proposed work. Finally, section 5 discusses the simulation methodologies and the obtained results.

2. Inventory Management and Optimization in the Era of Supply Chain 4.0

Industry 4.0 provides digitalization and automation to the inventory processes, which allows for automatic order tracking based on real-time data and information sharing. For example, it automates the widely-used ABC inventory classification for smart products that carry real-time location and valued information. In addition, it provides solutions for uncertainties in supply lead-times, customer demand estimation, and dynamic factors from suppliers [39].

Furthermore, industry 4.0 influences other critical parameters like buying cost, storage expense, procurement expense, and selling cost of the item, thanks to its ability to facilitate real-time data exchange. It also revolutionizes the inventory system review practices and encourages the transition from periodic to continuous reviews. Therefore, Industry 4.0 can improve stock management systems, and provide efficiency, accuracy, and adaptability in supply chain operations [39].

The main purpose of an inventory is to deal with uncertainties from providers and customers to meet demand. It's crucial to balance effectively the inventory levels. Keeping more inventory results in additional costs, on the other hand, having too little will not satisfy customers. Therefore, it's essential to optimize inventory to reduce costs and fulfill customer satisfaction. Inventory optimization involves determining when and how much to order, and forming an inventory strategy. This ideal strategy is established by solving an optimization problem, comprising an objective function and constraints that establish connections among system variables. The goal can either reduce the overall operational costs or enhance customer service levels, with decision parameters including order timing and order quantities. Data analytics can accurately formalize these problems in industry 4.0, where it can reach the best inventory policies. Moreover, inventory management can digitize and analyze supplier and customer data, which can minimize the operations cost of inventory management in supply chains [39].

3. A Taxonomy and Related Works

Demand forecasting and analysis can be used to improve inventory management in supply chains. This section presents various paradigms in the literature for demand analysis and forecasting.

The various aspects of demand analysis and forecasting methodologies are covered in the subsections. Every section focusses on particular domains related to inventory management. This taxonomy offers a thorough summary of the methods used to effectively regulate inventory by utilizing supply and demand data.

3.1. Demand Forecasting Techniques

There were several different demand forecasting methods proposed in [10, 11, 25, 29, 31]. The used techniques include machine learning techniques and Bayesian regression analysis.

The work of Stanelyte [31] focuses on developing an inventory optimization and demand forecasting model for the retail market. On the other hand, Gattinoni [11] suggest the possibility of using better forecasting to reorganize inventory management, the model examines the relationship between demand forecasting, inventory planning, and supply chain efficiency.

In another work, a comparative study of demand prediction techniques in leathercraft companies that are not large-sized is conducted by Purnamasari *et al*. [25]. However, Shao *et al*. [29] recommend an inventory optimization plan for automotive spare parts.

3.2. Demand Uncertainty Management

In inventory optimization, the problems of seasonality, demand variability, and uncertainty were addressed [8, 18, 23, 35]. Chandrika [4] conducted a simulation-based analysis of inventory optimization strategies for a retail supplier specializing in protective gear. It emphasizes managing demand variability and lead to time unpredictability.

The study presented by Li *et al*. [18] explores the integration of machine learning algorithms for demand analysis into cross-border e-commerce Enterprise Resource Planning (ERP) systems to refine inventory management tactics. In a similar work, Duc *et al*. [8] formulate a mixed-integer linear programming model, which operates under the base stock policy. It aims to minimize the total costs associated with safety inventory holding and shortage.

Nasution *et al*. [23] investigate the efficacy of machine learning techniques in demand forecasting to enhance inventory management within supply chains. Finally, Tirkolaee *et al*. [35] propose a robust optimization framework to solve the multi-echelon capacitated location-allocation-inventory problem, it aims at optimizing the design of supply chain networks for maximum efficiency.

3.3. Dynamic Pricing and Inventory Management

Decision-making based on dynamic pricing, real-time demand fluctuations and dynamic inventory management was studied [5, 10, 26, 33].

Chinello *et al*. [5] investigate the principal drivers of inventory optimization within supply chains. On the other hand, Teerasoponpong and Sopadang *et al*. [33] introduce a specialized Decision Support System (DSS), which is customized to the needs of Small-and Medium-sized Enterprises (SMEs).

A strong optimization framework is presented by Qiu *et al*. [26], which can be used for inventory decisions in a retail setting. The framework tackles the problem of uncertain demand across many product lines based on a constrained budget with various providers.

Finally, under uncertain demand, Gao *et al*. [10] survey the difficulties in dynamic pricing and dynamic inventory management. The fundamental goal of this work is to maximize total multi-period earnings.

3.4. Perishable Product Demand Forecasting

Several works aim to predict the demand for perishable products [1, 13, 20, 26, 34].

Gurnani *et al*. [13] focus on using machine learning techniques to predict the needs of a supermarket to maintain transactions in the future.

Similar to this work, a robust optimization framework for a retail environment is introduced by Qiu *et al*. [26]. The demand volatility is predicated on tight financial constraints for a range of vendors and goods.

However, Mankar and Khan [20] offer an overview of inventory control in the context of transportation and logistics. In addition, Afridi *et al*. [1] present a performance measurement method for Vendor-Managed Inventory (VMI). A Reinforcement Learning (RL) approach is suggested for the ideal replenishment strategy for the semiconductor industry.

Finally, the IACPPO inventory replenishment model is proposed by Tian *et al*. [34]. The model combines the Proximal Policy Optimization (PPO) and Advantage Actor-Critic (A2C) algorithms. With an emphasis on-demand analysis, the work aims to optimize inventory management within supply chains.

3.5. Cross-Border E-Commerce Demand Analysis

The demand patterns and variability were investigated [12, 17, 18, 32, 34] within cross-border e-commerce contexts.

Methodologies, such as big data analysis and predictive analytics, were employed to enhance efficiency in inventory management and alleviate stockouts.

Li *et al*. [18] focus on using machine learning algorithms within international online commerce ERP systems to refine inventory management strategies using demand analysis.

Similarly, Tang *et al*. [32] introduce an adaptive inventory management system utilizing Artificial Intelligence (AI) for cross-border e-commerce entities. The system aims to optimize inventory levels based on extensive data forecasting.

On the other hand, Kim and Jeong *et al*. [17] propose a demand forecasting plan incorporating ARIMA time series analysis and an Economic Order Quantity (EOQ) model. The objective is to optimize inventory management in mass-production manufacturing facilities, based on material supply excess and demand variability.

Guo *et al*. [12] attempt to optimize inventory management across supply chains by integrating demand prediction with web search data. The model employs a backpropagation neural network for model training.

Based on demand analysis, Tian *et al*. [34] introduce the IACPPO inventory replenishment model, which integrates A2C and PPO algorithms to refine inventory management within supply chains.

3.6. Retail Market Demand Analysis

The demand forecasting techniques and inventory management strategies for the retail sector were introduced in [10, 16, 25, 29, 31]. Methodologies are employed such as Bayesian methods, machine learning algorithms, and analytical tools for practical application. Specifically, Stanelyte [31] endeavors to construct a demand forecasting and inventory optimization model through the analysis of diverse techniques for demand prediction within the retail market. Similarly, Purnamasari *et al*. [25] conducts an initial comparative study of machine learning and traditional projection techniques for demand forecasting in small and medium-sized leathercraft enterprises, aiming to refine inventory management strategies. Additionally, Shao *et al*. [29] introduce an inventory optimization scheme for automotive spare parts, grounded in demand forecasting and lateral transfer strategies. Furthermore, Gao *et al*. [10] address the joint decision-making problem of adaptive pricing and flexible inventory control in apparel supply chains, with a focus on optimizing total multi-period profits while considering demand uncertainty. Lastly, Keith [16] investigates the influence of advanced forecasting techniques on inventory management performance within supply chains, drawing insights from global practices and empirical evidence.

3.7. Demand-Driven Replenishment Policies

The replenishment policies grounded in-demand data analysis were developed in [3, 8, 28, 30, 36]. Objectives include maintaining stockouts, reducing storage costs, and optimizing inventory ordering decisions across various supply chain contexts. Calero Mantilla and Esposito [3] propose an end-to-end approach to enhance runout inventory management in supply chains based on demand analysis. Similarly, Sihotang [30] delve into the application of exponential smoothing methods within a DSS to optimize inventory ordering decisions within the retail sector. Moreover, Valencia-Cardenas [36] undertakes a comparison of different forecasting models for demand prediction across multiple products in supply chains and subsequently optimizes inventory management policies for a gas service station in Colombia. Additionally, Duc *et al*. [8] develop a mixed-integer linear programming methodology for inventory positioning in a supply chain under the base stock policy, with the goal of minimizing total costs associated with safety inventory holding and shortages. Lastly, Dittrich and Fohlmeister [7] introduce a method for adaptive and continuously self-optimizing inventory control on a global scale utilizing deep q-learning techniques.

3.8. Demand Prediction Integration

The demand prediction techniques were integrated [12, 13, 23, 24, 28] with inventory management strategies, leveraging various methodologies such as web search data, machine learning algorithms, and ensemble deep learning-based methodologies to develop more effective inventory policies. Specifically, Seyedan *et al*. [28] investigated the utilization of ensemble deep learning-based forecasting methods to predict future demand within the online retail industry, aiming to enhance inventory optimization. Likewise, Gurnani *et al*. [13] explore the application of machine learning algorithms to predict the amount needed for a supermarket to maintain its upcoming sales, thereby contributing to inventory optimization within supply chains based on demand analysis. Furthermore, Poi and Opara [24] investigate the correlation between inventory optimization techniques, such as Just-In-Time (JIT) and ABC analysis, and customer satisfaction in petroleum marketing firms in Rivers State, Nigeria. Additionally, Nasution *et al*. [23] scrutinize the impact of applying machine learning techniques to demand forecasting to enhance inventory management practices within supply chains. Finally, Guo *et al*. [12] endeavor to optimize inventory management in supply chains by amalgamating demand prediction with web search data and employing a backpropagation neural network for training the prediction model.

3.9. Demand Pattern Analysis

The demand patterns and variation were analyzed [6, 15, 27, 38] to categorize inventory items into different product families, employing machine learning algorithms and optimization techniques to optimize inventory stocking and distribution. In the work of Kandemir [15], two algorithms are introduced: the first is based on statistical analysis. The second is based on unsupervised machine learning utilizing k-means clustering, to distinguish between seasonal and non-seasonal products for effective inventory management within supply chains. Sarath Kumar and Ramana [27] discusses the significance of business intelligence and analytics in operations management, emphasizing inventory optimization in supply chains based on demand analysis. Additionally, Wu *et al*. [38] introduce a Deep Reinforcement Learning (DRL) method tailored for supply chain management, aiming to optimize inventory management strategies based on demand analysis. Furthermore, Dehaybe *et al*. [6] explore the application of DRL in solving single-item lot sizing problems with predetermined expenses, lead time, and both unfulfilled orders and missed sales, ultimately aiming to optimize inventory management practices within supply chains.

3.10. Demand Adjustment and Collaboration

The collaborative approaches with key clients and the incorporation of forward-looking demand adjustments were advocated [17, 19, 21, 22, 25] into traditional inventory planning models, to enhance operational efficiency and financial performance.

Makarova [19] work, a collaborative approach to inventory optimization in supply chains is proposed, with a focus on improving operational and financial performance through robust collaboration with key Business-to-Business (B2B) clients. Purnamasari *et al*. [25] conduct an initial comparison of machine learning and traditional projection techniques for predicting demand in small and medium-sized leathercraft enterprises, aiming to bolster inventory management strategies. Furthermore, Mohamadi *et al*. [21] tackle the challenge of perishable inventory allocation within a two-echelon supply chain using DRL, intending to mitigate wastage and shortages while enhancing efficiency. Kim and Jeong *et al*. [17] present a demand forecasting plan utilizing ARIMA time series analysis and an EOQ-based model to optimize inventory management in response to excessive material supply and demand variability within a mass-production manufacturing factory.

Finally, Namir *et al*. [22] proposes a model that integrates time series analysis, machine learning algorithms, and combinatorial optimization.

4. Materials and Methods

Focused on predicting seasonal product quantities, the model explores the influence of seasons on anticipated product volumes. To prevent product shortages across various districts, it is essential to maintain inventory levels that align with customer needs during each season. The model undergoes a two-phase process: in the initial phase, it leverages quantities sold in the same season and districts five years ago as a reference point, while the second phase incorporates the estimation of the future quantities that must be fed into the inventory. The second phase takes the feedback from phase one. Figure 1 explains the abstract model that illustrates the general view of our model.

Before explaining the idea of the model, some terms must be clarified such as:

- OUT-inventory (demand): is the quantity sold from inventory.
- IN-inventory (order): is the quantity required to feed the inventory.

The model in Figure 1 illustrates a multi-step regression analysis process involving two phases. In the first stage, input features are used to perform regression analysis, aiming to predict OUT or demand value. This predicted value is subsequently utilized as an input in the second phase of regression analysis, along with additional features, to predict the final decision of determining the value or order (IN) quantities.

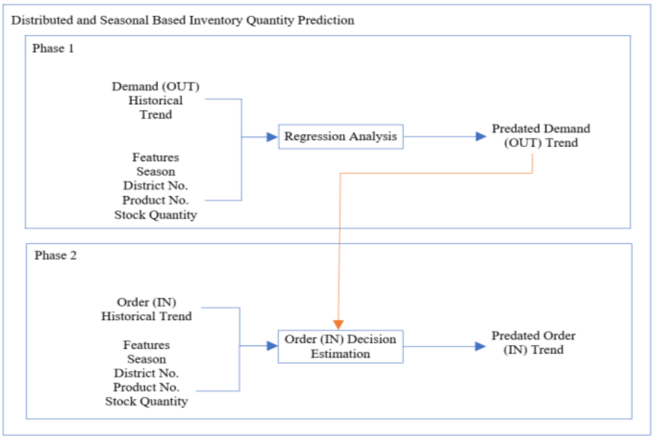


Figure 1. Inventory prediction model.

In essence, the process involves a cascading approach where the output from one regression analysis serves as an input for another regression analysis. This hierarchical structure allows for a more comprehensive understanding and prediction of the target variables, leveraging both the initial input features and the intermediate target variables generated in the first stage of analysis. Such an approach can be particularly useful in scenarios where the prediction of the second label is dependent on both the original input features and the predicted values from the first regression analysis.

Our dataset contains historical information from five years ago, covering the replenishment of ten products across five districts during four distinct seasons, along with the subsequent depletion of these products from the inventory.

For the First phase, the main feature that is considered as a core input for this phase is demand (OUT) historical inventory in addition to the rest of the features such as season, districts, products, and stock quantity.

Within the model, each facet of machine learning is carefully considered, elucidating the procedure for implementing it with an inventory dataset. This procedure entails managing the dataset, comprising attributes linked to inventory items alongside their corresponding outcomes. The preprocessing methodologies and model selection are tailored to the attributes within the dataset and the particular predictive task utilizing machine learning algorithms.

The steps that are taken into consideration while implementing the model are:

1. Data analysis: the data analysis of the inventory dataset aims to recognize features, identify

correlations, and gain insights into variable relationships.
2. Data cleaning and handling missing values and encoding: the dataset can be cleaned by specifying missing or inconsistent values. Other tasks include imputing missing values, removing outliers, and handling categorical variables. Districts, products, and seasons data were encoded to ensure proper interpretation.
3. Data division (training/testing): the dataset is divided into training (80%) and testing sets (20%) for the training and evaluation processes.
4. Data normalization (scaling, transformation): the features are normalized to ensure they are on a similar scale. Normalization can significantly reduce errors in regression algorithms.
5. Selecting regression model: a suitable regression algorithm should be selected, such as linear regression or decision tree regression, for predicting IN/OUT inventory values based on inventory features.
6. Model training: the selected regression model is trained using the training dataset. Random forest was used for input data prediction since it can handle input data complexities such as non-linear relationships and missing data, as well as its robustness to overfitting. As an ensemble method, it results in a more stable and accurate prediction.

In the feedback loop, we used linear regression that provides faster and more scalable prediction based on the smoothed, clean and well-preprocessed output of the forward path, and it is more suitable for real-time predictions.

7. Model evaluation: the model's effectiveness is assessed utilizing measurements like Mean Squared Error (MSE) or R-squared on the validation set.

Our inventory dataset is assumed to include features like the seasons, the products that must be provided to cover customer needs, the quantities inside the inventory and sold from inventory, and districts where inventories are located. The output labels are the predicted IN and OUT inventory quantities for each item in specific seasons and districts. Table 1 illustrates the encoded values of the dataset features.

Table 1. The encoded values of the dataset features.

Feature	Encoded value
Districts	1,2,3,4,5
Products	1,2,3,4,5,6,7,8,9,10
Seasons	1,2,3,4

Our dataset contains historical information from five years ago, covering the replenishment of ten products across five districts during four distinct seasons, along with the subsequent depletion of these products from the inventory.

The inventory dataset is also assumed to include features like item, district code, season code, product code, OUT quantity in (used as a label in first phase and as a feature in second phase), and quantity IN (used as a label in the second phase). Additionally, shortages and surpluses in quantities are calculated in the dataset, which supports our model in predicting future quantities.

5. Results and Discussion

As mentioned earlier, our model comprises two phases for prediction. One focuses on forecasting 'OUT-Inventory' products' quantities expected to be sold in the same season and district, while the other predicts 'IN-inventory' products' quantities. To replenish inventory within a specific season and district.

In regression machine learning, we evaluate our model's performance -after putting weights for some features to unify the features' forms-by calculating error, which is the difference between observed and forecasted values. Lower error values indicate improved predictive capabilities. Error metrics like MSE, Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) are commonly used in regression machine learning to measure prediction accuracy. The subsequent segments offer a brief explanation of these error metrics and their mathematical formulas.

a) MSE:

MSE is computed by taking the average of the squared differences between actual and predicted values as shown in Equation (1):

$$MSE = \left(\frac{1}{n}\right)\sum(actual\ value - predicted\ value)^2 \quad (1)$$

where n is the number of data entries.

b) RMSE:

RMSE is the square root of MSE and offers a metric of the average magnitude of the errors as shown in Equation (2):

$$RMSE = sqrt(MSE) \quad (2)$$

c) MAE:

MAE is computed by taking the average of the absolute differences between actual and predicted values as shown in the next equation:

$$MAE = \left(\frac{1}{n}\right)\sum|actual\ value - predicted\ value| \quad (3)$$

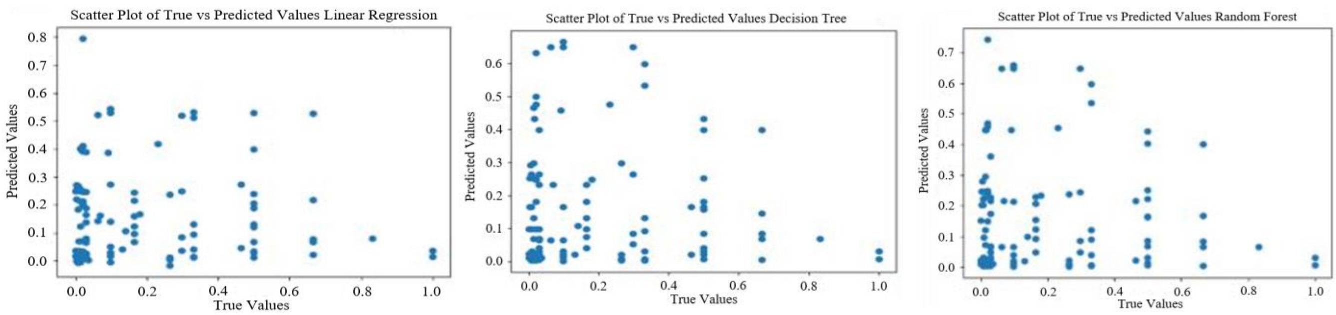
5.1. OUT Inventory Prediction

After using regression machine learning techniques (linear regression, decision tree, and random forest) for OUT Inventory and IN inventory, the tables and figures below show the results. Starting with OUT inventory prediction as the first phase of our model, Table 2 illustrates the error values between the forecasted OUT inventory entries and the actual entries.

In OUT prediction, random forest demonstrates lower error calculations in the comparison between actual and predicted values, indicating a high degree of proximity between the two sets of data. Figure 2 contains the three plotting figures of each algorithm used in regression for predicting the OUT inventory quantities.

Table 2. Error metrics for the three used machine-learning algorithms for OUT inventory prediction.

Algorithm	MSE	RMSE	MAE
Linear regression	0.002783	0.052757	0.035866
Decision tree	0.000786	0.028039	0.009716
Random forest	0.000517	0.022738	0.011559



a) Scatter plot of true vs. predicted values linear regression. b) Scatter plot of true vs. predicted values decision tree. c) Scatter plot of true vs. predicted values random forest.

Figure 2. The three plotting figures of each algorithm used for predicting the OUT inventory quantities.

As noticed, the random forest [2] algorithm demonstrates average lower error calculations in the comparison between the actual and the predicted values, indicating a high degree of proximity between the two sets of data.

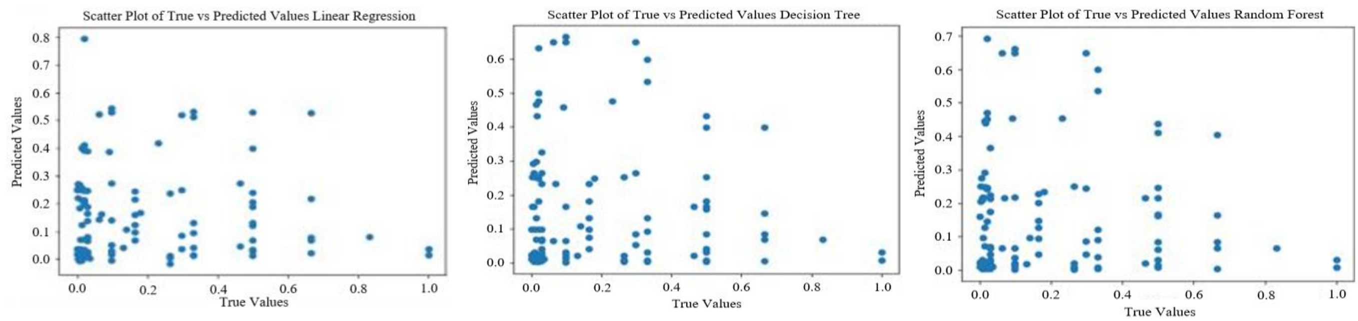
5.2. “IN” Inventory Prediction

The model first predicts the expected demand, which is then used as input for the order prediction process. Using demand trends as input in order predictions has several advantages: First, it improves forecast accuracy by learning from past errors, whether the demand was overestimated or underestimated. Second, this feedback process helps detect trends and seasonality, reflecting how customer behavior and preferences change over time due to factors like promotions, economic shifts, weather, or recurring events. As a result, it considers the relationships between demand and other influences such as pricing, marketing efforts, or external events. Finally, this feedback helps to smooth out order predictions, providing a reference point that reduces fluctuations in expectations.

Table 3. Error metrics for the three used machine-learning algorithms for IN inventory prediction.

Algorithm	MSE	RMSE	MAE
Linear regression	0.003456	0.058791	0.038242
Decision tree	0.002480	0.289271	0.196998
Random forest	0.001685	0.288430	0.197097

For phase two, the Table 3 displays the error values for three types of error metrics: MSE, RMSE, and MAE, which were promising. Figure 3 shows three plots, each representing an algorithm used in regression for predicting the OUT inventory quantities.



a) Scatter plot of true vs. predicted values linear regression. b) Scatter plot of true vs. predicted values decision tree. c) Scatter plot of true vs. predicted values random forest.

Figure 3. The three plots, for the three algorithms used in regression for predicting the IN inventory quantities.

As noticed, linear regression demonstrates lower error calculations in the comparison between actual and predicted values, indicating a high degree of proximity between the two sets of data. We compared our work with a study made to predict wheat quantities [40], but they used a Support Vector Machine (SVM) with an accuracy of 93.2% while our model provides more than 95% based on error metrics as shown in Table 4.

Table 4. Comparison between the previous model with the proposed model.

Inventory model	SVM model	The proposed model
No. of phases	Single phase	Two cascading phases
Approach	SVM	Random forest-linear regression
Accuracy	93.2%	More than 95%

6. Conclusions

In our supply chain 4.0 inventory management model, we integrated DS, predictive analytics, and big data to optimize inventory control. We carefully crafted a dataset and built a prediction model focused on anticipating IN (order) and OUT (demand) inventory quantities in specific districts during distinct seasons. We evaluated the model's effectiveness using error metrics such as MSE, RMSE, and MAE. Notably, the random forest algorithm performed well in predicting OUT inventory quantities with an average error of 0.011. While the linear regression algorithm showed promising results for predicting IN inventory quantities with an average error of 0.03.

Acknowledgment

This work was funded by the University of Jeddah, Jeddah, Saudi Arabia, under grant no. (UJ-23-DR-260). Therefore, the authors thank the University of Jeddah for its technical and financial support.

References

[1] Afridi M., Nieto-Isaza S., Ehm H., Ponsignon T., and Hamed A., "A Deep Reinforcement Learning Approach for Optimal Replenishment Policy in a Vendor Managed Inventory Setting for Semiconductors," *in Proceedings of the Winter Simulation Conference*, Orlando, pp. 1753-1764, 2020. DOI:10.1109/WSC48552.2020.9384048

[2] Baradieh K., Zainuri M., Kamari M., Yusof Y., Abdullah H., Zaman M., and Zulkifley M., "RSO Based Optimization of Random Forest Classifier for Fault Detection and Classification in Photovoltaic Arrays," *The International Arab Journal of Information Technology*, vol. 21, no. 4, pp. 636-660, 2024. https://doi.org/10.34028/iajit/21/4/8

[3] Calero Mantilla F. and Esposito A., End-of-Life Inventory Optimization during Runout Events for a Manufacturing Automotive Company, Master Thesis, Massachusetts Institute of Technology 2023. https://dspace.mit.edu/handle/1721.1/107234

[4] Chandrika V., A Simulation Based Inventory Optimization Heuristic for a Retail Supply Chain, Master Thesis, University of Cincinnati, 2015. http://rave.ohiolink.edu/etdc/view?acc_num=ucin1439309194

[5] Chinello E., Herbert-Hansen Z., and Khalid W., "Assessment of the Impact of Inventory Optimization Drivers in a Multi-Echelon Supply Chain: Case of a Toy Manufacturer," *Computers and Industrial Engineering*, vol. 141, pp. 106232, 2020. https://doi.org/10.1016/j.cie.2019.106232

[6] Dehaybe H., Catanzaro D., and Chevalier P., "Deep Reinforcement Learning for Inventory Optimization with Non-Stationary Uncertain Demand," *European Journal of Operational Research*, vol. 314, no. 2, pp. 433-445, 2024. https://doi.org/10.1016/j.ejor.2023.10.007

[7] Dittrich M. and Fohlmeister S., "A Deep Q-Learning-Based Optimization of the Inventory Control in a Linear Process Chain," *Production Engineering Research and Development*, vol. 15, no. 1, pp. 35-43, 2021. https://link.springer.com/article/10.1007/s11740-020-01000-8#citeas

[8] Duc N., Tai P., and Buddhakulsomsiri J., "Inventory Positioning in Supply Chain Network: A Service-Oriented Approach," *IEEE Access*, vol. 10, pp. 92986-93002, 2022. https://ieeexplore.ieee.org/abstract/document/9869807

[9] Fernandez-Carames T., Blanco-Novoa O., Froiz-Miguez I., and Fraga-Lamas P., "Towards an Autonomous Industry 4.0 Warehouse: A UAV and Blockchain-Based System for Inventory and Traceability Applications in Big Data-Driven Supply Chain Management," *Sensors*, vol. 19, no. 10, pp. 2394, 2019. https://www.mdpi.com/1424-8220/19/10/2394

[10] Gao J., Zheng Y., and Mao X., "A Joint Decision Model of Dynamic Pricing and Dynamic Inventory in an Apparel Supply Chain with Demand Learning," *in Proceedings of the International Conference on Logistics Systems and Intelligent Management*, Harbin, pp. 692-696, 2010. DOI:10.1109/ICLSIM.2010.5461329

[11] Gattinoni S., Supply Chain Planning: Inventory Optimization through the Demand Forecasting Tools and Algorithm, Master Thesis, Politecnico di Milano, 2022. https://hdl.handle.net/10589/209494

[12] Guo X., Liu C., Xu W., Yuan H., and Wang M., "A Prediction-Based Inventory Optimization Using Data Mining Models," *in Proceedings of the 7th International Joint Conference on Computational Sciences and Optimization*, Beijing, pp. 611-615, 2014. https://ieeexplore.ieee.org/document/6923759

[13] Gurnani P., Hariani D., Kalani K., Mirchandani P., and CS L., *Data Intelligence and Cognitive Informatics*, Springer, 2022. https://link.springer.com/chapter/10.1007/978-981-16-6460-1_41

[14] Jahani H., Jain R., and Ivanov D., "Data Science and Big Data Analytics: A Systematic Review of Methodologies Used in the Supply Chain and Logistics Research," *Annals of Operations Research*, pp. 1-58, 2023. https://doi.org/10.1007/s10479-023-05390-7

[15] Kandemir B., "A Methodology for Clustering Items with Seasonal and Non-Seasonal Demand Patterns for Inventory Management," *Journal of Advanced Research in Natural and Applied*

Sciences, vol. 8, no. 4, pp. 753-761, 2022. https://dergipark.org.tr/en/download/article-file/2409588

[16] Keith E., "Optimizing Inventory Management through Advanced Forecasting Techniques in Supply Chains," *European Journal of Supply Chain Management*, vol. 1, no. 1, pp. 22-30, 2023. https://forthworthjournals.org/journals/index.php/EJSCM/article/view/38

[17] Kim J. and Jeong J., "Simulation Evaluation for Efficient Inventory Management Based on Demand Forecast," *in Proceedings of the 18th International Conference on Computational Science and its Applications*, Melbourne, pp. 639-650, 2018. https://link.springer.com/chapter/10.1007/978-3-319-95162-1_44

[18] Li J., Wang T., Chen Z., and Luo G., "Machine Learning Algorithm Generated Sales Prediction for Inventory Optimization in Cross-Border E-Commerce," *International Journal of Frontiers in Engineering Technology*, vol. 1, no. 1, pp. 62-74, 2019. DOI: 10.25236/IJFET.2019.010107

[19] Makarova O., "Predictive Method of Inventory Optimization in Industrial Manufacturer Performance," *E3S Web of Conferences*, vol. 247, pp. 1-4, 2021. https://doi.org/10.1051/e3sconf/202124701066

[20] Mankar A. and Khan A., "A Review of Inventory Management Research in Transport and Logistics," *in Proceedings of the 14th International Conference on Computing Communication and Networking Technologies*, Delhi, pp. 1-7, 2023. DOI:10.1109/ICCCNT56998.2023.10307815

[21] Mohamadi N., Niaki S., Taher A., and Shavandi A., "An Application of Deep Reinforcement Learning and Vendor-Managed Inventory in Perishable Supply Chain Management," *Engineering Applications of Artificial Intelligence*, vol. 127, pp. 107403, 2024. https://doi.org/10.1016/j.engappai.2023.107403

[22] Namir K., Labriji H., and Ben Lahmar A., "Decision Support Tool for Dynamic Inventory Management Using Machine Learning, Time Series and Combinatorial Optimization," *Procedia Computer Science*, vol. 198, pp. 423-428, 2022. https://doi.org/10.1016/j.procs.2021.12.264

[23] Nasution A., Matondang N., and Ishak A., "Inventory Optimization Model Design with Machine Learning Approach in Feed Mill Company," *Jurnal Sistem Teknik Industri*, vol. 24, no. 2, pp. 254-272, 2022. https://doi.org/10.32734/jsti.v24i2.8637

[24] Poi E. and Opara B., "Inventory Optimization and Customer Satisfaction of Petroleum Marketing Firms in Rivers State," *Journal of Accounting, Marketing and Information Technology*, vol. 8, no. 2, pp. 118-133, 2021. https://www.researchgate.net/publication/356617026

[25] Purnamasari D., Permadi V., Saefudin A., and Agusdin R., "Demand Forecasting for Improved Inventory Management in Small and Medium-Sized Businesses," *JANAPATI*, vol. 12, no. 1, pp. 56-66., 2023. https://doi.org/10.23887/janapati.v12i1.57144

[26] Qiu R., Sun Y., and Sun M., "A Distributionally Robust Optimization Approach for Multi-Product Inventory Decisions with Budget Constraint and Demand and Yield Uncertainties," *Computers and Operations Research*, vol. 126, pp. 105081, 2021. https://doi.org/10.1016/j.cor.2020.105081

[27] Sarath Kumar B. and Ramana B., "Business Intelligence and Analytics in Supply Chain Management," *Turkish Journal of Computer and Mathematics Education*, vol. 10, no. 1, pp. 666-670, 2019. https://turcomat.org/index.php/turkbilmat/article/view/14266

[28] Seyedan M., Mafakheri F., and Wang C., "Order-up-to-Level Inventory Optimization Model Using Time-Series Demand Forecasting with Ensemble Deep Learning," *Supply Chain Analytics*, vol. 3, pp. 100024, 2023. https://doi.org/10.1016/j.sca.2023.100024

[29] Shao X., Chang D., and Li M., "Optimization of Lateral Transfer Inventory of Auto Spare Parts Based on Neural Network Forecasting," *Journal of Intelligent Systems and Control*, vol. 1, no. 1, pp. 1-17, 2022. https://doi.org/10.56578/jisc010102

[30] Sihotang J., "Optimization of Inventory Ordering Decision in Retail Business Using Exponential Smoothing Approach and Decision Support System," *International Journal of Mechanical Computational and Manufacturing Research*, vol. 12, no. 2, pp. 46-52, 2023. https://doi.org/10.35335/computational.v12i2.121

[31] Stanelyte G., Inventory Optimization in Retail Network by Creating a Demand Prediction Model, Master Thesis, Vilnius Gedimino Technicus University, 2021. https://vb.vgtu.lt/object/elaba:80509157/

[32] Tang Y., Chau K., Lau Y., and Zheng Z., "Data-Intensive Inventory Forecasting with Artificial Intelligence Models for Cross-Border E-Commerce Service Automation," *Applied Science*, vol. 13, no. 5, pp. 1-13, 2023. https://doi.org/10.3390/app13053051

[33] Teerasoponpong S. and Sopadang A., "Decision Support System for Adaptive Sourcing and Inventory Management in Small and Medium-Sized Enterprises," *Robotics and Computer-Integrated Manufacturing*, vol. 73, pp. 102226,

2022. https://doi.org/10.1016/j.rcim.2021.102226

[34] Tian R., Lu M., Wang H., Wang B., and Tang Q., "IACPPO: A Deep Reinforcement Learning-Based Model for Warehouse Inventory Replenishment," *Computers and Industrial Engineering*, vol. 187, pp. 109829, 2024. https://doi.org/10.1016/j.cie.2023.109829

[35] Tirkolaee E., Mahmoodkhani J., Bourani M., and Tavakkoli-Moghaddam R., "A Self-Learning Particle Swarm Optimization for Robust Multi-Echelon Capacitated Location-Allocation-Inventory Problem," *Journal of Advanced Manufacturing Systems*, vol. 18, no. 4, pp. 677-694, 2019. https://doi.org/10.1142/S0219686719500355

[36] Valencia-Cardenas M., Diaz-Serna F., and Correa-Morales J., "Multi-Product Inventory Modeling with Demand Forecasting and Bayesian Optimization," *DYNA*, vol. 83, no. 198, pp. 235-243, 2016. https://revistas.unal.edu.co/index.php/dyna/article/view/51310/57375

[37] Waller M. and Fawcett S., "Data Science, Predictive Analytics, and Big Data: A Revolution that will Transform Supply Chain Design and Management," *Journal of Business Logistics*, vol. 34, no. 2, pp. 77-84, 2013. https://doi.org/10.1111/jbl.12010

[38] Wu G., De Carvalho-Servia, and Mowbray M., "Reinforcement Learning for Inventory Management in Multi-Echelon Supply Chains," *in Computer Aided Chemical Engineering*, vol. 52, pp. 795-800, 2023. https://doi.org/10.1016/B978-0-443-15274-0.50127-X

[39] Yuan X., *Impact of Industry 4.0 on Inventory Systems and Optimization*, IntechOpen, 2020. https://www.intechopen.com/chapters/70417

[40] Zhao Z. and Wu C., "Wheat Quantity Monitoring Methods Based on Inventory Measurement and SVR Prediction Model," *Applied Sciences*, vol. 13, no. 23, pp. 12745-12745, 2023. https://doi.org/10.3390/app132312745

Emna Bouazizi is an Assistant Professor at the University of Jeddah, College of Computer Science, Khulis Branch, KSA, since 2016. Her research interests include Quality of Service in Real-Time Systems, Big Data, Data Warehouse and related fields.

Ayman Khedr is currently a Professor with the University of Jeddah. He has been the Vice Dean of Post-Graduation and Research and the Head of the Information Systems Department, Faculty of Computers and Information Technology, Future University in Egypt. He is also a Professor with the Faculty of Computers and Information, Helwan University, Egypt. Previously, he was the General Manager of the Helwan E-Learning Center. His research interests include the Themes (Scientific) Data and Model Management, Data Science, Big Data, the IoT, E-Learning, Data Mining, Bioinformatics, and Cloud Computing. (Based on document published on 6 November 2023).

Sherin Elfaiomy is a Lecturer Assistant at the Egypt University of Informatics, Faculty of Computing and Information Sciences, Cairo, Egypt. She obtained her Master's degree in Information Systems and Artificial Intelligence field from the Arab Academy for Science and Maritime Technology and obtained her Bachelor's degree from Modern Academy in Maadi.

Mohamed Belal is a Professor of Computer Science, specializing in Computational Intelligence, at the Faculty of Computers and Artificial Intelligence, Helwan University, Egypt. He earned his Ph.D. degree from the Faculty of Engineering at Cairo University and his B.Sc., with honors, from the Computer Engineering Department at the Faculty of Engineering, Alexandria University. Professor Belal served as a Visiting Professor at the High-Performance Computing Lab (HPCL) at George Washington University (GWU), USA, from October 2002 to August 2003. He is also a consultant to the Ministry of Communications and Information Technology for AI-Based Semantic Search. His administrative roles have included serving as the Dean of the Faculty of Computers and Artificial Intelligence at Helwan University from 2012 to 2015 and at Beni-Suef University from 2016 to 2020. He was the Vice President of Beni-Suef Technological University from 2020 to 2021 and is currently the Vice President of the 6th October Technological University, Egypt.