OPERATIONS AND SUPPLY CHAIN MANAGEMENT
Special Issue on AI, Cyber-Physical Systems, and Decision Making in OSCM
Vol. 18, No. 2, 2025, pp. 161 – 173, ISSN 1979 – 3561 | EISSN 2579 - 9363

Forecasting and Multi-Objective Optimization Model for Supplier Selection and Order Allocation at Lubricants Distributor

Huu-Tin To
Department of Industrial System Engineering, Faculty of Mechanical Engineering, Ho Chi Minh University of Technology, Vietnam National University, Vietnam
Email: tohuutin13579@gmail.com

Mai-Ha Phan
Department of Industrial System Engineering, Faculty of Mechanical Engineering, Ho Chi Minh University of Technology, Vietnam National University, Vietnam
Email: ptmaiha@hcmut.edu.vn (*Corresponding Author*)

ABSTRACT

Some distributors face challenges in selecting suppliers and determining the quantity of orders to meet customer demand adequately. The root cause of this issue lies in the lack of application of accurate forecasting techniques and insufficient consideration of uncertainty factors. Therefore, this study proposes a model that incorporates forecasting techniques and accounts for the uncertainty of input parameters. First, applying and comparing ARIMA, Holt-Winter, and ANN methods to forecast future demand for datasets collected that are both trendy and seasonal. The forecasted demand is one of the input parameters for the proposed Multi-objective Supplier Selection and Order Allocation model. Use the Mixed Integer Linear Programming (MILP) method to resolve uncertainty in supplier purchasing prices and the weighted sum approach to combine the objective functions. The proposed model establishes Supplier Selection and Order Allocation planning for the distributor in the lubricant industry, resulting in a total cost decrease of 12.30% and a damaged product decrease of 5.63% within four weeks compared to the actual. Moreover, the model is capable of solving more complex scenarios, including cases with up to 15 suppliers and 10 product types over 8 weeks across various scenarios. However, the solution time is still considerable, suggesting that exploring metaheuristic optimization algorithms could be beneficial to improve the efficiency of the model in handling such large-scale problems, as well as considering additional constraints related to sustainability and other uncertainty factors to better reflect real-world conditions.

Keywords: *forecasting techniques, multi-objective, order allocation, supplier selection, uncertainty, weighted sum*

1. INTRODUCTION

Supply chain management (SCM) is crucial in manufacturing, trading, and construction enterprises. When competition in the market increases, selling and purchasing prices are more and more strictly managed. Supply chain management significantly affects the ability of an organization to expand; tiny and medium-sized firms (SMEs) are vital to the nation's economy (Odei & Hamplová, 2022), so SMEs need to start paying attention to SCM. The supply chain is the chain of activities from procuring raw materials to producing products and transporting them to the end customer. If the business can manage these activities well, it can earn high profits and outpace competitors in the industry. However, it is a big challenge for every business; procurement is also one of the significant challenges for companies regarding SCM because it makes decisions that affect businesses a lot, such as deciding on input materials, supplier selection, pricing, and order allocation (Douaioui *et al.*, 2021; Dey & Saha, 2018).

Suppliers play a role in ensuring the supply of goods and materials for business activities. In business, each enterprise often has many suppliers in charge of supplying input materials, products, or services for businesses to put into production, business, and bring to consumers. Selecting suppliers and placing efficient quantities of goods will help companies ensure the quality of products created and give them advantages in setting competitive prices compared to competitors in the same industry (Siemieniako *et al.*, 2023). Therefore, many businesses have their purchasing departments in charge of supplier selection and order allocation, but this is ineffective because companies still need to assess the capacity among suppliers, causing shortages or excess input materials, significantly affecting production.

The study is focused on the company's distribution of lubricants to car repair dealers, gasoline dealers, and retail stores. The distributor is experiencing supply chain inefficiencies characterized by supplier non-compliance during peak demand periods and elevated damage rates. The consequence is frequent stockouts, forcing the distributor to place expedited orders at higher costs. To mitigate these risks, the company aims to implement a demand forecasting solution to enable strategic supplier selection and order allocation optimization. Such a need motivated the authors to study the application of forecasting techniques to the multi-objective Supplier Selection and Order Allocation (SSAOA) model. The model not only has the capability to integrate forecasting techniques but can also address uncertainty in input parameters, enabling effective solutions when applied in practice across various fields for distributors.

This study proposes a novel approach by integrating forecasting, supplier selection, and order allocation into a unified model, addressing a critical gap in existing research. The proposed approach stands out by employing advanced machine learning techniques to handle complex time series data that exhibit both trends and seasonality, surpassing traditional methods. Previous studies have often examined these aspects separately. However, this model acknowledges their interdependence and underscores the necessity for a comprehensive decision-making process. Forecasting plays a crucial role in predicting demand, these methods such as ARIMA, Holt-Winter, and ANN are utilized to improve forecast accuracy. Supplier selection ensures the identification of optimal partners based on multiple criteria, and order allocation optimizes distribution while considering constraints such as costs, warehouse capacity, and supplier capabilities. Given the increasingly volatile and uncertain nature of supply chains, this integrated approach is essential for improving operational efficiency, mitigating risks, and enhancing supply chain adaptability.

Furthermore, incorporating these three elements into a single model enables synchronized decision-making, optimizing inventory costs, ordering costs, and operational expenses. The model integrates real-world constraints such as warehouse conditions, distributor budgets, and supplier production capacities, making it particularly applicable and reflective of practical operational scenarios. Despite the growing complexity of modern supply networks, very few studies have explored this holistic perspective, making our research a significant contribution to the field.

This study consists of 5 sections. Section 1 is the introduction, section 2 is the literature review and the research methodology. section 3 discuss the Model Application with the Demand forecasting model; the Supplier Selection and Order Allocation model, respectively. Section 4 presents the results and discussions, and finally, the conclusion is mentioned in section 5.

2. LITERATURE REVIEW AND METHODOLOGY

2.1 Literature Review

In supply chain management, not only forecasting but also supplier selection and order allocation are critical research issues that are frequently examined. Numerous studies have investigated these issues individually in practical applications with distributors across various industries. Regarding forecasting techniques, Rumbe *et al.* (2024) compared the accuracy of the Holt-Winters method and Artificial Neural Network (ANN) under seasonal influences at a tent manufacturing company. Their findings highlight the ability of ANN to capture complex demand patterns, offering a promising approach for improving forecasting accuracy. A comparison between the Autoregressive Integrated Moving Average model with Exogenous Inputs (ARIMAX) and machine learning (ML) models was conducted by Gonçalves *et al.* (2021) in the assembly industry, showing that ARIMAX predicts demand signals better in the early life-cycle, while ML models outperform it in later stages. Rachidi *et al.* (2021) assesses the global production trends and supply sustainability of cobalt, a key energy-transition metal (ETM), using various forecasting methods. Results from the ARIMA and Holt's models indicate a short-term linear increase in cobalt production, while the Hubbert model predicts a decline starting in the late 2010s. Several studies highlight the popularity of traditional methods such as Holt-Winters, ARIMA, and some machine learning approaches in supply chain management. However, each method has its characteristics and suitability for different data patterns, so careful consideration is needed before selection.

Supplier selection (SS) is vital to every business's supply chain management strategy. Supplier selection strategy involves evaluating and selecting potential suppliers and establishing contracts with those suppliers. Agrawal *et al.* (2021) used the FTOPSIS technique to select the primary ingredients for laundry powder (surfactants) based on the criteria of timely arrival, expense, material quality, consistency, etc. Supplier selection problems are also considered uncertainty factors to be able to solve these problems close to reality. Büyüközkan & Göçer (2017) have an intuitionistic fuzzy axiomatic design (IFAD) and intuitionistic fuzzy analytic hierarchy process (IFAHP) to select suppliers under the uncertainty of the environment at a sporting company in Tuckey. For the consideration of project life cycles and demand uncertainty in the issue of multi-criteria supplier selection between multiple projects and different phases, Zheng *et al.* (2021) have introduced the multi-stage mixed-integer stochastic programming methodology to be able to evaluate the right supplier for the project. De Oliveira *et al.* (2023) simulated eight scenarios of the influence of criteria and alternatives in supplier selection in a service company, two techniques: Extended Hesitant Fuzzy Linguistic VIKOR (EHFLVIKOR) and Hesitant Fuzzy Linguistic Term Set VIKOR (HFLT VIKOR) were respectively applied to solve the problem. The comparative results showed that EHFLVIKOR was better suited to that company in the SS. Uncertainty in decision-makers judgment and related supplier parameters integrated with inventory management was addressed by Saputro *et al.* (2023) through the FAHP and interval TOPSIS methodology. During the Covid pandemic, Joy *et al.* (2023) developed the AHP methodology for potential supplier selection for the Glove industry.

In recent years, decision-makers have also considered sustainability or Green Supplier selection to limit negative impacts on ecological, economic, and social aspects. This is a massive challenge for businesses because it is necessary to ensure the requirements when choosing suppliers while ensuring sustainability or "green". The proposed approach in the paper of Vaezi *et al.* (2024) integrates sustainability and blockchain-related factors into the decision-making process. The model uses the Best Worst Method (BWM) to calculate the importance weights of supplier selection criteria and applies the MULTIMOORA method to evaluate suppliers, along with a bi-objective Mixed Integer Linear Programming (MILP) model to optimize order allocation. Research by Liou *et al.* (2021) applied machine learning methods to select green suppliers; the methods used are support vector machine (SVM), fuzzy best worst method (FBWM), and FTOPSIS. The three methods reduced 25 criteria to 13 criteria by analyzing the historical data of a multinational electronics manufacturer. Hajiaghaei-Keshteli *et al.* (2023)

combined environmental performance factors and traditional criteria to reduce environmental impacts while keeping lower costs; the authors came up with the Pythagorean Fuzzy TOPSIS (PF-TOPSIS) technique to be able to select green suppliers based on requirements.

Machine learning is also used and combined with the Dynamic decision support system (DSS) to select sustainable suppliers at petrochemical companies when considering the strategies of circular supply chains in the research paper of Alavi *et al.* (2021). Recognizing that Chemical Industry is a field of "high-risk, high-pollution, and high-efficiency," Wu *et al.* (2021) came up with a solution to apply Fuzzy Grey Relational Analysis (FGRA), Failure Mode and Effects Analysis (FMEA), Cloud computing-entropy weight method (EWM), and Decision-making Trial and Evaluation Laboratory (DEMATEL) to select sustainable suppliers based on social, environmental and economic dimensions. The model gives positive results when applied in a Chinese petrochemical company. Decision-makers consider supplier selection to broaden the issue. A crucial element in reducing manufacturing costs is figuring out the right order quantity in addition to assessing possible suppliers. Numerous writers have examined the topic of supplier selection and order allocation due to its significance. In order to address the SSAOA problem, Firouzi & Jadidi (2021) investigated the Multi-Objective Mixed Integer Linear Programming (MOMILP) technique, which transforms the multi-objective model into a single-objective model under uncertain situations. Jia *et al.* (2020) studied the distributionally robust goal programming technique (DRGPM) in a steel firm to propose an optimal solution when taking sustainability and uncertainty into account. Mixed integer nonlinear programming models were introduced by Ventura *et al.* (2021) to choose possible suppliers and figure out order quantities. Deterministic approaches were taken into consideration by Ventura *et al.* (2021). Kaur & Singh (2021) proposed a multi-stage hybrid method that combines supplier segmentation, selection, and order allocation while accounting for risks and interruptions. This method solves SSAOA by combining the Mixed Integer Program (MIP) and the Fuzzy Analytical Hierarchical Process TOPSIS. Ali *et al.* (2023) also consider sustainable issues when selecting suppliers and allocating orders, authoring research on the Fuzzy analytical hierarchy process, multi-objective linear programming (MOLP), and Fuzzy compromise programming to provide the best solution. The Best-worst method (BWM) is often used because it is easy to use and gives highly reliable results. Hosseini *et al.* (2022) studied BWM, Evidential reasoning (ER), and Dynamic programming (DP) for supplier evaluation. Shidpour *et al.* (2023) also used BWM provider assessments based on Corporate Social Responsibility (CSR). Meanwhile, Nayeri *et al.* (2023) developed the Stochastic Fuzzy Best Worst Method (SFBWM) to increase accuracy in the provider evaluation process; Fuzzy Robust Stochastic (FRS) was also studied to provide solutions to uncertainty in a healthcare system. Ordu *et al.* (2021) address healthcare system challenges by integrating hybrid models using ARIMA, Exponential Smoothing (ES), and Stepwise Linear Regression (SLR) for forecasting. Although not centered on supplier selection, these methodologies align well with concepts in order allocation by enhancing resource management through demand forecasting and linear optimization by using forecasting-simulation-optimization (FSO) approach, thereby supporting strategic planning. An advanced approach to enhancing guaranteed delivery models in online advertising is presented by Zhang *et al.* (2020). The study tackles industry challenges using spatial-temporal tensor factorization (ST-TF), CNN, and LSTM for precise forecasting. Heuristic-High Water Mark (H-HWM) and FIFO are employed for efficient order allocation. Implemented in Tencent's platform, this integrated system effectively manages resources and maximizes revenue in the complex digital advertising landscape. However, while these reviews combine forecasting with order allocation models, they do not consider the domain of supplier selection.

Islam *et al.* (2021) were the first to apply forecasting techniques to the SSAOA model. They compared the performance of ARIMA, Holt, and the Relational Regressor Chain (RRC) method, finding that the RRC method provided the most accurate results. They then used the forecasted results as one of the parameters to solve the SSAOA problem. In 2024, Jafari-Raddani *et al.* (2024) introduced the model sustainable supplier selection and order allocation considering demand forecasting, employing a rigorous three-stage methodology. Specifically, the initial stage utilized fuzzy AHP and fuzzy TOPSIS for ranking sustainable suppliers, followed by the application of polynomial regression (PR) for demand forecasting in the second stage. Notably, this model did not undertake a comparative analysis of various forecasting methods, thus lacking a comprehensive assessment of forecasting accuracy. Furthermore, the model did not incorporate considerations of uncertainty. These gaps underscore the distinct contribution of this study, which integrates both forecast accuracy evaluation and the management of uncertainty within the framework of supplier selection and order allocation. Building on model of Islam *et al.* (2021), the study introduces a more advanced forecasting approach, integrating Holt-Winter, ARIMA and Artificial Neural Networks (ANN) to better capture the trend and seasonal nature of the time series data. Unlike their study, which focused solely on improving forecasting accuracy, this appoach extends the SSAOA model by incorporating uncertainty in supplier purchasing costs, addressing real-world price fluctuations that impact decision-making. Additionally, we enhance the model's practical applicability by introducing new constraints, such as warehouse capacity limitations and supplier capability constraints, ensuring that the supplier selection and order allocation decisions are not only cost-effective but also feasible under real operational conditions. These improvements make the proposed model more robust and adaptable to the dynamic challenges faced by the lubricant industry, bridging the gap between theoretical optimization and practical supply chain management. The reviewed literature is summarized in Table 1, which outlines the methods used to address the issues of supplier selection and order allocation across 23 relevant studies. It is evident that few studies have explored the integration of demand forecasting, supplier selection, and order allocation in supply chains. Given the current complexities, there is a clear need for research that combines these issues to effectively address business challenges on time. In Table 1, uncertainty is primarily considered within the criteria of supplier, where many studies incorporate purchasing cost variability, supplier reliability, production capacity fluctuations, and quality variations as key decision-making factors. These uncertainties play a crucial role in evaluating suppliers and determining order allocation, directly influencing the effectiveness of supply chain optimization.

Table 1 Summary of SSAOA literature review.

Authors	Forecasting	Supplier Selection	Order Allocation	Uncertainty	Application Field
Büyüközkan & Göçer (2017)		IFAD, IFAHP		Fuzzy	Sporting company
Jia *et al.* (2020)		DRGPM	DRGPM	Robust	Steel company
Zhang *et al.* (2020)	ST-TF, CNN, LSTM		H-HWM, FIFO		Advertising company
Islam *et al.* (2021)	Holt, ARIMA, RRC	SMILP	SMILP	Stochastic	Food industry
Kaur & Singh (2021)		FAHP-TOPSIS	MIP	Fuzzy	Automobile company
Ordu *et al.* (2021)	ARIMA, ES, SLR		FSO		Healthcare industry
Ventura *et al.* (2021)		MINPM	MINPM	Deterministic	Retail industry
Wu *et al.* (2021)		FGRA, FMEA, EWM, DEMATEL		Fuzzy	Chemical industry
Firouzi & Jadidi (2021)		MOMILP	MOMILP	Fuzzy	Generated data
Agrawal *et al.* (2021)		FTOPSIS			Surfactants company
Alavi *et al.* (2021)		DSS			Petrochemical company
Zheng *et al.* (2021)		M-MISPP		Stochastic	Auto industry
Liou *et al.* (2021)		SVM, FBWM, FTOPSIS			electronics manufacturer
Hosseini *et al.* (2022)		BWM, ER	DP	Stochastic	Composite company
De Oliveira *et al.* (2023)		PDHFLVIKOR, EHFLVIKO		Fuzzy	Service company
Joy *et al.* (2023)		AHP			Glove industry
Saputro *et al.* (2023)		FAHP and interval TOPSIS		Fuzzy	Manufacturing industry
Hajiaghaei-Keshteli *et al.* (2023)		PF-TOPSIS			Food industry
Shidpour *et al.* (2023)		BWM	MOO, BWM	Linguistic terms and fuzzy	Engine/industrial oil industry
Nayeri *et al.* (2023)		SFBWM	MOM	FRS	Healthcare system
Ali *et al.* (2023)		FAHP	MOLP, FCP	Fuzzy	Textile industry
Jafari-Raddani *et al.* (2024)	PR	fuzzy AHP, fuzzy TOPSIS	SMILP		Food and beverage
Vaezi *et al.* (2024)		BWM, MULTIMOORA	MILP		Blockchain company
Proposed model	**Holt-Winter, ARIMA, ANN**	**MILP**	**MILP**	**Stochastic**	**Lubricant industry**

2.2 Methodology

A proposed model tackles the identified problem by integrating forecasting techniques while accounting for the uncertainty of purchasing prices to minimize total costs and damage rates. This model is developed based on the work of Islam *et al.* (2021). The proposed model consists of 2 phases (Figure 1). In phase 1, forecasting demand using techniques based on input data is a time series. According to the review by Aamer *et al.* (2020), spanning the decade from 2010 to 2019, artificial neural networks (ANN) and neural networks (NN) constituted 48% of the algorithms applied in machine learning for demand forecasting. This trend underscores the increasing application of neural network algorithms within the context of supply chain management. Furthermore, Seyedan & Mafakheri (2020) conducted a comprehensive review over 15 years, establishing ARIMA as the predominant traditional method. Consequently, both ANN and ARIMA methodologies have been chosen for demand forecasting in this study. Holt-Winter was also selected for comparison alongside ARIMA and ANN due to its popularity and ease of use, as well as its ability to improve accuracy in short-term and seasonal forecasts. In addition, ARIMA has the great advantage of handling stationary and less seasonal time series while providing reliable long-term forecasting, while ANN is highly appreciated in forecasting problems because of its ability to optimize forecasting models with large and non-linear data. The comparison of these three methods will help to have a multi-dimensional perspective, combining the advantages of each method to achieve more accurate and comprehensive forecasting results in research. For ARIMA, after analyzing the data, if it is found to be non-stationary, differencing will be applied before proceeding with the method. For ANN, various network structures need to be tested to identify the most suitable model for the given dataset. The evaluation is divided into three cases:

Case 1: Testing the Rectified Linear Unit (ReLU) and Sigmoid activation functions.

Case 2: Testing different numbers of neurons in a single hidden layer (SLP).

Case 3: Testing different numbers of neurons in two hidden layers (MLP).

This study leverages advanced modifications to both ARIMA and ANN models, elevating their effectiveness beyond basic methodologies. The ARIMA model incorporates a seasonal component (SARIMA) to adeptly capture and forecast seasonal patterns in the data, which significantly enhances forecast accuracy compared to traditional ARIMA models that do not account for seasonality. ANN model is designed with two hidden layers, providing a deeper architecture capable of capturing complex nonlinear relationships in the data. This configuration offers a substantial improvement in predictive performance over basic ANN models with a single hidden layer. These sophisticated adaptations underscore the robustness of this proposed approach in accurately forecasting demand in variable and complex supply chain environments.

After comparison of accuracy, choose the method that best suits the dataset; the forecast result of phase 1 is one of the input parameters of phase 2.

In phase 2, the Supplier Selection and Order Allocation model will be developed, incorporating the parameter uncertainty factor; the proposed model is a multi-objective framework aimed at minimizing total cost and product damage, employing the weighted sum method to integrate the two objectives. After assessing and analyzing the model's effectiveness in addressing the supplier's issues, a comparative analysis is conducted between the model, accounting for the parameter uncertainty scenarios, and the deterministic model. This study implements a comprehensive stochastic programming model that integrates multiple scenarios within a single model, rather than addressing each scenario separately. This approach allows us to manage uncertainties, optimize overall costs, and ensure compliance with strict budget constraints specific to supplier selection. By considering all scenarios, the solution is robust and adaptable to various real-world situations, such as price fluctuations or sudden spikes in demand (Miranda & Saucedo, 2021; Cheraghalipour & Farsad, 2018). This integrated model provides a comprehensive view of the supplier selection and order allocation, which is crucial for distributors operating under limited budgets. It aids in improving negotiation terms with suppliers through framework agreements, where purchasing a reserved capacity over a certain period can result in more favorable pricing (Nasiri *et al.*, 2018). Additionally, practical comparisons are conducted to evaluate the real-world applicability and effectiveness of the proposed model.

To the best of our knowledge, since the first study in 2021, there have been very few related studies on this topic. Most of the existing literature focuses solely on either forecasting techniques or the SSAOA model, without addressing the integration of both aspects. These studies tend to treat forecasting and optimization separately, overlooking the potential benefits of combining accurate demand forecasting with multi-objective optimization for supplier selection and order allocation in real-world applications. The proposed model introduces significant innovations compared to previous studies by comparing traditional methods with machine learning approaches to handle complex time series data that exhibit both trend and seasonality. Additionally, the SSAOA model incorporates various constraints, such as warehouse conditions, distributor budget, and supplier production capacity, to better reflect real-world scenarios.

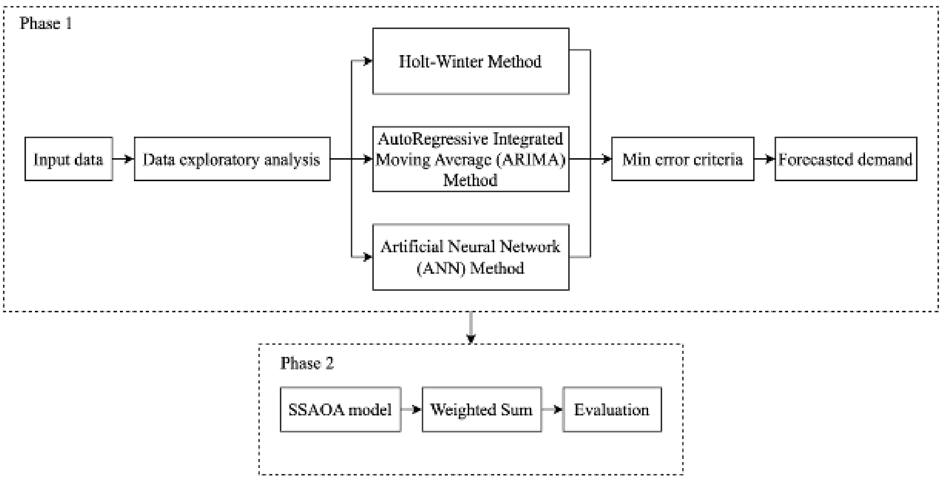

Figure 1 Solution approach to solve the problem.

3. MODEL APPLICATION

3.1 Demand Forecasting Model

The study was conducted at a lubricant distributor for vehicles and engines in the Mekong Delta region, Vietnam This distributor offers 89 product, classified into three product types based on size: Bottle, Can, and Bucket. Since this study does not account for brand factors, the average data for each type – Bucket, Bottle, and Can – will be used as the foundation for analysis. This approach enables a more focused application of forecasting techniques, allowing us to capture the general trends and patterns across product categories that ensure a more streamlined and efficient application of forecasting models, without being influenced by brand differences. Each supplier capable of providing these three product categories to the distributor has differing production capacities, selling prices, and transportation costs. Additionally, products may be damaged during transportation, storage, or due to temperature conditions, leading to unmet customer demand and compensation issues. Therefore, the damage rate across suppliers is a critical criterion in the supplier selection process. The goal of the supplier selection and order allocation plan is to minimize not only total costs but also the damage rate. Customer demand for the three product categories fluctuates weekly, so data was collected weekly over four years, from 2020 to 2024. Moreover, product prices are subject to change due to economic factors, making purchasing costs an uncertain parameter, while other factors are assumed to be fixed.

Phase 1 presents the first part of the solution model, which is the application of forecasting techniques to manage product demand. The forecasting output is information about product demand in the first 4 weeks of 2024, serving as input data for the subsequent steps of the solution model. Use error criteria to evaluate and compare the forecasting techniques, this study uses four coefficients: *MAE, MSE, RMSE,* and *MAPE* to evaluate forecasting models.

3.1.1 Data Exploratory Analysis

Analyze data characteristics to be able to choose appropriate forecasting methods (Micus *et al.*, 2023), but to be able to evaluate the performance of the proposed forecasting model and compare the forecast results of each type of method, "test" data is required. Therefore, it is necessary to divide the data into two parts. Part 1 includes the first 80% of data used for analysis and calculation and as a basis for building the model. Part 2 consists of the last 20% of data used to evaluate the model and compare forecast results. If the training dataset does not follow the normal distribution, it is necessary to conduct a Box-cox transformation to let the data follow the normal distribution. Table 2 shows the normal distribution test results of Bucket, Bottle, and Can's train dataset.

Table 2 Normal distribution test results of bucket, bottle, and can's train dataset.

Product	Conclusion
Bucket	Follow a normal distribution
Bottle	Follow a normal distribution
Can	Do not follow a normal distribution

Since Can's train data does not follow the normal distribution, it is necessary to use the Box-cox method to transform the data, using Minitab software with the Box-cox transformation function to conduct data conversion. From the results after Box-cox transformation, it can be seen that the value of λ is 0.5, which means the value has been transformed by 0.5, expressed by the formula (1) which y' is the value after the transformation, y is the actual value and λ is the parameter of the Box-Cox transformation.

$$y' = y^{\lambda} = y^{0.5} \quad (1)$$

The data patterns of a data series over time will include trending, seasonality, and stationary; it is possible to determine the data patterns of the data by observing the autocorrelations of that data (Guijarro *et al.*, 2015), the data patterns, following analysis, are presented in Table 3.

Table 3 Data patterns of the training datasets for bucket, bottle, and can.

Product	Trend	Seasonal
Bucket	Increasing trend	Seasonal length = 12
Bottle	Increasing trend	Seasonal length = 6
Can	Increasing trend	No seasonal

Subsequently, the stationarity of the data is assessed using the Augmented Dickey-Fuller test for each of the three training datasets corresponding to the three product types. For the Augmented Dickey-Fuller test, the null hypothesis states that the data are non-stationary. Therefore, comparing the P-values is necessary to determine the stationarity of the data; Table 4 is the result of the Augmented Dickey-Fuller test.

Table 4 Augmented dickey-fuller test for the training datasets of bucket, bottle, and can.

Null hypothesis: Data are non-stationary
Alternative hypothesis: Data are stationary

Product	Test Statistic	P-Value	Conclusion
Bucket	-3.79124	0.003	Stationary
Bottle	-2.44532	0.129	No stationary
Can	-3.60229	0.006	Stationary

3.1.2 The Holt-Winter Method

To forecast using the Holt-Winter method, determining the data pattern of the dataset is necessary to appropriately choose between the Holt or Winter method. The Winter method is selected for the datasets of Bucket and Bottle, while the Holt method is chosen for the Can dataset. For the Winter method, both additive and multiplicative models are used to select the most suitable model (Omar *et al.*, 2021; Tratar *et al.*, 2016). Additionally, the weights for level, trend (if any), and seasonal (if any) are visually determined, so multiple scenarios will be compared to determine the coefficients as well as the best model for each product dataset. Table 5 is the result after comparing Additive and Multiplicative models to come up with a suitable model for the three product types.

Table 5 The suitable holt–winter model and error criteria for the three product types.

Product	Model	α	β	γ	MAE	MSE	RMSE	MAPE(%)
Bucket	Winter - Additive	0.05	0.05	0.05	6.874	71.515	8.457	16.698
Bottle	Winter - Multiplicative	0.40	0.40	0.40	58.149	4764.481	69.025	11.869
Can	Holt	0.20	0.20	0.00	2.776	14.668	3.830	10.725

3.1.3 The Arima Method

One of the essential conditions for implementing the ARIMA forecasting method is that the dataset must be stationary. However, the training dataset for the Bottle product is not stationary. Therefore, it is necessary to differentiate the data to ensure the stationarity of the dataset. After differencing the dataset, it is noted that the data has become stationary, and further differencing is unnecessary.

To determine the auto regressive (p) and moving average (q) parameters, one needs to observe the ACF and

PACF plots. This method requires experience and sometimes may not accurately determine the best ARIMA model for forecasting. If the dataset exhibits seasonality, it becomes even more challenging to determine the parameters for a seasonal ARIMA (SARIMA) model, including seasonal *AR*, seasonal *MA*, *D*, and seasonal Length, solely based on ACF and PACF plots (Wang *et al.*, 2012; Putro *et al.*, 2016). Therefore, in this research, the Akaike Information Criteria (AIC) is used to identify the best ARIMA model. AIC is a measure of the relative quality of a statistical model for a given dataset. It is considered a useful tool for model selection, as well as determining the lags of an ARIMA model. The better an ARIMA model fits the dataset, the lower the AIC value. The model results and error criteria for Bucket, Bottle, and Can shown in Table 6.

Table 6 The model results and error criteria for bucket, bottle, and can.

Product	AIC	ARIMA	MAE	MSE	RMSE	MAPE(%)
Bucket	523.9	(2,0,0)(0,1,1)	6.131	50.827	7.129	14.879
Bottle	866.7	(0,1,0)(0,1,1)	51.207	5872.005	76.629	9.369
Can	176.8	(1,0,1)	2.451	10.097	3.178	10.224

3.1.4 The ANN Method

Artificial Neural Network (ANN) machine learning is a technique to predict the datasets related to the three different product categories. The input layer, hidden layers, and output layer are the three primary components of an ANN's overall structure (Rumbe *et al.*, 2024; Dercole *et al.*, 2020). The performance of the model with a single hidden layer (Single-layer Perceptron, or SLP) and two hidden layers (Multilayer Perceptron, or MLP) is compared and evaluated in this research. To choose the best activation function for the hidden layers, two activation functions: Sigmoid and Rectified Linear Unit (ReLU) are assessed and compared (Oostwal *et al.*, 2021). This ANN model is appropriate for regression issues since it does not employ a linear function as the activation function for the output layer. One popular non-linear activation function for hidden layers is the ReLU activation function. The formula (2) defines the ReLU activation function. If the input value is negative, ReLU sets the output value to 0; if the input value *x* is positive, it keeps the value unchanged. The Sigmoid activation function is defined in formula (3).

$$f(x) = max\,(0, x) \tag{2}$$

$$(x) = \frac{1}{1 + e^{-x}} \tag{3}$$

In equations (2) and (3), *x* is the input of the function and e is Euler's number, approximately equal to 2.71828. Experiments to determine the optimal design for the network are used to determine the proper number of neurons to include in the hidden layer (Tealab *et al.*, 2017; Pektaş *et al.*, 2013).

a. Testing the Rectified Linear Unit (ReLU) and Sigmoid activation functions

Perform training of the model sequentially with two activation functions, Rectified Linear Unit (ReLU) and Sigmoid, in the case of having an equal number of neurons in the hidden layer on the Bucket's training dataset with 200 epochs. However, since the AME algorithm initializes some random parameters during the weight computation process, each code run will yield slightly different results. Therefore, for each case of changing the number of neurons, the code will be executed ten times, and the average value of the error criteria will be calculated. The average results of MAE, MSE, RMSE, and MAPE for the Bucket's test dataset are shown in Table 7.

Table 7 The average results of MAE, MSE, RMSE, and MAPE for the bucket's test dataset.

Neuron in SLP	Functions	MAE	MSE	RMSE	MAPE(%)
15	ReLU	4.01	25.35	5.03	15.22
	Sigmoid	26.41	723.80	26.90	66.18
20	ReLU	4.50	28.49	5.34	15.72
	Sigmoid	25.4	671.67	25.92	63.59
25	ReLU	3.72	21.06	4.59	15.48
	Sigmoid	21.77	500.43	22.37	54.24
30	ReLU	3.75	22.82	4.78	15.77
	Sigmoid	19.85	420.18	20.50	49.27
100	ReLU	3.62	22.84	4.78	15.73
	Sigmoid	7.48	73.96	8.60	19.34

Table 8 The average MAE, MSE, RMSE, and MAPE values of the Bucket's test dataset when varying the number of neurons in the SLP.

Neuron in SLP	MAE	MSE	RMSE	MAPE(%)
5	5.38	42.77	6.54	16.41
9	4.34	27.19	5.21	15.54
10	4.20	26.39	5.14	15.71
15	4.01	25.35	5.03	15.22
20	4.50	28.49	5.34	15.72
25	3.72	21.06	4.59	15.48
30	3.75	22.82	4.78	15.77
31	3.77	22.94	4.79	15.78
32	3.92	22.36	4.73	15.76
32	3.92	22.36	4.73	15.76

From Table 7, it can be observed that for the collected dataset, the ANN model using the ReLU activation function produces more accurate prediction results than the Sigmoid activation function when the number of neurons is equal in the hidden layer. As the number of neurons increases, the accuracy of the Sigmoid activation function also increases, but it may lead to overfitting, longer training times, increased costs, and difficulty in optimization. Therefore, choosing the ReLU activation function for the ANN model is preferable.

b. Testing the number of neurons in a single hidden layer (SLP)

Train the model based on the Bucket's training dataset, then predict the results on the test dataset to calculate the error criteria. Use these error criteria to compare and evaluate the performance of different numbers of neurons if there is only one hidden layer in the network structure. Table 8 summarizes the average MAE, MSE, RMSE, and MAPE values of the Bucket's test dataset when varying the number of neurons in the SLP. When using SLP, the optimal number of neurons in the hidden layer is 25 for predicting the Bucket dataset.

c. Testing the number of neurons in two hidden layers (MLP)

Proceeding similarly to part a testing SLP, after training the ANN model with the training dataset, calculate the average error criteria values based on the test dataset. Evaluate the error criteria to select the appropriate number of neurons for the dataset. Table 9 presents the average error criteria values based on the Bucket's test dataset when varying the number of neurons in two hidden layers. Although there is not much difference in the Error criteria across different numbers of neurons, it can be observed that having 25 neurons in the first hidden layer and ten neurons in the second hidden layer yields the best results. Therefore, the Rectified Linear Unit (ReLU) activation function with 25 neurons in the first hidden layer and 10 neurons in the second hidden layer will provide the best prediction for the Bucket dataset.

From Table 10, it can be observed that using the ANN method for forecasting the datasets of Bucket and Bottle yields the highest accuracy. In contrast, the dataset Can benefit more from the ARIMA method. This difference can be attributed to the Can dataset having fewer random fluctuations and noise, making the ARIMA method more suitable for forecasting as it requires less training data and is less prone to overfitting. After analyzing and selecting the most efficient method for each dataset, the next step is to forecast the demand for the next four weeks. The forecasted results will be parameters when building and running the Supplier Selection and Order Allocation (SSAOA) model.

Table 9 The average MAE, MSE, RMSE, and MAPE values based on the bucket's test dataset when varying the number of neurons in two hidden layers.

1st hidden layers	2nd hidden layers	MAE	MSE	RMSE	MAPE(%)
5	5	4.02	23.01	4.80	15.60
10	5	3.75	22.53	4.75	15.69
15	5	3.73	22.02	4.69	15.38
15	10	3.77	21.07	4.59	15.60
20	5	3.79	20.35	4.51	15.53
20	10	3.70	23.57	4.85	15.60
20	15	3.77	23.29	4.83	15.39
25	5	3.71	24.91	4.99	15.71
25	**10**	**3.65**	**22.09**	**4.70**	**15.37**
25	15	3.70	22.86	4.78	15.37
25	20	3.85	21.03	4.59	15.66

Table 10 The error criteria result for the test datasets of bucket, bottle, and can.

Product	Method	MAE	MSE	RMSE	MAPE(%)
Bucket	ARIMA	6.13	50.83	7.13	14.88
	Winter - Additive	6.87	71.51	8.46	16.70
	ANN	**3.65**	**22.09**	**4.70**	**14.37**
Bottle	ARIMA	51.21	5872.00	76.63	9.37
	Winter - Multiplicative	58.15	4764.48	69.03	11.87
	ANN	**46.13**	**3222.99**	**56.77**	**9.14**
Can	**ARIMA**	**2.45**	**10.10**	**3.18**	**10.22**
	Holt	2.78	14.67	3.83	10.72
	ANN	2.95	11.97	3.46	12.33

Let t_1, t_2, t_3, t_4 are the next four weeks following the last week (in 2024) in the dataset. Forecasting is conducted using the ANN method for the Bucket and Bottle datasets, and ARIMA for the Can dataset, resulting in the outcomes presented in Table 11, these values serve as input parameters for the FD_{pt} in the SSAOA framework conducted in Section 3.2.

Table 11 Demand forecasting for three product types in the next four weeks.

Product	Method	t_1	t_2	t_3	t_4
Bucket	ANN	412.669	415.951	425.256	423.584
Bottle	ANN	5310.808	5181.854	5139.035	5104.421
Can	ARIMA	278.718	276.309	276.078	263.948

3.2 Supplier Selection and Order Allocation Model

The model builds on the distributor's current requirements, including optimizing purchasing and ordering costs, inventory costs, and transportation costs, and considering damage rate options among suppliers. Thus, to meet the criteria, it is necessary to build a multi-objective model based on a Mixed-Integer Linear Programming method with the required input data sets for the Supplier Selection and Order Allocation (SSAOA) model, including forecasted customer needs; cost of ordering, purchasing, transportation, and inventory; inventory volume, inventory capacity, the production capacity of suppliers along with some other related parameters.

Set:

Suppliers, $S = \{1 \ldots s \ldots S\}$
Products, $P = \{1 \ldots p \ldots P\}$
Time periods, $T = \{1 \ldots t \ldots T\}$
Scenarios, $K = \{1 \ldots k \ldots K\}$

Parameters

The following parameters are the input parameters used to run the model, collected from the available data of the distributor or supplier or calculated from other data. Some parameters change based on each type of product *p*, supplier *s,* or period *t,* especially Purchasing Cost which is a parameter that changes according to scenario *k*.

O_{st} : Ordering cost of supplier *s* in time period *t*
C_{pstk} : Purchasing price of product *p* from supplier *s* in time period *t* in scenario *k*
H_p : Holding cost of product *p*
T_{st} : Transportation cost in each distance unit from supplier *s* in time period *t*
FD_{pt} : Forecasting demand of product *p* in time period *t*
DR_{ps} : Damage rate of product *p* from supplier *s*
CP_{ps} : Capacity to produce product *p* of supplier *s*
Z_k : Probability of scenario *k*, the value ranges from 0 to 1
B_t : The budget to maintain procurement and inventory activities in time period *t*
WH_t : Warehouse capacity in time period *t*
V_p : Volume of product *p*
DT_s : Distance from supplier *s* to the Business

Decision Variables

X_{pstk} : Order quantity of product *p* from supplier *s* in time period *t* in scenario *k*
Y_{st} : 1 if purchase any product from supplier *s* in time period *t*; 0 otherwise
h_{ptk} : Level of inventory of product *p* at the end of time period *t* in scenario *k*

Objective Function

$$min\ y_1 = \sum_p \sum_s \sum_t \sum_k Z_k X_{pstk} C_{pstk} + \sum_s \sum_t O_{st} Y_{st} + \sum_p \sum_s \sum_t \sum_k Z_k X_{pstk} T_{st} DT_s + \sum_p \sum_t \sum_k Z_k h_{ptk} H_p \quad (4)$$

$$min\ y_2 = \sum_p \sum_s \sum_t \sum_k Z_k X_{pstk} DR_{ps} \quad (5)$$

Constraints

$$\sum_s X_{pstk}(1 - DR_{ps}) + h_{p(t-1)k} = h_{ptk} \ \ \forall p, \quad (6)$$

$$\sum_p h_{ptk} V_p \le WH_t \ \ \forall t, k \quad (7)$$

$$X_{pstk} \le CP_{ps} Y_{st} \ \ \forall p, s, t, k \quad (8)$$

$$\sum_p \sum_s X_{pstk} C_{pstk} + \sum_s O_{st} Y_{st} + \sum_p \sum_s X_{psti} + \sum_p h_{ptk} H_p \le B_t \ \ \forall t, k \quad (9)$$

$$X_{pstk} \ge 0 \ \ \forall p, s, t, k \quad (10)$$

$$h_{ptk} \ge 0 \ \ \forall p, t, k \quad (11)$$

$$Y_{st} = 0 \text{ or } 1 \ \ \forall p, s, t \quad (12)$$

The model has two objective functions expressed through equations (4) and (5). In particular, (4) represents the goal of minimizing costs from the ordering process to inventory including four types of costs: ordering cost (when the distributor orders from each supplier at a time *t* has to pay a cost to that supplier), purchasing cost (the purchase price from the supplier varies according to the scenarios), transportation cost (This cost depends on the quantities of products when ordering at each supplier, distance, and cost per unit distance of that supplier) and inventory cost. The target of minimum supplier-damaged products during selection is shown in (5).

Model constraints are expressed from (6) to (12) considering inventory, production capacity, order cost, and value type of decision variable. (6) Shows that the inventory value at time *t* depends on the inventory at time *t* – 1, the order quantity that was not damaged, and the forecasted demand. (7) warehouse capacity constraints. The quantity ordered by each supplier must meet its production capacity or be considered as the maximum supply per order of that supplier, as shown in (8). (9) Ordering, purchasing, transportation, and holding costs must be less than the budget. Constraints (10) and (11) stipulate that order quantities and levels of inventory must be non-negative integers, and (12) bind binary variables for supplier selection. After establishing all the constraints for the model, the authors will use the weighted sum technique to solve the SSAOA multi-objective problem detailed in the following sections.

$$min\ y_3 = w_1\Big(\sum_p \sum_s \sum_t \sum_k Z_k X_{pstk} C_{pstk} + \sum_s \sum_t O_{st} Y_{st} + \sum_p \sum_s \sum_t \sum_k Z_k X_{pstk} T_{st} DT_s + \sum_p \sum_t \sum_k Z_k h_{ptk} H_p\Big) + w_2 DC_p \sum_p \sum_s \sum_t \sum_k Z_k X_{pstk} DR_{ps} \quad (13)$$

However, the two objective functions γ_1 and γ_2 are not the same unit, so it is impossible to use the weighted sum approach; the two objective functions must be converted to the same unit. Therefore, it is necessary to multiply the

Objective γ_2 by DC_p (DC_p is the damage cost per damaged product p). At this point, both objective functions have the same unit $, which can use the weighted sum approach. From there, the new objective function (13) is calculated.

4. RESULT AND DISCUSSION

The parameter purchasing price (C_{pstk}) is uncertainty, and each scenario of this fluctuates by 10%. Scenario 1 is the main scenario when the purchasing price value does not change; the purchasing price of suppliers in scenarios 2 and 3 changes randomly by 10% of the original value. The probability of the main scenario accounted for the highest ratio of 0.5; the other two ratios were 0.3 and 0.2, respectively. Purchasing price and probability values are shown in Table 12. Additionally, to evaluate the effectiveness of the model, this study focuses only on five suppliers from Southern Vietnam with similar geographic characteristics, excluding foreign suppliers. The analysis uses the average data from the three product groups: Bucket, Bottle, and Can.

Use the weighted sum approach to solve multi-objective problems with the weight coefficient ($w_1 = 0.75$; $w_2 = 0.25$). After utilizing IBM ILOG CPLEX Optimization Studio 22.1.1 to solve the Mixed Integer Linear Programming problem based on the proposed model, the results indicate a total cost of $2,554,560 for a 4-week ordering period, 1190 damaged products, and an aggregated objective function value of $1,921,870 under the weighted sum approach (Table 13).

Table 12 Purchasing price and probability values.

Scenario	s_1	s_2	s_3	s_4	s_5	Probability
k_1	C_{p1t1}	C_{p2t1}	C_{p3t1}	C_{p4t1}	C_{p5t1}	0.50
k_2	$0.9C_{p1t1}$	$1.1C_{p2t1}$	$1.1C_{p3t1}$	$0.9C_{p4t1}$	$0.9C_{p5t1}$	0.30
k_3	$1.1C_{p1t1}$	$0.9C_{p2t1}$	$0.9C_{p3t1}$	$1.1C_{p4t1}$	$1.1C_{p5t1}$	0.20

Table 13 Objective function results.

w_1	w_2	γ_1	γ_2	γ_3
0.75	0.25	2,554,560	1190	1,921,870

4.1 Solution Cases Comparison

To clarify the effectiveness of the model when considering the uncertainty of purchasing price (Hosseini *et al.*, 2022; Goodarzi *et al.*, 2022), it is necessary to compare objective function values with non-uncertainty considerations (Hammami *et al.*, 2014). The uncertainty model is called the scenario model (SM) and the non-uncertainty model of Purchasing price is called the deterministic model (DM). Although uncertainty is not considered when running the model, still considers the fluctuation of the purchasing price of each supplier in weight; this weight is based on the probability of the scenario. For example, for supplier 1, the purchasing price in this case would be:

$$C_{p1t}^{DM} = 0.5 \times C_{p1t1} + 0.3 \times 0.9 \times C_{p1t1} + 0.2 \times 1.1 \times C_{p1t1} = 0.99C_{p1t1}$$

Summing up the purchasing price values of each supplier in DM case yields in Table 14. From Table 15, it can be observed that the scenario where SM considers the uncertainty of purchasing prices results in a lower total cost (2,554,560) and objective (weighted sum approach) (1,921,870) compared to the case where DM does not consider uncertainty, demonstrating that the SM approach produces better results. The DM model shows an increase of over 16% in both total cost and objective compared to the SM model, the number of damaged products also increased by more than 32 units. This highlights the effectiveness of the scenario model that accounts for uncertainty. From building a predictive model to developing a Mixed Integer Linear Programming model, the distributor has been able to implement short-term Supplier Selection and Order Allocation planning, effectively addressing the distributor's challenges.

Table 14 Purchasing price values of each supplier in DM.

Scenario	s_1	s_2	s_3	s_4	s_5
DM	$0.99C_{p1t1}$	$1.01C_{p2t1}$	$1.01C_{p3t1}$	$0.99C_{p4t1}$	$0.99C_{p5t1}$

Table 15 The objective results of SM and DM.

Model	w_1	w_2	γ_1	γ_2	γ_3
SM	0.75	0.25	2,554,560	1190	1,921,870
DM	0.75	0.25	2,974,291	1222	2,236,828

4.2 Solution Evaluation

Conduct a collection of actual data for the distributor's first four weeks of 2024, including the number of products sold each, total cost (ordering cost, purchasing cost, transportation cost, and holding cost), and damaged products. Then, the actual figures will be compared against the results of the proposed solution. Table 16 shows the accuracy of the forecast value of 4 weeks compared to reality. It can be seen that the forecast recommendation methods are quite good, and the error criteria are low within the acceptable range, but there are still errors affecting the SSAOA model, although it is difficult to forecast with 100% accuracy, in the future, it is necessary to consider improving the accuracy of forecasting methods. Compare the difference between the total cost and damaged products of the proposed and actual solution to clearly see the effectiveness of the solution, as shown in Table 17.

Table 16 The accuracy of the forecast value of 4 weeks compared to reality.

Error criteria	p_1	p_2	p_3
MAE	1.75	4.75	2.50
RMSE	2.06	5.68	2.92
MAPE (%)	4.02	0.92	10.28

Table 17 Compare the difference between total cost and damaged products in 2 cases.

	Total Cost	Damaged Product
Actual	2,912,839	1261
SSAOA	2,554,560	1190

From Table 17, it can be observed that both the total cost and the number of damaged products in the proposed solution are lower than the actual figures, resulting in a 12.30% decrease in the total cost and a 5.63% decrease in damaged products over four weeks, compared to the actual values for the average of 89 products after being classified into three product types. This demonstrates the effectiveness of the solution, proving that it can address the distributor's challenges and achieve economic efficiency. However, the solution has some limitations, inaccuracies in forecasts can cause the distributor to place short or overstock, leading to increased inventory and costs. The above comparison is only a reference result to evaluate the solution compared to the distributor's current approach. In the future, it is necessary to improve forecasting accuracy or some objectives and constraints suitable to the distributor, economic and social situation. The distributor applying the solution at the present time needs to consider the specific situation and make the right decision based on the results of the Supplier Selection and Order Allocation planning in this study.

4.3 Numerical Experiments

To evaluate the computational efficiency and scalability of the proposed model, a series of numerical experiments were conducted using the CPLEX solver across 16 different examples. Each example progressively increases the number of suppliers, product types, time periods, and demand variations, allowing us to assess both the model's effectiveness and computational feasibility. Performance is measured based on runtime (s), explored nodes, and GAP (%), where GAP serves as an indicator of objective value deviations from the optimal solution (Nasiri *et al.*, 2018), providing insights into how closely the obtained solution approximates the theoretical best outcome.

Table 18 Summary of numerical experiment results.

Experiments	Number of				Runtime (s)	Explored Nodes	GAP (%)
	Suppliers	Product types	Time period	Scenarios			
1	3	2	2	2	3.47	10,456	0
2	3	3	3	2	12.39	33,027	0
3	3	3	3	3	26.58	493,701	0
4	4	3	3	3	67.98	890,521	0
5	4	3	4	3	195.91	1,673,827	0
6	5	4	4	3	343.56	2,500,308	0
7	5	5	4	3	422.50	2,874,755	0
8	6	6	4	3	978.43	3,135,111	0
9	6	7	6	3	1610.52	4,287,355	0
10	8	7	6	4	13,963.81	11,481,027	0.28
11	8	8	6	4	19,951.15	18,540,103	0.33
12	9	8	6	4	23,219.59	21,806,749	0.70
13	10	8	8	4	34,439.22	32,945,547	4.91
14	10	10	8	4	47,410.16	43,961,969	6.95
15	15	10	8	4	66,634.74	60,651,712	6.12
16	15	15	8	5	78,983.82	68,833,560	7.42

The results from Table 18 reveal a notable increase in computational effort when transitioning from 3 to 4 demand scenarios, leading to a sharp rise in runtime while maintaining a low GAP (<1%) when considering 8 - 9 suppliers, 7 - 8 product types over a 6 - 8 week period. This demonstrates that the model remains computationally efficient within this range, as the obtained solutions are very close to optimal. This proposed model efficiently processes datasets comprising up to 9 suppliers and 8 product types, demonstrating its superior scalability and robustness. In contrast, the model presented by Islam *et al.* (2021) was validated with a more limited dataset, involving only 7 suppliers and 3 product types. This distinction highlights the advanced capability of our model, making it more suited for complex and diverse operational environments. This comparison will be clearly articulated to underscore the qualitative advancements offered by the methodology.

However, as the problem size expands beyond 10 suppliers and 10 product types, runtime exceeds 9 hours, and GAP surges past 4.9%, indicating not only longer computation times but also a greater deviation from the optimal solution. This suggests that while the proposed MILP model effectively handles moderate-scale problems, its computational burden and solution accuracy deteriorate for larger datasets. To enhance scalability and maintain solution quality, future research could explore metaheuristic optimization techniques, such as Genetic Algorithms (GA) or Particle Swarm Optimization (PSO), to achieve near-optimal solutions within practical time constraints.

5. CONCLUSION

This research integrates demand forecasting with supplier selection and order allocation to solve a complex problem in the supply chain. Thanks to the demand forecasting with high accuracy, it increases the reliability of the SSAOA model. The proposed solution consists of 2 phases, in phase 1, after comparing the three methods, Holt-Winter, ARIMA, and ANN, it was found that the ANN method is suitable for the fluctuating data series as much as the Bucket and Bottle data, whereas the ARIMA method is suitable for Can's data, the forecasting result is used as input parameter for the model in phase 2. The Multi-Objective Supplier selection and Order allocation model is built in phase 2 to minimize total cost (ordering cost, purchasing cost, transportation cost, and holding cost) and damaged products. The model reduced the total cost by 12.30% and the damaged product rate by 5.63% compared to reality, demonstrating its significant efficiency improvement and helping lubricant distributors save costs. After running experiments, the model was able to solve the problem with up to 15 suppliers and 10 product types, delivering optimal solutions within 8 weeks across multiple scenarios. This showcases the potential to apply the model to other industries and fields, making it a valuable tool for supply chain management in various sectors. In the future, it is possible to expand the development direction of research by considering green supply chain, sustainable constraints as well as discount factors when purchasing, then it is necessary to consider non-linear solutions. Additionally, incorporating metaheuristic optimization algorithms could further enhance the model's efficiency.

ACKNOWLEDGEMENTS

We acknowledge Ho Chi Minh City University of Technology (HCMUT), VNU-HCM for supporting this study.

REFERENCES

Aamer, A., Eka Yani, L., & Alan Priyatna, I. (2020). Data analytics in the supply chain management: Review of machine learning applications in demand forecasting. *Operations and Supply Chain Management: An International Journal*, *14*(1), pp. 1-13.

Agrawal, V., Dixit, J. K., & Agarwal, S. (2021). FTOPSIS approach for material supplier selection: A study. *Materials Today: Proceedings, 45*, pp. 5334-5337.

Alavi, B., Tavana, M., & Mina, H. (2021). A dynamic decision support system for sustainable supplier selection in circular economy. *Sustainable Production and Consumption, 27*, pp. 905-920.

Ali, H., & Zhang, J. (2023). A fuzzy multi-objective decision-making model for global green supplier selection and order allocation under quantity discounts. *Expert Systems with Applications, 225*, p. 120119.

Büyüközkan, G., & Göçer, F. (2017). Application of a new combined intuitionistic fuzzy MCDM approach based on axiomatic design methodology for the supplier selection problem. *Applied Soft Computing, 52*, pp. 1222-1238.

Cheraghalipour, A., & Farsad, S. (2018). A bi-objective sustainable supplier selection and order allocation considering quantity discounts under disruption risks: A case study in plastic industry. *Computers & Industrial Engineering*, *118*, pp. 237-250.

De Oliveira, M. E. B., Lima-Junior, F. R., & Galo, N. R. (2023). A comparison of hesitant fuzzy VIKOR methods for supplier selection. *Applied Soft Computing, 149*, p. 110920.

Dercole, F., Sangiorgio, M., & Schmirander, Y. (2020). An empirical assessment of the universality of ANNs to predict oscillatory time series. *IFAC-PapersOnLine*, 53(2), pp. 1255-1260.

Dey, K., & Saha, S. (2018). Influence of procurement decisions in two-period green supply chain. *Journal of cleaner production*, *190*, pp. 388-402.

Douaioui, K., Fri, M., Mabrouki, C., & Semma, E. A. (2021). A multiobjective integrated procurement, production, and distribution problem of supply chain network under fuzziness. *IFAC-PapersOnLine*, *54*(1), pp. 1104-1111.

Firouzi, F., & Jadidi, O. (2021). Multi-objective model for supplier selection and order allocation problem with fuzzy parameters. *Expert Systems with Applications, 180*, p. 115129.

Gonçalves, J. N., Cortez, P., Carvalho, M. S., & Frazao, N. M. (2021). A multivariate approach for multi-step demand forecasting in assembly industries: Empirical evidence from an automotive supply chain. *Decision Support Systems*, *142*, p. 113452.

Goodarzi, F., Abdollahzadeh, V., & Zeinalnezhad, M. (2022). An integrated multi-criteria decision-making and multi-objective optimization framework for green supplier evaluation and optimal order allocation under uncertainty. *Decision Analytics Journal*, *4*, p. 100087.

Guijarro, R., Trujillo-Santos, J., Bernal-Lopez, M. R., de Miguel-Díez, J., Villalobos, A., Salazar, C., ... & Monreal, M. (2015). Trend and seasonality in hospitalizations for pulmonary embolism: a time-series analysis. *Journal of Thrombosis and Haemostasis*, *13*(1), pp. 23-30.

Hajiaghaei-Keshteli, M., Cenk, Z., Erdebilli, B., Özdemir, Y. S., & Gholian-Jouybari, F. (2023). Pythagorean fuzzy TOPSIS method for green supplier selection in the food industry. *Expert Systems with Applications, 224*, p. 120036.

Hammami, R., Temponi, C., & Frein, Y. (2014). A scenario-based stochastic model for supplier selection in global context with multiple buyers, currency fluctuation uncertainties, and price discounts. *European Journal of Operational Research*, 233(1), pp. 159-170.

Hosseini, Z. S., Flapper, S. D., & Pirayesh, M. (2022). Sustainable supplier selection and order allocation under demand, supplier availability and supplier grading uncertainties. *Computers & Industrial Engineering, 165*, p. 107811.

Islam, S., Amin, S. H., & Wardley, L. J. (2021). Machine learning and optimization models for supplier selection and order allocation planning. *International Journal of Production Economics, 242*, p. 108315.

Jafari-Raddani, M., Asgarabad, H. C., Aghsami, A., & Jolai, F. (2024). A hybrid approach to sustainable supplier selection and order allocation considering quality policies and demand forecasting: a real-life case study. *Process Integration and Optimization for Sustainability*, *8*(1), pp. 39-69.

Jia, R., Liu, Y., & Bai, X. (2020). Sustainable supplier selection and order allocation: Distributionally robust goal programming model and tractable approximation. *Computers & Industrial Engineering, 140*, p. 106267.

Joy, T. M., Aneesh, K. S., & Sreekumar, V. (2023). Analysis of a decision support system for supplier selection in glove industry. *Materials Today: Proceedings, 72*, pp. 3186-3192.

Kaur, H., & Singh, S. P. (2021). Multi-stage hybrid model for supplier selection and order allocation considering disruption risks and disruptive technologies. *International Journal of Production Economics, 231*, p. 107830.

Liou, J. J., Chang, M. H., Lo, H. W., & Hsu, M. H. (2021). Application of an MCDM model with data mining techniques for green supplier evaluation and selection. *Applied Soft Computing, 109*, p. 107534.

Micus, C., Schramm, S., Boehm, M., & Krcmar, H. (2023). Methods to analyze customer usage data in a product decision process: A systematic literature review. *Operations Research Perspectives*, p. 100277.

Miranda, J. L. M., & Saucedo, J. A. M. (2021). A Multi-product Stochastic Programming Model for Supplier Selection in a Humanitarian Relief Chain. In *Computer Science and Health Engineering in Health Services: 4th EAI International Conference, COMPSE 2020, Virtual Event, November 26, 2020, Proceedings 4* (pp. 31-51). Springer International Publishing.

Nasiri, M. M., Rahbari, A., Werner, F., & Karimi, R. (2018). Incorporating supplier selection and order allocation into the vehicle routing and multi-cross-dock scheduling problem. *International Journal of Production Research*, *56*(19), pp. 6527-6552.

Nayeri, S., Khoei, M. A., Rouhani-Tazangi, M. R., GhanavatiNejad, M., Rahmani, M., & Tirkolaee, E. B. (2023). A data-driven model for sustainable and resilient supplier selection and order allocation problem in a responsive supply chain: A case study of healthcare system. *Engineering Applications of Artificial Intelligence, 124*, p. 106511.

Odei, S. A., & Hamplová, E. (2022). Innovations in small businesses: do public procurement contracts and intellectual property rights matter?. *Heliyon*, *8*(9).

Omar, M. S., & Kawamukai, H. (2021). Prediction of NDVI using the Holt-Winters model in high and low vegetation regions: A case study of East Africa. *Scientific African*, 14, e01020.

Oostwal, E., Straat, M., & Biehl, M. (2021). Hidden unit specialization in layered neural networks: ReLU vs. sigmoidal activation. *Physica A: Statistical Mechanics and its Applications*, *564*, p. 125517.

Ordu, M., Demir, E., Tofallis, C., & Gunal, M. M. (2021). A novel healthcare resource allocation decision support tool: A forecasting-simulation-optimization approach. *Journal of the Operational Research Society*, *72*(3), pp. 485-500.

Pektaş, A. O., & Cigizoglu, H. K. (2013). ANN hybrid model versus ARIMA and ARIMAX models of runoff coefficient. *Journal of Hydrology*, *500*, pp. 21-36.

Putro, S. P., Koshio, S., & Oktaferdian, V. (2016). Implementation of Arima model to asses seasonal variability macrobenthic assemblages. *Aquatic Procedia*, *7*, pp. 277-284.

Rachidi, N. R., Nwaila, G. T., Zhang, S. E., Bourdeau, J. E., & Ghorbani, Y. (2021). Assessing cobalt supply sustainability through production forecasting and implications for green energy policies. *Resources Policy*, *74*, pp. 102423.

Rumbe, G., Hamasha, M., & Al Mashaqbeh, S. (2024). A comparison of Holts-Winter and Artificial Neural Network approach in forecasting: A case study for tent manufacturing industry. *Results in Engineering*, *21*, p. 101899.

Saputro, T. E., Figueira, G., & Almada-Lobo, B. (2023). Hybrid MCDM and simulation-optimization for strategic supplier selection. *Expert Systems with Applications, 219*, p. 119624.

Seyedan, M., & Mafakheri, F. (2020). Predictive big data analytics for supply chain demand forecasting: methods, applications, and research opportunities. *Journal of Big Data*, *7*(1), p. 53.

Shidpour, H., Shidpour, M., & Tirkolaee, E. B. (2023). A multi-phase decision-making approach for supplier selection and order allocation with corporate social responsibility. *Applied Soft Computing, 149*, p. 110946.

Siemieniako, D., Makkonen, H., & Mitręga, M. (2023). Buying center-selling center interaction as a driver for power dynamics in buyer-supplier relationships. *Industrial Marketing Management*, *114*, pp. 94-109.

Tealab, A., Hefny, H., & Badr, A. (2017). Forecasting of nonlinear time series using ANN. *Future Computing and Informatics Journal*, *2*(1), pp. 39-47.

Tratar, L. F., & Strmčnik, E. (2016). The comparison of Holt–Winters method and Multiple regression method: A case study. *Energy*, *109*, pp. 266-276.

Vaezi, A., Rabbani, E., & Yazdian, S. A. (2024). Blockchain-integrated sustainable supplier selection and order allocation: A hybrid BWM-MULTIMOORA and bi-objective programming approach. *Journal of Cleaner Production*, *444*, p. 141216.

Ventura, J. A., Bunn, K. A., Venegas, B. B., & Duan, L. (2021). A coordination mechanism for supplier selection and order quantity allocation with price-sensitive demand and finite production rates. *International Journal of Production Economics, 233*, p. 108007.

Wang, Y., Wang, J., Zhao, G., & Dong, Y. (2012). Application of residual modification approach in seasonal ARIMA for electricity demand forecasting: A case study of China. *Energy Policy*, *48*, pp. 284-294.

Wu, C., Lin, Y., & Barnes, D. (2021). An integrated decision-making approach for sustainable supplier selection in the chemical industry. *Expert Systems with Applications, 184*, p. 115553.

Zhang, H., Zhang, L., Xu, L., Ma, X., Wu, Z., Tang, C., ... & Yang, Y. (2020, August). A request-level guaranteed delivery advertising planning: Forecasting and allocation. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2980-2988).

Zheng, M., Zhou, H., Jiang, P., Pan, E., Zhao, S., & Wu, K. (2021). Supplier selection problem for multiple projects with uncertain demand and project life cycles. *Computers & Operations Research, 132*, p. 105312.

Huu-Tin To, was born in Vietnam in 2002. He applied to the Ho Chi Minh City University of Technology (HCMUT), VNU-HCM in 2020 and has since graduated with a degree in Logistics - Supply Chain. He is particularly interested in the fields of Supplier, Procurement, and Forecasting within the supply chain. This topic is developed from his graduation thesis.

Mai-Ha Phan, was born in Vietnam in 1977 and is an Associate Professor of Industrial System Engineering Department at Ho Chi Minh City University of Technology (HCMUT), VNU-HCM. She got a Bachelor's degree with a major in Industrial Management from Ho Chi Minh University of Technology in 2000. She got a master's degree in Industrial Engineering from the Asian Institute of Technology, Thailand in 2001. She finished PhD program in Industrial Engineering in 2015 at Pusan National University, South Korean. Her research focuses on issues in the design and operation of logistics systems.