J. COMBIN. MATH. COMBIN. COMPUT. 127a (2025) 5871--5894

Journal of Combinatorial Mathematics
and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Optimization Research of Enterprise Supply Chain Demand Forecasting and Inventory Cost Control Based on Machine Learning Models

Teng Zhang[1,✉], Guoqiang Hao[1], Zhenhua Zhang[2], Chenyu Song[2], Chenxin Cui[2]

[1] *Economics and Management School, Taiyuan University of Technology, Taiyuan, Shanxi, 030024, China*

[2] *Software School, Taiyuan University of Technology, Taiyuan, Shanxi, 030024, China*

ABSTRACT
Market economy is characterized by the uncertainty of supply and demand, so enterprises can realize the optimization of inventory cost control only by reasonably forecasting the demand of supply chain. This paper studies a supply chain demand forecasting method based on machine learning. The factors affecting supply chain demand are collected and analyzed, and the ARMA model, which combines autoregressive model and moving average model, is used to forecast supply chain demand. Then, through the introduction of procurement cost, storage cost and time cost, a multi-level inventory model is established, and the immune genetic algorithm is used to solve the model to find the optimal inventory cost. The experimental results show that the prediction model has good forecasting performance. After using the optimized scheme, the total inventory cost of the enterprise supply chain is reduced by 17.35% and 13.69% respectively. It can be seen that, on the whole, the method in this paper has a good effect of supply chain demand forecasting and cost control.

Key words: ARMA, Machine learning, Supply chain demand forecasting, Inventory cost control, Immune genetic algorithm

1. Introduction

In the 21st century, the competitive situation of enterprises has undergone profound changes. It is no longer limited to the competition among single enterprises, but has evolved into the multi-dimensional comprehensive competition of talents, technology, innovation and supply chain [1-4]. In this diversified competitive landscape, the importance of supply chain, as a key link connecting internal and external resources of enterprises and driving business development, has become increasingly prominent. The supply chain is not only a simple logistics chain, but also a

✉ Corresponding author.
E-mail address: 15031123257@163.com (T. Zhang).
Received 12 January 2024; Revised 25 March 2024; Accepted 25 December 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-325

complex and precise network chain structure led by core enterprises and surrounded by many "satellite" enterprises [5]. In this network, the core enterprises have a decisive impact on the overall operational efficiency and competitiveness of the supply chain by virtue of their market position, technical strength and management ability [6-7].

The management of supply chain involves the comprehensive control of information flow, logistics and capital flow. From the purchase of raw materials to the manufacture of intermediate products to the sale of final products, every link is closely linked, and together constitute a complete functional network chain model [8-9]. The operation efficiency and capability of this mode are directly related to the rapid development of an enterprise's business and the improvement of its market competitiveness. Among them, the accuracy of supply chain demand forecast, as a key link, directly affects the rapid response of supply and demand relationship and the reduction of supply chain cost, and then affects the final cost benefit [8]. Accurate demand forecasting can help enterprises better grasp the market dynamics, optimize inventory management, reduce waste and improve resource utilization efficiency.

In the management of supply chain, the core enterprise's choice and management of suppliers is particularly important. Different supplier selection strategies will directly affect the procurement cost and inventory management cost of the core enterprise [10-12]. As an important part of supply chain management, inventory is not only related to the storage of goods and the management of resources, but also directly affects the flow of funds and cost control of enterprises. Core enterprises' management of the sales of finished products, such as the quantity, purchase mode and inventory management of sales enterprises, will indirectly affect the management cost of finished products of production enterprises, and then affect the total inventory cost of core enterprises [13-14].

Inventory cost, as an important expenditure in enterprise management, includes ordering cost, storage cost, out-of-stock cost, expired loss cost, capital cost and other aspects, and is affected by many factors such as demand fluctuation, supply uncertainty and product life cycle [15-16]. In the total cost of an enterprise, inventory cost often occupies a considerable proportion. Studies have shown that the inventory cost of an enterprise usually accounts for about 30% of the total cost, and in the logistics cost, the inventory cost is as high as 80%-90%[17-19]. This means that effective control of product inventory costs is of great importance to the cost control of enterprises.

Traditional inventory management methods often rely on manual experience and simple mathematical model, which is difficult to deal with the complex and changeable market environment and supply chain network. With the rapid development of big data and machine learning technology, there are new ideas and tools for supply chain management and inventory cost control. By analyzing and learning a large amount of historical data, machine learning algorithms can dig out potential laws and patterns, and provide more accurate and efficient decision support for demand forecasting, supplier selection, inventory management and other aspects. Therefore, this paper aims to explore the research on enterprise supply chain and inventory cost based on machine learning, optimize supply chain management and inventory cost control strategies by introducing machine learning technology, improve the market competitiveness and economic benefits of enterprises, and deeply analyze the application prospects and challenges of machine learning in supply chain demand forecasting, supplier selection, inventory management and other aspects. Corresponding solutions and research are proposed to provide new ideas and methods for supply chain management and inventory cost control of enterprises.

This paper selects the supply chain data of a company G as the data set of machine learning research. First, according to the characteristics of the data set, we conduct in-depth analysis of the data from four dimensions: warehouse, weeks, distribution and procurement. Then, based on ARMA (autoregressive moving average) method, we construct the supply chain demand forecast model. In the process of model construction, ADF (unit root test) was used to test the smoothness of the data to ensure the applicability of the data. Then, the specific form of the model was determined by analyzing the autocorrelation coefficient and partial autocorrelation coefficient. In order to optimize the model, the criterion function ranking method is used to rank several candidate models. On this basis, the least square method is used for accurate estimation. Then, considering multiple factors such as procurement cost, inventory holding cost and time cost, we put forward relevant assumptions and build a multi-level inventory model of supply chain. To solve this complex model, we introduce immune genetic algorithm, which can effectively find the optimal solution of the model. Finally, we verify the predictive performance of the constructed supply chain demand forecasting model and the effectiveness of the multi-level inventory model in inventory cost control optimization through an example analysis.

2. Overview

In today's globalized business environment, the competition between enterprises has gone beyond a single dimension contest, but has evolved into a comprehensive competition in many aspects such as talent, technology, innovation and supply chain. Supply chain, as a key link connecting internal and external resources of enterprises and driving business development, its management efficiency and cost control ability are directly related to the market competitiveness and economic benefits of enterprises. As the core link of supply chain management, supply chain demand forecasting plays a vital role in optimizing inventory, reducing cost, increasing sales, improving profits and customer loyalty.

In recent years, with the rapid development of big data and machine learning technology, the methods and technologies of supply chain demand forecasting are constantly innovating and progressing. Literature [20] improves the current supply chain demand forecasting system by means of support vector regression, deep learning models and new integrated strategies, which significantly improves the accuracy of forecasting. This research result provides supply chain managers with a more accurate demand forecasting tool, which helps them better grasp the market dynamics and optimize inventory management strategies. Literature [21] incorporated time series and explanatory factors into the machine learning supply chain demand forecasting model, which not only realized the accurate prediction of demand, but also revealed the variability of savings in the supply chain performance results. This finding provides supply chain managers with deeper insight, helps them better understand the internal mechanism of supply chain operation, and then formulate more effective management strategies. Literature [22] uses machine learning, time aggregation mechanism and UNISON data-driven to build a framework for supply chain demand forecasting of electronics companies, and carries out realistic verification of this method. The results show that this method has good robustness and can maintain stable forecasting performance under different market environments. The research results provide a feasible demand forecasting solution for supply chain managers in the electronics industry. Literature [23] uses the time series data of products in the supply chain to train these data for prediction through machine learning, and learns inventory data, supply chain structure information, etc., to further improve the forecasting performance. This study demonstrates the strong potential of machine learning in supply chain demand forecasting, providing supply chain managers with more diversified data analysis and forecasting tools. Literature [24] summarizes AI-driven supply chain demand forecasting methods, including deep learning algorithms, product neural networks, recurrent neural networks, reinforcement learning algorithms, etc., and their hybrid models. These methods accurately predict demand in a real supply chain environment, optimizing cost-effective inventory levels. In particular, the application of hybrid models has further improved the accuracy and reliability of forecasts. Literature [25] constructed a hybrid model to predict baseline demand and promotion demand under the supply chain, and verified the accuracy of the model and inventory performance through statistics and machine learning. This model can accurately predict the fluctuating demand series with different coefficient of variation, and achieve low inventory cost. The research results provide supply chain managers with more flexible and adaptable demand forecasting tools.

Most of these researches make predictions based on product time series data, supply chain information data, deep learning algorithms, machine learning algorithms, etc., and gradually improve the prediction accuracy. At the same time, they also carry out simple optimization of inventory costs, but they still need to be explored in depth.

In terms of inventory cost, literature [26] verifies the price, weight and volume of goods in inventory under hypothetical conditions, and finds that enterprises or warehouses with equal input costs have no significant differences in inventory holding costs, while enterprises or warehouses with multiple products have significant differences. This finding provides a basis for inventory managers to adjust inventory management strategies and helps to optimize inventory holding costs. In literature [27], immune genetic algorithm was used to calculate the time cost of inventory management at multiple levels of supply chain with the support of Internet of Things technology, and inventory cost control was realized. This research result provides a new inventory cost control method for supply chain managers, which is helpful to improve the overall efficiency of supply chain. Literature [28] envisages an ant colony algorithm and fuzzy model to support supply chain joint inventory management and optimize the cost, so as to solve the inventory control problem and improve the supply chain integration. Although this hypothesis has not been verified yet, it provides new ideas and directions for supply chain inventory management.

The accurate prediction of supply chain demand plays a vital role in inventory cost control. Cost control is the key for enterprises to maintain competitiveness and improve profitability in the fierce market competition. In the current research, machine learning has achieved remarkable results in supply chain demand forecasting, but relatively few applications have been made in inventory cost control.

This paper applies machine learning algorithm to supply chain demand forecasting and inventory cost control, aiming to save development cost and improve supply chain management efficiency by improving forecasting accuracy and optimizing inventory management strategy. This paper believes that integrating machine learning technology into supply chain management and inventory cost control can not only improve the accuracy of forecasting and the efficiency of inventory management, but also help enterprises better cope with market changes, reduce operating costs and improve profitability. Therefore, this paper argues that enterprises should actively explore and apply the potential of machine learning technology in supply chain demand forecasting and inventory cost control, so as to improve their market competitiveness and economic benefits.

3. Data source and pre-processing

3.1. Data source and data set description

Part of the data of the food supply chain of G Company is selected as the basic experimental data of this study. In order to have a comprehensive and in-depth understanding of the variable distribution characteristics and statistical rules of the selected data set, this paper carries out a detailed statistical analysis of the data set involved, and presents the statistical results clearly in the form of tables, as shown in Table 1. After careful combing and analysis, we found that the supply chain data covered the period from the 4th week to the 10th week of production activities of the enterprise, and the data in this period of time could fully reflect the running state of the supply chain in a specific production cycle. In terms of warehouse, there is a wide range of warehouse ids with replenishment behavior, extending from 1123 to 23548, which shows the diversity and extensibility of warehouse distribution in the supply chain.

Table 1. Statistical characterization of data sets

	Time/week	Warehouse ID	Distribution route ID	Client ID	Product ID	Demand
Count	74357	74214	75642	74621	74218	74652
Mean	5.39	2539.21	2113.8	1824682.2	20764.69	7.67
Min	4	1123	1	58	55	0
Last quartile	5	1321	1139	339871.39	1238	3
Median	7	1628	1287	1206687	30984	4
Lower quartile	9	2033	2874	2387216	37891	8
Max	10	23548	9964	11654873	49871	1721

In terms of demand records, a total of 75,642 demand records were counted in this paper, reflecting the demand of each node of the supply chain. Through calculation, the average node demand of the supply chain is 7.67, which reveals the average level of the whole supply chain demand. Further analysis shows that the real demand of most samples is less than 8, which indicates that the supply chain demand is kept in a relatively stable range in most cases. However, there are also some cases where the demand of sales nodes is abnormally high, among which the maximum demand reaches 1721, suggesting that we need to pay special attention to these high-demand nodes in supply chain management to ensure the smooth operation of the supply chain.

This paper also makes statistics on the number of customers and products in the system, and the number of existing products also reaches 49,871, showing the richness and diversity of product types in the supply chain. At the same time, the number of customers in the system is as high as 116,4873, providing a broad market space for the supply chain. These data provide a solid foundation and rich material for the follow-up research of this paper.

The description of key fields is shown in Table 2. From the description of the data set, it can be seen that the task of the competition is to organize and analyze the information of different channels, car yards, customers and products within several weeks, form the corresponding demand prediction model, and accurately predict the customer demand in the future time. Node demand is the target variable that needs to be predicted, while other variables exist to reveal the real demand fluctuation. In-depth analysis of the correlation between other variables can help policy makers grasp the changes of the influencing factors of demand, so as to dig out the real demand fluctuation pattern.

Table 2. Description of key fields

Name	Description
Semana	Weeks
Agencia_ID	Warehouse ID
Canal_ID	Purchase channel ID
Ruta_SAK	Distribution route ID
Cliente_ID	Client ID
NombreCliente	Client name
Producto_ID	Product ID
NombreProducto	Product name
Venta_uni_hoy	The number of sales units this week
Venta_hoy	Sales this week
Dev_uni_proxima	The number of units returned next week
Dev_proxima	Return next week
Demanda_uni_equil	Node demand

The correlation analysis of the data set can provide in-depth insight into the customer characteristics and provide a strong basis for the segmentation of customer groups. By tailoring an exclusive service to each specific customer group, its needs can be met more effectively. At the same time, this analysis also helps to simulate the real market environment and quickly catch and respond to emerging changes in demand, thereby improving the return rate of the enterprise. In addition, in-depth analysis of relevant factors can strengthen the cooperation among enterprises in the supply chain, optimize the supply chain management process, reduce service costs, and enhance the overall competitiveness of the supply chain. More importantly, in-depth analysis can also help enterprises to explore potential profit growth points and promote enterprises to achieve innovation or transformation and upgrading. Therefore, this paper will carry out a detailed analysis of some of the key influencing factors.

3.2. Warehouse Analysis

When analyzing the operation of the warehouse, we use agentcia_id (warehouse ID) as the keyword to group and summarize the data in detail. Through this step, we can clearly calculate the order attributes processed by each warehouse every week, including the number of orders, order types and other key indicators. As can be seen from Figure 1, when the weekly sales volume reaches 100 units, the number of warehouses is relatively large. With the gradual increase of sales, the number of warehouses began to decrease. This finding provides us with important information about the distribution of warehouse sales, which helps us to better understand the operation status and sales characteristics of warehouses.

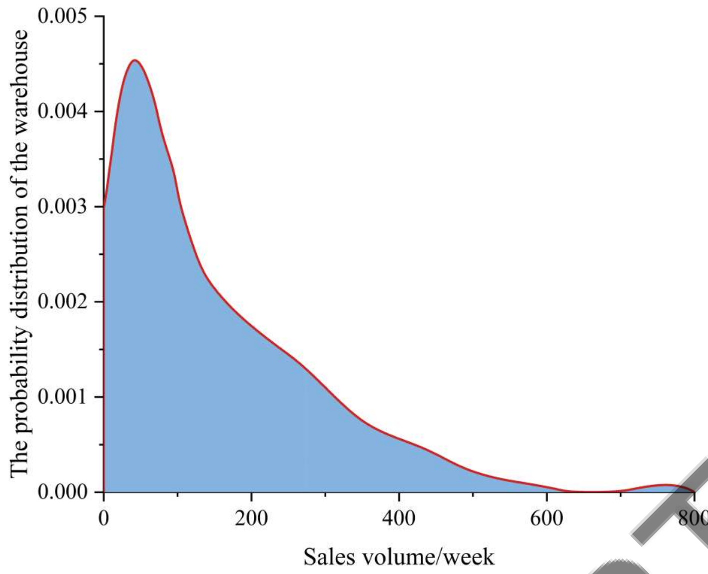

Fig. 1. Distribution of warehouse demand

3.3. Weekly Analysis

The weekly distribution of product orders is shown in detail in Figure 2, the core of which is the key attribute of "Semana" (i.e. weekly). From the figure, we can clearly see that the distribution of the data is mainly concentrated in the period from week 4 to week 10. During this period, the total number of weekly orders showed a relatively stable trend, and the increase or decrease was not large on the whole. This small increase or decrease may be affected by a variety of factors, such as market demand, promotional activities, supply chain conditions, etc., but generally speaking, the order volume of products has maintained a relative stability during this period of time, which provides a strong data support for the production plan and inventory management of enterprises.

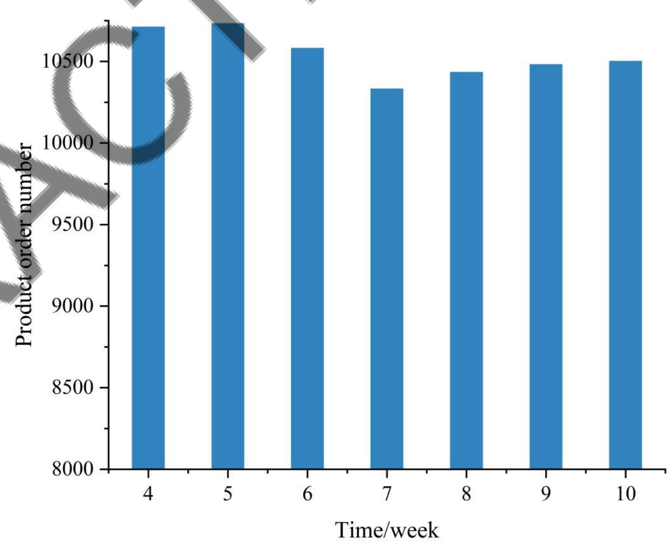

Fig. 2. Distribution of order weeks

3.4. Distribution route analysis

The dataset is grouped and summarized with the keyword "Ruta_SAK" (distribution routes), and it can be seen from the raw data that some distribution routes do not exist, but their functions are directly exercised by the sales warehouse. The number of orders processed per week for each distribution route is calculated to obtain the distribution of the number of orders in the distribution channel, i.e. Figure 3. As can be seen from the figure, the weekly order processing capacity of the distribution route in weeks 4-10 is mainly concentrated in the interval [0, 150]. As the number of orders processed per week increases, the corresponding number of distribution routes gradually decreases.

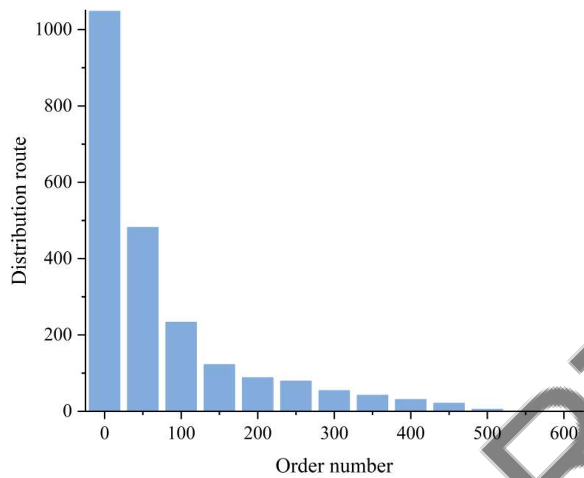

Fig. 3. Distribution of the distribution channel's orders

3.5. Analysis of procurement channel

As the main source of warehouse goods, the incoming material channel plays a crucial role in supply chain management. Therefore, when analyzing the incoming material channel, we pay special attention to the corresponding relationship between the channel and the warehouse, and carry on the in-depth analysis. In order to reveal this correspondence, we use the method of statistical warehouse corresponding to the number of incoming channels, in order to quantify and describe the degree of correlation between the two.

The statistical results are shown in Figure 4. By observing the chart, it can be found that among all warehouses, quite a few warehouses have only one supply channel, which reflects the concentrated characteristics of supply channels. As the number of supply channels increases, the corresponding number of warehouses gradually decreases, which indicates that there are relatively few warehouses with multiple supply channels. This finding is of great significance for optimizing supply chain management, improving warehouse operation efficiency and reducing risks, and reminds us that in the future supply chain layout, it is necessary to allocate the relationship between warehouses and supply channels more reasonably to achieve more efficient and stable logistics operation.

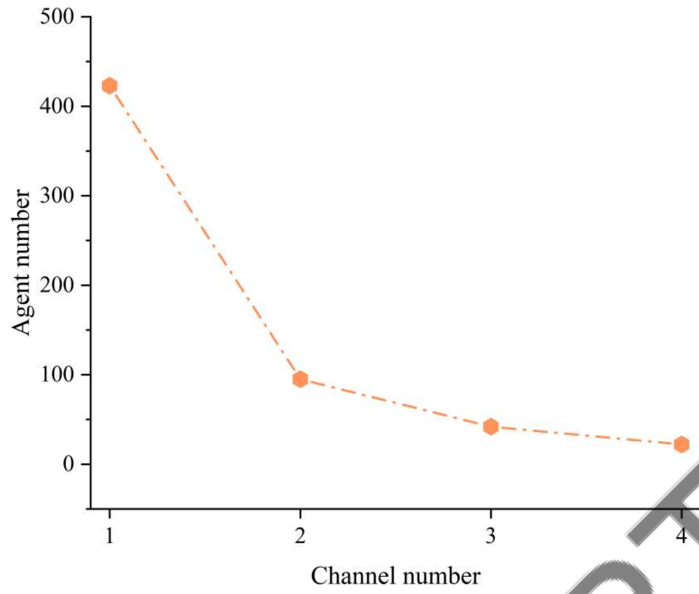

Fig. 4. Statistical result

4. Construction of supply chain demand forecasting model based on ARMA model

4.1. Forecasting Modeling

4.1.1. Time series modeling. Enterprise supply chain has the characteristics of changing with time, so it can be predicted and analyzed by using the relevant prediction method of time series model, so as to enhance the accuracy of prediction and improve the operation efficiency of enterprises.

In the field of time series analysis and prediction, autoregressive moving average (ARMA) model has become one of the mainstream methods due to its high accuracy and wide applicability. By integrating the mathematical expressions and characteristics of autoregressive (AR) and moving average (MA) models, this model can capture both autocorrelation and random perturbation of time series. Compared with AR or MA model alone, ARMA shows better adaptability in the parameter fitting stage, especially in the spectrum analysis, more detailed decomposition effect can be obtained [29]. Specifically, the autoregressive part (AR) of the model is used to quantify the correlation between current and historical observations, while the moving average part (MA) is used to eliminate random error terms in the time series. The model is generally expressed as ARMA(p,q), and its mathematical expression is as follows:

$$y_t = \gamma_0 + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \ldots + \gamma_p y_{t-p} + \omega_t + \theta_1 \omega_{t-1} + \theta_2 \omega_{t-2} + \ldots + \theta_q \omega_{t-q} \tag{1}$$

Since AR(p) and MA(q) models are simplified forms of ARMA (p, q) models in specific cases, ARMA (p, q) models have a wider applicability than AR(p) and MA(q) models, and can be used to predict and analyze any smooth time series.

4.1.2. ARMA modeling and construction. The premise of ARMA model is that the time series should be stationary time series, so the process of establishing the model should first detect the stationarity of the time series. If the time series is non-stationary time series, it needs to conduct differential stationarity for the non-stationary time series, and then identify the required time series model. After recognition, the order of the selected time series model is determined, that is, the model parameters of the model are determined, and then the model diagnosis is carried out on the built model, and finally the time series prediction is carried out through the built model.

1) Smoothness test. Statistical measure test This hypothesis test is a way to judge the smoothness of time series more accurately than graph test. Among them, the unit root test is the most commonly used method and is widely used in the smoothness test. The unit root test analyzes the position relationship between the feature root and the unit circle by constructing the autoregressive feature equation. If the eigenroots are located in the unit circle, the unit roots are considered to exist. If there is no unit root, then the sequence is smooth. ADF test is the most commonly used method in unit root test. This paper uses this method to judge the smoothness of sequence.

If there is a unit root in the autoregression process, the relationship between the independent variable and the dependent variable is unclear, and the value of the residual will not decrease with the increase of the sample size, which means that the residual cannot be eliminated from the model, and this kind of regression is called pseudo-regression. If there is a unit root, it indicates that the process is a random walk.

The three scenarios for the ADF unit root test are as follows:

$$\begin{cases} y_t = \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \ldots + \gamma_p y_{t-p} + \omega_t \\ y_t = a_0 + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \ldots + \gamma_p y_{t-p} + \omega_t \text{(Contains intercept entries)} \\ y_t = a_0 + a_2 t + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \ldots + \gamma_p y_{t-p} + \omega_t \\ \text{(Contains the intercept item and the time trend item)} \end{cases} \tag{2}$$

Namely:

$$\begin{cases} \Delta y_t = \phi y_{t-1} + \eta_1 \Delta y_{t-1} + \ldots + \eta_{p-1} \Delta y_{t-p-1} + \omega_t \\ \Delta y_t = a_0 + \phi y_{t-1} + \eta_1 \Delta y_{t-1} + \ldots + \eta_{p-1} \Delta y_{t-p-1} + \omega_t \\ \Delta y_t = a_0 + a_0 t + \phi y_{t-1} + \eta_1 \Delta y_{t-1} + \ldots + \eta_{p-1} \Delta y_{t-p-1} + \omega_t \end{cases} \tag{3}$$

Where. $\phi = \gamma_1 + \gamma_2 + \ldots \gamma_{p-1}, \eta_i = -(\gamma_{i+1} + \gamma_{i+2} + \gamma_p)$

The original and alternate hypotheses are:

$$\begin{cases} H_0 : \phi = 0 \\ H_1 : \phi < 0 \end{cases} \tag{4}$$

In the ADF test method, the null hypothesis (hypothesis) is set as the existence of the unit root. H_o In this case, the ADF test statistic is calculated. If the calculated statistic is less than the critical values of the three confidence levels of 10%, 5%, and 1% given in advance, then there is a 90%, 95%, and 99% probability of rejecting the null hypothesis, respectively, which means that the series is stable [30].

2) Model identification and order determination. Determine the model using the autocorrelation coefficient (ACF) and partial autocorrelation coefficient (PACF) as follows:

If the ACF of the observations shows that the former q term falls outside the range, and the ACF of more than 95% of the observations falls within that range, and the ACF oscillates rapidly as it decreases to near 0, the ACF exhibits a truncation feature. $2(1/\sqrt{T})$ In this case, the predictive model suitable for this time series can be identified as the MA(q) model. If PACF also exhibits the same characteristics, then the prediction model for that time series is the AR(p) model.

If the ACF and PACF of the observed value exceed the range by more than 5%, or the ACF and PACF gradually decrease to 0 and show small amplitude oscillations in a smooth process, then the ACF and PACF show a trailing feature, and the prediction model suitable for this time series is the ARMA model. $2(1/\sqrt{T})$

In this paper, the criterion function ranking method is used to select the order of the model. The commonly used criterion functions in criterion function ranking method include AIC and BIC. When using this method to set the order of the model, the characteristics of the order of the model are usually first determined by sample autocorrelation and partial autocorrelation graphs, so as to select the parameters to be determined. If only one set of parameters is identified, no order setting is required; If multiple parameter groups are identified, and each group of parameters satisfies the basic conditions of the ARMA model, then the AIC or BIC value of each group of parameters is calculated according to the criterion function, and the one with the smallest AIC or BIC value is selected as the final order of the model. The specific expressions of AIC and BIC are as follows:

AIC (Minimum information criterion) function:

$$AIC = \ln \bar{\sigma}_a^2 + \frac{2(p+q)}{N} \tag{5}$$

BIC (Bayesian Information Criterion) function:

$$BIC = \ln \bar{\sigma}_a^2 + \frac{(p+q)\ln(N)}{N} \tag{6}$$

3) Parameter estimation and model diagnosis. In this paper, the least square method is used for parameter estimation. The core idea of the least square method is to square and sum the difference between the true value and the fitted value. When the sum reaches the minimum, the corresponding parameter is the estimated value of the model. LSE is one of the common methods used to estimate the parameters of a time series model. It assumes that the value of a random variable not observed in the historical data is 0, and thus deduces, i.e. : $y_i = 0, \forall t \le 0$

$$\varepsilon_{\mathrm{t}} = \frac{\Gamma(B)}{\Theta(B)} y_t - \sum_{i=1}^{l} \pi_i y_{t-i} \tag{7}$$

Then the sum of squares of residuals at this time is:

$$S(\hat{\gamma}_1, \dots, \hat{\gamma}_p, \hat{\theta}_1, \dots, \hat{\theta}_q) = \sum_{t=1}^{r} (y_t - \sum_{i=1}^{r} \pi_t y_{t-1})^2 \tag{8}$$

The ARMA model needs to be diagnosed and tested after the completion of modeling, mainly including the following points:

In, its minimum point can be regarded as the least squares estimate of ARMA (p, q) parameters. $S(\hat{\gamma}_1, \dots, \hat{\gamma}_p, \hat{\theta}_1, \dots, \hat{\theta}_p)$ $\hat{\gamma}_1, \dots, \hat{\gamma}_p, \hat{\theta}_1, \dots, \hat{\theta}_p$ When the sample size is large enough, the properties of the least squares estimator are similar to the moment estimator, such as conforming to the asymptotic normal distribution, being an unbiased estimator of the parameters, etc.

The ARMA model needs to be diagnosed and tested after the completion of modeling, mainly in the following aspects:

(1) Smooth reversibility test: the premise of establishing ARMA model is that the time series has stationarity. Therefore, it is necessary to calculate the polynomial roots of the autoregressive coefficient and moving average coefficient of the series. If the modulus of the root is greater than or equal to 1, that is, the test fails, and the difference is needed to make it smooth.

(2) Residual sequence test: After modeling, the fitting effect of the model should be tested, which can be judged by the randomness of the model prediction error sequence. If the model can extract effective information from the observation sequence well, the predicted residual sequence should be a pure random sequence; If it is not pure random, the model needs to be re-identified.

4.2. Analysis of prediction effect

4.2.1. Time series smoothness detection. The smoothness of time series means that the features of each time point in the time series are evenly distributed, stable and unchanged, and do not change significantly with time. In general, testing the smoothness of time series depends on the statistical verification of the significance of data smoothness, and the unit root test is especially used. If the

P-value obtained by the test is lower than 0.05, it means that the series has no unit root and meets the criteria for smooth series. Conversely, if the P-value is higher than 0.05, the sequence is considered non-stationary.

Table 3 details the results of the unit root test. Augmented Dickey-Fuller Test (ADF), , was used to test the stationarity of the time series data, and obtained a P-value of 0.5345, which did not support the original hypothesis, indicating that the smoothness of the data was not significant, and the tested series was non-stationarity. Therefore, differential processing should be taken as the first step when studying this kind of series. Further, if the p-value of the first difference sequence is less than 0.05, the original assumption is wrong and it has smoothness. In addition, the results of white noise test also believe that the original hypothesis is wrong, and the conclusion that the sequence after first-order difference is not white noise is drawn .

Table 3. Unit root test

			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-0.395634	0.5345
Test critical values:	1% level		-2.636369	
	5% level		-1.951658	
	10% level		-1.610628	
Variable	Coefficient	Std. Error	t-Statistic	Prob.
X(-1)	-0.056981	0.142628	-0.395328	0.6982
D(X(-1))	-0.498217	0.183337	-2.719628	0.0125
D(X(-2))	-0.595395	0.223851	-2.663982	0.0131

4.2.2. Model identification and prioritization. The first step is to preprocess the time series data to obtain a smooth sequence. Figure 5 shows a partial autocorrelation plot for this shipment A, and we can see that both the autocorrelation plot and the partial autocorrelation plot show a trailing feature, so that the possible value range of the autocorrelation order p is between 1 and 2 orders, and the possible value range of the moving average order q is between 1 and 3 orders.

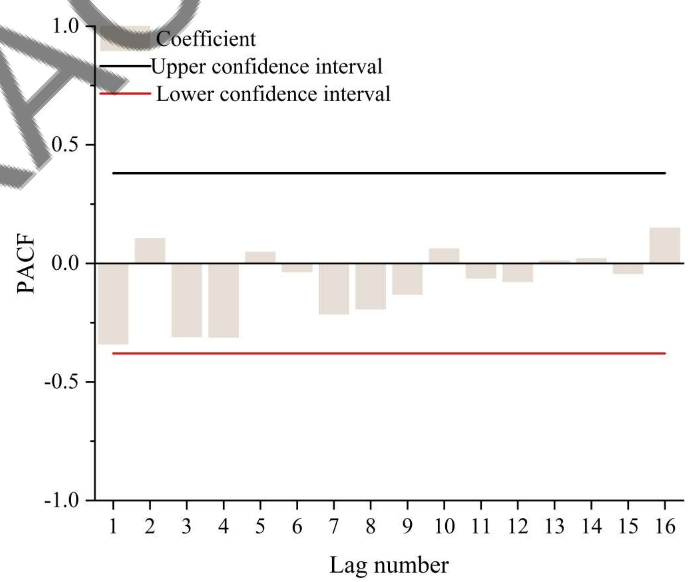

Fig. 5. A material actual shipping partial self-correlation diagram

AIC criterion is based on the concept of entropy, which can weigh the complexity of the estimated model and the superiority of the data fitted by the model, aiming to select a model that can fit the data well and has appropriate complexity, so as to avoid overfitting. We use AIC criterion to determine the optimal values of p and q, and we find that AIC reaches the minimum when p=2 and q=3. Based on this, we construct the ARIMA (2,1,3) model. In the preliminary modeling phase, we set the initial value of the parameter p to be less than zero, and then reconstructed the model. Finally, the parameters of the resulting model are detailed in Table 4 (where 2 represents the autoregressive order, 1 represents the difference order, and 3 represents the moving average order).

Table 4. Coefficient table

Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(1)	-0.6923548	0.145391	-4.746648	0.0000
AR(2)	-0.893984	0.1484218	-6.003163	0.0000
MA(3)	-0.681321	0.216982	-3.146151	0.0035
R-squared	0.465987	Mean dependent var		443.7162
Adjusted R-squared	0.431346	S.D. dependent var		24739.33
S.E. of regression	18662.28	Akaike info criterion		22.593921
Sum squared resid	1.09E+10	Schwarz criterion		22.72625
Log likelihood	-381.0329	Hannan-Quinn criter.		22.63726
Durbin-Watson stat	1.868398			
Inverted AR Roots	0.36 + 0.87 I	- 0.33-0.86 I		
Inverted MA Roots	0.89	- 0.45-0.77 I	0.43 + 0.75 I	

4.2.3. Model fitting and validation. We conducted residual white noise test on the constructed model. White noise test is a statistical test method used to judge whether the time series data are purely random, that is, whether the data in the series have no predictable pattern or trend. The test results are shown in FIG. 6. By observing the autocorrelation function (ACF) and partial autocorrelation function (PACF) graphs of the residuals, we can find that the first-order correspondences are both very small and do not exceed their respective confidence interval boundaries. Therefore, we judge that the residual sequence conforms to the characteristics of the white noise sequence, that is, there is no correlation. This indicates that our model performs well.

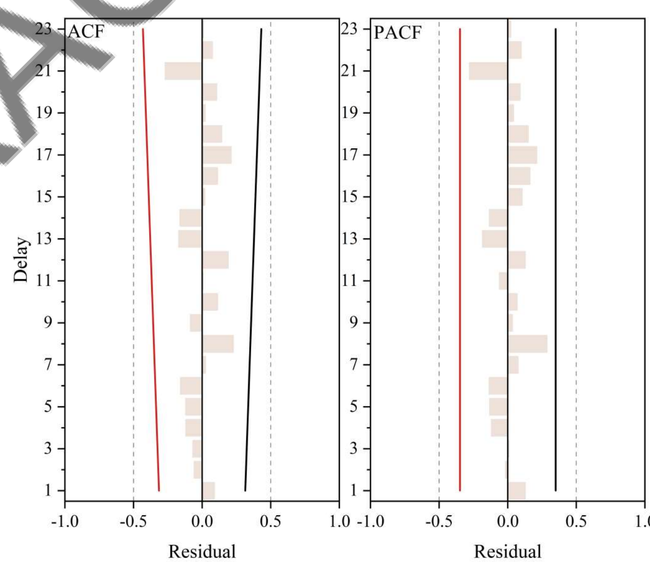

Fig. 6. Residual test

In the field of enterprise supply chain management, we understand the importance of accurately forecasting future demand. In order to more accurately reflect the actual performance of the traditional supply chain in the production and forecasting process, we have studied and optimized many times, and finally determined and applied the optimized ARIMA (2,1,3) model to forecast the probability distribution of demand in the coming year. ARIMA (2,1,3) model, as a classic time series analysis model, its parameter Settings have been carefully adjusted and verified by us to ensure that the model can fully capture the time series characteristics in the data. Through this model, we are able to make a detailed forecast of the demand trend in the coming year, which provides powerful data support for enterprises' supply chain decisions.

In order to verify the validity of the model, we conducted a comprehensive comparative analysis between the actual observed values and the predicted values of the model. After careful comparison, we were pleased to find that the model successfully predicted key trends in demand over the next 12 months. Both the overall demand trend and the specific demand at each point in time, the model was able to give forecasts that were highly consistent with the actual situation. More importantly, there is no significant deviation between the actual observed value and the predicted value. This result not only further validates the validity and accuracy of the model, but also strengthens our confidence in the model's predictive ability. We can be sure that the model has high reliability and practicability in predicting future demand.

Therefore, the ARIMA (2,1,3) forecasting model proposed in this paper has good reference value in terms of demand trend. It can not only provide powerful data support for enterprises' supply chain decision-making, but also help the raw material suppliers in the supply chain to grasp the market demand dynamics more accurately. Based on the forecast results of the model, raw material suppliers can prepare the corresponding raw materials in advance to ensure timely supply when the demand peak arrives, so as to avoid production interruption or delay caused by raw material shortage. The ARIMA (2,1,3) model we used has performed well in predicting future demand, providing strong support for supply chain management and raw material procurement of enterprises. We believe that in future practice, the model will continue to give play to its advantages and contribute more to the sustainable development of enterprises.

5. Multi-level inventory cost control optimization model based on supply chain demand forecast

5.1. Multi-level inventory cost model hypothesis

5.1.1. Model assumptions. The modeling assumptions of this paper are as follows:

1) Constancy of the order lead time: it is assumed that the order lead time is constant and does not change with time, which simplifies the logistics process and is convenient to accurately predict the inventory and meet the demand.

2) the fixed price of goods: the price of goods is fixed, and each level involves only one commodity, simplifying the price strategy, focusing on inventory management and demand forecasting.

3) Fixed unit inventory cost: Unit inventory cost is a fixed constant value, ignoring cost changes, which is convenient to analyze the impact of inventory cost on supply chain performance.

4) Irregular demand distribution: Daily demand follows Poisson distribution, captures the irregularity of demand, and reflects the volatility of market demand.

These assumptions not only conform to the characteristics of actual supply chain management, but also simplify the complexity of the model, enabling us to focus on the study of key issues and optimization strategies in the supply chain. Through these assumptions, we construct a supply chain management model with theoretical value and practical significance.

The model adopts a multi-layer structure, in which arrows indicate the direction of the flow of goods for supply, production, distribution and retail enterprises that are not on one level. Specifically, in the supply chain multi-level inventory cost model we constructed, nodes at layer k are able to send goods to any node at layer k+1, but no logistics activity occurs if the two nodes are on the same level or at different levels but far apart. We then view the manufacturing enterprise as a dividing line, with the part above the dividing line being the network for supplying goods and the part below the dividing line being the network for distributing goods. If layer k is a manufacturing enterprise, then all nodes from layer 1 to K-1 form the network for supplying goods to the manufacturing enterprise, and nodes from layer k+1 to the end form the network for distributing goods to the manufacturing enterprise. In the model construction and research of this paper, the problem we discuss is how to determine the optimal logistics path, so that the total annual inventory cost of the system can be minimized.

5.1.2. Parameter annotation of the model. The representation of supply chain multi-level inventory cost model in this section is divided into inventory cost, time cost model representation and other parameter model representation, as follows:

Pr_w -- Product price,. w $w = 1, 2, 3, \cdots\cdots, W$

Pr_{k-l} -- the price of the goods. $k-l$ $l = 1, 2, 3, \cdots\cdots, k-1$

$T_k^{(g,h)}$ -- The purchase period from the floor to the previous floor. k h g

$T_{k+l}^{(i,j,w)}$ -- Order the purchasing cycle from the previous layer to the previous layer. $k+l$ i j w

$Am_k^{(g,h)}$ -- Purchase cost of ordering to a higher tier. k h g

$Am_{k+l}^{(h,i,w)}$ - The purchase cost of the tier to the order of the previous tier. k h g w

$\lambda_k^{(g,h)}$ -- If, then the product purchased in a certain layer is relative to the previous layer. $\lambda_k^{(g,h)} = 1$ h k gDo not make a purchase. $\lambda_k^{(g,h)} = 0$

$\lambda_{k+l}^{(i,j,w)}$ -- If, layer before layer order, when not making a purchase. $\lambda_k^{(g,h)} = 1$ $k+1$ i j w $\lambda_k^{(g,h)} = 0$

$Sl_k^{(g,h)}$ -- k h g

$Ed_k^{(g,h)}$ -- The average number of goods required per day for each layer vs. the previous layer. k h g

$Lt_k^{(g,h)}$ - Delivery times required for tier purchase orders to higher levels. k h g

$Lt_{k+l}^{(i,j,w)}$ - The purchase lead time ordered compared to the previous tier. $k+l$ i j w

$Hm_k^{(g,h)}$ -- The cost of warehousing each item prior to the floor. k h g

$Hm_{k+l}^{(i,w)}$ -- Tier retains unit warehousing costs. $k+l$ i w

$X_{k-l}^{(g,h)}$ -- The amount of goods purchased between tiers is less than expected by the next tier. $k-l$ g $(T_{k-l}^{(i,j,w)} + Lt_{k-l}^{(g,h)})$ h

$X_{k+l}^{(i,j,w)}$ -- - $k-l$ i $(T_{k+l}^{(i,j,w)} + L_{k+l}^{(i,j,w)})$ j w

$Bm_h^{(g,h)}$ -- Layer The out-of-stock penalty for each item on the next layer. k h g

$Bm_{k+l}^{(i,j,w)}$ -- Out-of-stock penalty for the previous layer of each item. $k+l$ i j w

$Y_k^{(g,h)}$ -- Calculate, layer by layer, the cost of shipping time for each item at each point in time. k h $k+1$ g

$t_k^{(g,h)}$ -- Layer Indicates the transportation time of nodes in the next layer. k h g

t_k^h -- Average time required for the Layer to store goods. k h

C_k^h - The cost of the layer to manufacture at each point in time. k h

c_k^h -- The cost of delayed manufacturing at each point in time of the layer. k h

V_k^h -- the number of goods manufactured per point in time}. k h

v_k^h -- Layer number The number of goods delayed in production at each point in time. k h

L_k^h - Layer The cost of storing each item at each point in time. k h

$y_k^{(g,h)}$ - Layer each point in time delay shipping cost The next level of each commodity futures. k h g

$TP_k^{(g,h)}$ Store the cost of each point in time for the next level. k h g

5.2. Supply chain multi-level inventory model construction

In order to carry out in-depth internal analysis and derivation of inventory cost, this paper divides inventory cost into three components, namely procurement cost, inventory holding cost and time cost, and adopts centralized inventory control strategy and cyclic test inventory strategy (t, R, S), and finally uses mathematical expectation method to find the total inventory cost.

5.2.1. Procurement cost. The one-year procurement cost of each level is as follows: Specify the procurement cost of the core manufacturer:

$$O_m = \sum_{g=1}^{m_{k-1}} \sum_{h=1}^{m_k} \frac{Am_k^{(g,h)}}{T_k^{(g,h)}} \lambda_k^{(g,h)} \tag{9}$$

Procurement cost of designated supply network:

$$O_s = \sum_{l=1}^{k-2} \sum_{f=1}^{m_{k-l-1}} \sum_{g=1}^{m_{k-l}} \sum_{h=1}^{m_{k-l+1}} \frac{Am_{k-l}^{(f,g)}}{T_{k-l}^{(f,g)}} \lambda_{k-l}^{(f,g)} \lambda_{k-l+1}^{(g,h)} \tag{10}$$

Distribution network procurement costs are:

$$O_d = \sum_{l=1}^{N-k} \sum_{h=1}^{m_{k+l-2}} \sum_{i=1}^{m_{k+l-1}} \sum_{j=1}^{m_{k+l}} \sum_{w=1}^{W} \frac{Am_{k+l}^{(i,j,w)}}{T_{k+l}^{(i,j,w)}} \lambda_{k+l-1}^{(h,i)} \lambda_{k+l}^{(i,j,w)} \tag{11}$$

Then specify the total procurement cost for the entire system:

$$O = O_m + O_s + O_d \tag{12}$$

5.2.2. Inventory carrying cost. If the level of inventory at the time a Tier 1 business receives an order from a Tier 1 business is equal to the level of inventory before the order is received in the next ordering cycle, the average level of inventory is as follows: h k g $k-1$

$$\begin{aligned} &0.5 \times [Sl_k^{(g,h)} - Ed_k^{(g,h)} Lt_k^{(g,h)} + Sl_k^{(g,h)} - Ed_k^{(g,h)} Lt_k^{(g,h)} - Ed_k^{(g,h)} T_k^{(g,h)}] \\ &= Sl_k^{(g,h)} - Ed_k^{(g,h)} Lt_k^{(g,h)} - 0.5 \times Ed_k^{(g,h)} T_k^{(g,h)} \end{aligned} \tag{13}$$

The 1-year total inventory carrying cost at each level is as follows:
The cost of carrying inventory for core producers is as follows:

$$\begin{aligned} H_m = \sum_{g=1}^{m_{k-1}} \sum_{h=1}^{m_k} [&Sl_k^{(g,h)} - Ed_k^{(g,h)} Lt_k^{(g,h)} - \\ &0.5 \times Ed_k^{(g,h)} T_k^{(g,h)}](Hm_k^{(g,h)} + \Pr_{k-1}^{g}) \lambda_k^{(g,h)} \end{aligned} \tag{14}$$

Inventory carrying costs for the designated supply network:

$$\begin{aligned} H_s = \sum_{l=1}^{k-2} \sum_{f=1}^{m_{k-l}} \sum_{g=1}^{m_{k-l}} \sum_{h=1}^{m_{k-l+1}} [&Sl_{k-l}^{(f,g)} - Ed_{k-l}^{(f,g)} Lt_{k-l}^{(f,g)} - \\ &0.5 \times Ed_{k-l}^{(f,g)} T_{k-l}^{(f,g)}](Hm_{k-l}^{(f,g)} + \Pr_{k-l-1}^{f}) \lambda_{k-l}^{(f,g)} \lambda_{k-l+1}^{(g,h)} \end{aligned} \tag{15}$$

The inventory carrying cost for the distribution network is:

$$H_d = \sum_{l=1}^{N-k} \sum_{i=1}^{m_{k+l-1}} \sum_{j=1}^{m_{k+l}} \sum_{h=1}^{m_{k+l-2}} \sum_{w=1}^{W} [Sl_{k+l}^{(j,w)} - Ed_{k+l}^{(j,w)} Lt_{k+l}^{(i,j,w)} - 0.5 \times Ed_{k+l}^{(j,w)} T_{k+l}^{(i,j,w)}](H\eta_{k+l}^{(i,w)} + \mathrm{Pr}_w) \lambda_{k+l}^{(i,j,w)} \lambda_{k+l-1}^{(h,j)} \tag{16}$$

Then specify the inventory carrying cost for the entire system:

$$H = H_m + H_s + H_d \tag{17}$$

5.2.3. Time cost. The time cost that needs to be consumed in the product preparation time, such as the labor cost of workers, preparation cost, equipment cost, etc. If the delivery is delayed, the transportation process from the supply company to the production company will be penalized. If the delivery is delayed, the transportation process from the manufacturing business to the distribution business will also be penalized. The cost of temporary storage of goods by the manufacturing business. The cost of time is made up of the components mentioned above, and they all need to be converted into costs.

The costs involved in stocking up are as follows:

$$T_p = \sum_{g=1}^{m_{k-i}} \sum_{h=1}^{m_k} TP_k^{(g,h)} Ed_k^{(g,h)} \frac{\lambda_k^{(g,h)}}{T_k^{(g,h)}} \tag{18}$$

The total time cost of each layer for 1 year is as follows:

The cost of manufacturing enterprises is as follows:

$$T_m = \sum_{g=1}^{m_{k-l}} \sum_{h=1}^{m_k} [(C_k^h \times (Ed_k^{(g,h)} - v_k^h) / V_k^h + (c_k^h \times v_k^h) / V_k^h + Ed_k^{(g,h)} \times t_k^h \times L_k^h + Ed_k^{(g,h)} \times y_k^{(g,h)} \times t_k^h + KEd_k^{(g,h)} Lt_k^{(g,h)}) \times \frac{\lambda_k^{(g,h)}}{T_k^{(g,h)}}] \tag{19}$$

These include:

$$Ed_k^{(g,h)} = \sum_{i=1}^{m_{k+1}} \sum_{w=1}^{W} Ed_{k+1}^{(h,i,w)} \lambda_{k+1}^{(h,i,w)} \tag{20}$$

The cost of the supply network is specific as:

$$T_s = \sum_{l=1}^{k-2} \sum_{f=1}^{m_{k-l-1}} \sum_{g=1}^{m_{k-l}} \sum_{h=1}^{m_{k-l+1}} [(Y_k^{(g,h)} \times Ed_{k-l}^{(f,g)} \times t_k^{(g,h)} + Ed_{k-l}^{(f,g)} \times y_k^{(g,h)} \times t_k^h + KEd_{k-l}^{(f,g)} Ll_{k-l}^{(f,g)}) \frac{\lambda_{k-l}^{(f,g)} \lambda_{k-l+1}^{(g,h)}}{T_{k-l}^{(f,g)}}] \tag{21}$$

The cost of the distribution network is specific to:

$$T_d = \sum_{l=1}^{N-k} \sum_{i=1}^{m_{k+l-1}} \sum_{j=1}^{m_{k+l}} \sum_{h=1}^{m_{k+l-2}} \sum_{w=1}^{W} [Y_k^{(g,h)} \times Ed_{k+l}^{(j,w)} \times t_k^{(g,h)} + Ed_{k+l}^{(j,w)} \times y_k^{(g,h)} \times t_k^h + KEd_{k+l}^{(j,w)} Lt_{k+l}^{(i,j,w)}) \frac{\lambda_{k+l}^{(i,j,w)} \lambda_{k+l-1}^{(h,j)}}{T_{k+l}^{(i,j,w)}}] \tag{22}$$

The cost of the entire system is specific as:

$$T = T_p + T_m + T_s + T_d \tag{23}$$

To sum up, the total inventory cost of the system is as follows:

$$TC = \alpha(O + H + Q) + \beta T \tag{24}$$

The minimum inventory cost is specified as:

$$\min TC = \alpha(O + H + Q) + \beta T \tag{25}$$

The constraints in the model are specified as follows:

$$\begin{aligned}
&\alpha + \beta = 1; \sum_{g=1}^{m_{k-1}} \lambda_k^{(g,h)} = 1 \\
&\sum_{f=1}^{m_{k-l-1}} \lambda_{k-l}^{(f,g)} = 1, g = 1, 2, \ldots, m_{k-l}; l = 1, 2, \ldots, k-2; \\
&\sum_{i=1}^{m_{k+l-1}} \lambda_{k+l}^{(i,j)} = 1, j = 1, 2, \ldots, m_{k+l}; l = 1, 2 \ldots, N-k
\end{aligned} \tag{26}$$

5.3. Solving steps of immune genetic algorithm

The basic steps of immune genetic algorithm solution are as follows [31] :

1) Identify the antigen. Antigens for immune genetic algorithms include objective functions and various constraints in the model.

2) Generate the initial antibody population. Within a certain range, according to the characteristics of the desired problem, according to the corresponding antibody coding method, the initial antibody population is randomly generated one by one, and then the generated antibodies are tested to see if they meet the requirements, if one antibody does not meet the requirements, new antibodies need to be generated until all antibodies meet the requirements.

3) Calculate the adaptive value of the antibody. The adaptive value of the antibody is related to the objective function, when the problem is maximized, the adaptive value of the antibody is equal to the value of the objective function, and when the problem is minimized, the value of the objective function is converted to the adaptive value, so as to ensure the maximum adaptive value of the optimal individual.

4) Update the memory unit. Add antibodies from the original high-affinity population to the memory cells, because the population size of memory cells is limited, replace the antibodies in the original memory cells with antibodies from the high-affinity population.

5) Genetic manipulation. New antibodies are produced to the next generation through selection, crossover, and mutation operations in gene manipulation. After the two antibodies are selected and the corresponding mutation operation is performed according to the predetermined mutation probability, the crossover operation is performed.

6) Immunization. Immunization is mainly the suppression of the same or similar number of antibodies produced by genetic manipulation using concentration control method. In the optimization process, when the concentration of a certain antibody in the population is too high, the algorithm is easy to fall into local optimization, resulting in prematurity. Repeat steps 3 to 6 until the set termination conditions are met.

7) Output results. From the last generation of antibodies in the population, select the antibody with the highest ability to bind to the antigen as the optimal solution to the optimization problem.

6. Case application analysis

In the context of increasingly fierce market competition, the importance of enterprise supply chain management has become increasingly prominent. Through a detailed analysis of an actual case of an auto parts manufacturing company, this section focuses on relevant optimization strategies based on machine learning models, shows how to use machine learning models to optimize supply chain demand forecasting and inventory cost control, and deeply analyzes the application effect and value of this model method to improve the overall operational efficiency of enterprises. This case not only demonstrates the potential of data-driven forecasting methods in improving supply chain efficiency, but also further validates the application value of multi-level inventory cost model in practical operations.

6.1. Data processing and experimental setup

In this study, we selected a modern enterprise integrating R&D, production, sales and service of auto parts as the research object. In order to construct an accurate prediction model, we collected the product sales data of two major customers of the enterprise for 6 consecutive months in 2023, and divided them into training sets and test sets according to the ratio of 8:2 to ensure the generalization ability of the model. The forecast length is set to n=8, that is, the product demand in the next 8 days is predicted.

When solving the multi-level inventory model, we use genetic algorithm, and set individual parameters carefully, which can balance the search efficiency and solving accuracy of the algorithm, and ensure that the inventory strategy obtained can meet the requirements of service level and realize cost optimization. The population size is set to 50, the number of iterations is 500, and the crossover probability and mutation probability are adjusted to 0.8 and 0.1, respectively, to ensure the effective convergence of the algorithm. In addition, we set the maximum number of order batches for manufacturers to order raw materials and customers to order finished products at 4,000, as well as the minimum service level for manufacturers and customers at 95 percent to meet the actual operational needs of enterprises.

6.2. Solve the model and perform the analysis

MATLAB software was used to realize the algorithm solving process, and the solution results as shown in Table 5 were obtained. The algorithm solving process is realized by using MATLAB software. The results show that the order number of the company's raw material supplier is 2598, 2835, 3313 pieces, and the safety factor of the manufacturer's raw material is 2.82, 2.45, 2.09 in turn; The order batches of customers to the company were 3973 and 3995 pieces, and the safety coefficients of customer products were 2.63 and 2.75, respectively. At this time, the optimal solution of the objective function (that is, the total inventory cost per day) was 3891.3 yuan/day. This result shows that under the current parameter setting, the inventory strategy obtained by machine learning model and genetic algorithm can achieve relatively low inventory cost on the basis of meeting the service level.

Table 5. The result of the solution

Enterprise	Order batch/piece	Safety coefficient
Manufacturer's raw materials 1	2835	2.45
Manufacturer's raw materials 2	2598	2.82
Manufacturer's raw materials 3	3313	2.09
Client 1	3995	2.75
Client 2	3973	2.63

It is worth noting that in actual operation, the delivery time is often affected by a variety of factors and presents uncertainty. In order to be closer to the realistic situation, we further consider the impact of the uncertainty of shipping time and the difference in order processing time of the parent company on the delivery time. Among them, the transportation time from supplier to manufacturer and from manufacturer to customer follows the gamma distribution with parameters, and the order processing time is only related to the order batch size. (α,β) Based on this, the formula for calculating the average order delivery time of the manufacturer and the customer is derived:

$$U_L = \frac{\alpha}{\beta} + \varepsilon Q^* \tag{27}$$

The formula represents the order processing time parameter, which is related to the order batch of the next node of the enterprise. When the order batch is fixed, the order processing time is a fixed value. ε The variance of the order lead time of the manufacturer and the customerL2 is σ=β/2α. When the delivery time and demand are not fixed, the order point of the manufacturer and the order point of the customer are calculated:

$$R_{mi} = \bar{d}_{mi}\mu_L + z_{mi}\sqrt{\mu_L {\sigma_{mi}}^2 + (\bar{d}_{mi}\sigma_L)^2} \tag{28}$$

$$R_{cj} = \bar{d}_{cj}\mu_L + z_{cj}\sqrt{\mu_L {\sigma_{cj}}^2 + (\bar{d}_{cj}\sigma_L)^2} \tag{29}$$

Set the time parameter for order processing to ε=2.0×10-4, and the shipping time follows the gamma distribution of parameters α=0.25 and β=0.75. Genetic algorithm is used to solve the problem, and the optimal inventory cost is 4121.3 yuan/day. The gamma-ray distribution parameters satisfied by the transportation time are shown above. The transportation time parameters are calculated, and the total inventory cost and the order point between the manufacturer and the customer are solved according to the calculation results. The solution results of the inventory model under different parameters are shown in Table 6.

Table 6. The solution of the inventory model under different parameters

No.	R_{m1} /item	R_{m2} /item	R_{m3} /item	R_{c1} /item	R_{c2} /item	TC yuan/day
1	1822	1975	1919	692	1308	4063.8
2	2971	2140	2315	863	1612	4124.6
3	3070	3090	3103	1441	2814	4225.4
4	5781	5384	6044	1988	3788	4524.5
5	10278	11155	10473	4892	6843	5144.8
6	1694	1452	1558	653	1238	4014.6
7	2086	1707	1967	815	1650	4067.7
8	3406	2411	2561	989	1647	4129.4
9	3166	3424	3272	1446	2056	4237.8
10	4586	4343	4453	1511	3410	4343.4
11	8069	9125	7815	2742	4934	4765.1

When the fixed delivery time is 2 days, the optimization result of the total inventory cost per unit time obtained by genetic algorithm is 3891.3 yuan/day; When the delivery time is not fixed, the optimization result is 4063.8 yuan/day. Compared with the company's original inventory strategy, the total inventory cost per unit time is 4,708.39 yuan/day. The optimization strategy based on the machine learning model reduces the total inventory cost of the supply chain by 17.35% and 13.69% respectively under the premise of ensuring the company's minimum service level. From the theoretical analysis, it can be seen that with the increase of α/β value, the delivery time of manufacturers and customers' orders is extended, and the demand increases throughout the delivery period. In order to meet the demand, manufacturers and customers need to increase the order point. When the α/β^2 value increases, the enterprise needs to set up a larger safety inventory to deal with the uncertainty of the lead time. In addition, we also notice that the solution results of the inventory model under different parameters show a certain regularity. As the α/β and α/β^2 values increase, the total cost of inventory rises, and so do the order points for manufacturers and customers. This fully reflects the significant advantages of the multilevel inventory cost model built based on machine learning model in practical application, with high practical significance and application value, for enterprises to effectively reduce costs and improve the efficiency of supply chain management provides a strong theoretical support and practical guidance.

7. Conclusion

This study mainly realizes the demand prediction of enterprise supply chain through machine learning algorithm, constructs the inventory cost control optimization model based on supply chain demand, and solves the inventory cost control model through genetic algorithm, so as to realize the optimal allocation of supply chain inventory. White noise test is carried out on the residual error after modeling, and the corresponding value of the first order is relatively small, and does not exceed the confidence interval boundary, which indicates that the residual error of the model is not correlated. At the same time, the trend predicted by the model is closer to the real value, indicating that the predictive modeling is more effective. The simulation experiment of case application is carried out with MATLAB. In the case of fixed delivery time of 2 days, the total inventory cost of supply chain is reduced by 17.35% and 13.69% respectively after adopting the method in this paper, which reflects the optimization effect of the new method on inventory cost control. In other words, when the probability of an emergency is constant, the prediction results obtained by the model are not accurate enough. Therefore, it is necessary to continue to optimize the model in the follow-up research.

Based on the above research, this paper successfully constructs an optimization framework for enterprise supply chain demand forecasting and inventory cost control based on machine learning model, and verifies its effectiveness and feasibility through empirical analysis. This framework not only improves the accuracy of supply chain demand forecasting, but also optimizes inventory cost effectively, providing a new idea and method for enterprise supply chain management.

From the point of view of enterprise management, the importance of machine learning technology in supply chain management is emphasized. In the era of big data, enterprises are faced with massive data and information, how to effectively use these data to improve the management efficiency of supply chain has become an urgent problem to be solved. This research proves through practice that the machine learning model can deeply dig the potential law in the data, and provide strong support for the supply chain demand forecast and inventory control of enterprises. In addition, this study provides specific implementation paths and methods for enterprises. By building a machine learning model, enterprises can more accurately predict the market demand, so as to formulate more reasonable production plans and inventory strategies. This can not only reduce the inventory cost of enterprises, but also improve the response speed and flexibility of the supply chain, and enhance the market competitiveness of enterprises. Inventory cost is an important part of enterprise operating cost. How to effectively control inventory cost is of great significance for improving enterprise profitability and market competitiveness. By optimizing the inventory control strategy, this study realizes the effective reduction of inventory cost, and provides a new idea and method for the enterprise's cost management.

The research of this paper not only provides a new theory and method for supply chain demand forecasting and inventory cost control, but also provides beneficial enlightenment for enterprise management. It tells us that in the era of big data and machine learning, enterprises should make full use of these advanced technologies to improve the efficiency of supply chain management and optimize inventory costs, thus enhancing the market competitiveness and profitability of enterprises. We believe that with the continuous progress of technology and the deepening of application, supply chain demand forecasting and inventory cost control optimization based on machine learning model will play a more important role in enterprise management.

References

[1] Li, G., Wang, X., Su, S., & Su, Y. (2019). How green technological innovation ability influences enterprise competitiveness. Technology in Society, 59, 101136.

[2] Ni, G., Xu, H., Cui, Q., Qiao, Y., Zhang, Z., Li, H., & Hickey, P. J. (2020). Influence mechanism of organizational flexibility on enterprise competitiveness: The mediating role of organizational innovation. Sustainability, 13(1), 176.

[3] Lee, H. H., Zhou, J., & Wang, J. (2018). Trade credit financing under competition and its impact on firm performance in supply chains. Manufacturing & Service Operations Management, 20(1), 36-52.

[4] Zhang, L., Xu, T., Zhu, H., Qin, C., Meng, Q., Xiong, H., & Chen, E. (2020, April). Large-scale talent flow embedding for company competitive analysis. In Proceedings of The Web Conference 2020 (pp. 2354-2364).

[5] Chauhan, C., & Singh, A. (2020). A review of Industry 4.0 in supply chain management studies. Journal of Manufacturing Technology Management, 31(5), 863-886.

[6] Qi, Y., Huo, B., Wang, Z., & Yeung, H. Y. J. (2017). The impact of operations and supply chain strategies on integration and performance. International Journal of Production Economics, 185, 162-174.

[7] Hong, J., Zhang, Y., & Ding, M. (2018). Sustainable supply chain management practices, supply chain dynamic capabilities, and enterprise performance. Journal of Cleaner Production, 172, 3508-3519.

[8] Tolonen, A., Haapasalo, H., Harkonen, J., & Verrollot, J. (2017). Supply chain capability creation–The creation of the supply chain readiness for a new product during product development process. International Journal of Production Economics, 194, 237-245.

[9] Reefke, H., & Sundaram, D. (2017). Key themes and research opportunities in sustainable supply chain management–identification and evaluation. Omega, 66, 195-211.

[10] Um, K. H., & Kim, S. M. (2019). The effects of supply chain collaboration on performance and transaction cost advantage: The moderation and nonlinear effects of governance mechanisms. International Journal of Production Economics, 217, 97-111.

[11] Benton Jr, W. C. (2020). Purchasing and supply chain management. Sage Publications.

[12] Singh, D., & Verma, A. (2018). Inventory management in supply chain. Materials Today: Proceedings, 5(2), 3867-3872.

[13] Alrjoub, A. M. S., & Ahmad, M. A. (2017). Inventory management, cost of capital and firm performance: evidence from manufacturing firms in Jordan. Investment management and financial innovations, (14,№ 3), 4-14.

[14] Milewski, D. (2020). Total costs of centralized and decentralized inventory strategies—including external costs. Sustainability, 12(22), 9346.

[15] Bensoussan, A., Skaaning, S., & Turi, J. (2018). Inventory control with fixed cost and price optimization in continuous time. J. Appl. Anal. Comput, 8, 805-835.

[16] Alfares, H. K., & Ghaithan, A. M. (2022). A generalized production-inventory model with variable production, demand, and cost rates. Arabian Journal for Science and Engineering, 47(3), 3963-3978.

[17] Palanivel, M., & Uthayakumar, R. (2017). A production-inventory model with promotional effort, variable production cost and probabilistic deterioration. International Journal of System Assurance Engineering and Management, 8, 290-300.

[18] Bhunia, A. K., Shaikh, A. A., & Cardenas-Barron, L. E. (2017). A partially integrated production-inventory model with interval valued inventory costs, variable demand and flexible reliability. Applied Soft Computing, 55, 491-502.

[19] Pando, V., San-Jose, L. A., & Sicilia, J. (2020). A new approach to maximize the profit/cost ratio in a stock-dependent demand inventory model. Computers & Operations Research, 120, 104940.

[20] Kilimci, Z. H., Akyuz, A. O., Uysal, M., Akyokus, S., Uysal, M. O., Atak Bulbul, B., & Ekmis, M. A. (2019). An improved demand forecasting model using deep learning approach and proposed decision integration strategy for supply chain. Complexity, 2019(1), 9067367.

[21] Feizabadi, J. (2022). Machine learning demand forecasting and supply chain performance. International Journal of Logistics Research and Applications, 25(2), 119-142.

[22] Fu, W., & Chien, C. F. (2019). UNISON data-driven intermittent demand forecast framework to empower supply chain resilience and an empirical study in electronics distribution. Computers & Industrial Engineering, 135, 940-949.

[23] Zhu, X., Ninh, A., Zhao, H., & Liu, Z. (2021). Demand forecasting with supply‐chain information and machine learning: Evidence in the pharmaceutical industry. Production and Operations Management, 30(9), 3231-3252.

[24] Kondapaka, K. K. (2021). Advanced Artificial Intelligence Techniques for Demand Forecasting in Retail Supply Chains: Models, Applications, and Real-World Case Studies. African Journal of Artificial Intelligence and Sustainable Development, 1(1), 180-218.

[25] Yu, W., Hou, G., & Li, J. (2019). Supply chain joint inventory management and cost optimization based on ant colony algorithm and fuzzy model. Tehniki vjesnik, 26(6), 1729-1737.

[26] Gurtu, A. (2021). Optimization of inventory holding cost due to price, weight, and volume of items. Journal of Risk and Financial Management, 14(2), 65.

[27] Wang, Y., Geng, X., Zhang, F., & Ruan, J. (2018). An immune genetic algorithm for multi-echelon inventory cost control of IOT based supply chains. IEEE Access, 6, 8547-8555.

[28] Abolghasemi, M., Beh, E., Tarr, G., & Gerlach, R. (2020). Demand forecasting in supply chain: The impact of demand volatility in the presence of promotion. Computers & Industrial Engineering, 142, 106380.

[29] IgnacioIllan Gala. (2025). Alzheimer's and Related disorders Multicenter Archive (ARMA): a collaborative platform to advance ADRD diagnosis and treatment in Spain. Alzheimer's & Dementia(S6),e085813-e085813.

[30] Lund University,Maastricht University,Department of Economics, Lund University & Centre for Financial Econometrics, Deakin University. (2019). Lag truncation and the local asymptotic distribution of the ADF test for a unit root. Statistical Papers(6),2109-2118.

[31] Hamedi Amir Hossein,Abolghasemi Hossein,Shokri Saeid,Nia Hadi Jafar & Moayedi Farshad. (2023). Integrating Artificial Immune Genetic Algorithm and Metaheuristic Ant Colony Optimizer with Two-Dose Vaccination and Modeling for Residual Fluid Catalytic Cracking Process. Arabian Journal for Science and Engineering(12),16329-16341.