CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

In today's fast-paced, high-demand, and competitive markets, efficient supply chain management ensures smooth progress in the seamless flow of products and customer services. The latest technologies of calculations and forecasting have enabled enterprises to adopt predictive modeling techniques to address challenges such as inventory shortages, demand fluctuations, and logistics optimization. However, despite there are advances in technologies used, many enterprises still face the problems of supply chain systems that are difficult to adapt to volatile and unstable factors. It leads to handling some specific issues hardly, and it will collapse after continuing to snowball. This literature review will refer to the theme of delayed prediction of inventory shortages in sports equipment supply chains to explore the latest methods of predictive modeling in supply chain management. This chapter will focus on traditional models and machine learning methods used for this project. By reviewing these methods, this study aims to identify existing gaps to improve the resilience and efficiency of supply chain management.

2.2 State-of-the-Art Approaches

This section provides an overview of the state-of-the-art methods for delay prediction on inventory shortage. These methods are divided into two categories, which are traditional methods and machine learning methods. Each model has its strengths and limitations, and understanding these models will help in selecting or combining techniques to achieve high prediction accuracy.

2.2.1 Traditional Methods

Traditional methods usually use mathematical and statistical models and make predictions based on historical data. The traditional method used in this project to solve the inventory shortage in the sports equipment supply chain is the autoregressive integrated moving average model (ARIMA).

2.2.1.1 ARIMA (Autoregressive Integrated Moving Average):

ARIMA was developed by the founders George Box and Gwilym Jenkins for time series forecasting in the 1970s. It is also called the Box-Jenkin Method. The ARIMA model is the most suitable option to capture the linear trends between the time series model and the time series' past values. Therefore, the ARIMA model is commonly used in econometrics to predict event development using relevant data such as the logistics supply chain data used for this project. The ARIMA is used as a combination of the autoregressive model (AR) and moving average model (MA). The autoregressive model is used to decide the order in the current value influenced by the lag observation numbers or find the p-value from the partial autocorrelation function (PACF) graph. Likewise, the moving average model can describe the order in lagged forecast errors or look for the q-value from the autocorrelation function (ACF) graph. Based on the graph needed, the ARIMA model will combine three parameters: p-value, q-value, and d-value. The d-value is the different value that makes the time series from non-stationary data into stationary data. In conclusion, the ARIMA model focuses on the relationship between autocorrelation and linear dependencies. The following are the equations and explanations of the ARIMA model:

- $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + \theta_1 \epsilon_{t-1} + ...$

  This is the ARIMA model's equation, where

  - $\phi_1, \phi_2, ...$: Autoregression coefficients (AR) to define the influence from past values to current value

- X_t: The value at the time t
- $X_{t-1}, X_{t-2}, \ldots$: The past value of time series (autoregressive component)
- θ_1, θ_2: Moving average coefficients (MA) to revise the predictions from past errors
- ϵ_t: Random noise (white noise)
- $\epsilon_{t-1}, \epsilon_{t-2}$: Past error terms (moving average component)

2.2.2 Machine Learning Methods

Machine learning methods help to customize the nonlinear data modeling by using the data-driven method. These methods are more resilient than traditional models and can handle large data sets with multiple variables. The machine learning methods used in this project to solve inventory shortages in the sports equipment supply chain are the seasonal autoregressive integrated moving average model (SARIMA), extreme gradient boosting model (XGBoost), long short-term memory model (LSTM), and a hybrid model consisting of autoregressive integrated moving average and extreme gradient boosting (ARIMA+XGBoost).

2.2.2.1 SARIMA (Seasonal Autoregressive Integrated Moving Average):

SARIMA is an extension of ARIMA developed during the same period by the founders George Box and Gwilym Jenkins to take seasonality into account, which adds the seasonal autoregressive function to handle the periodic pattern. The periodic pattern is also known as the symbol 's', measured in month units, and increased as consideration to the ARIMA model to become (p, d, q, s). The following are the equations and explanations of the SARIMA model:

- $X_t = \phi(B)X_t + \phi(B^S)X_t + \theta(B)\epsilon_t + \theta(B^S)\epsilon_t$

  This is the SARIMA model's equation, where

- B: Lag operator ($BX_t = X_{t-1}$)
- $\phi(B)$: Regular autoregressive terms
- $\phi(B^s)$: Seasonal autoregressive terms
- $\theta(B)$: Regular moving average terms
- $\theta(B^s)$: Seasonal moving average terms
- ϵ_t: Error term

2.2.2.2 XGBoost (Extreme Gradient Boosting):

XGBoost was developed by its founder Tianqi Chen in 2014 to optimize the residual problem to achieve the highest delay prediction accuracy. XGBoost is a nonlinear model that forecasts complex patterns using gradient boosting and the decision trees method to handle complex feature interactions. In this project, XGBoost will help to modify those factors that will be influencing the inventory system such as order volume, shipping time, and weather. The following are the equations and explanations of the XGBoost models:

- $y_i = \sum_{k=1}^{K} f_k(x_i), f_k \in F$

  This is the XGBoost model's equation, where

  - y_i: Predicted value for the i^{th} sample
  - f_k: PRediction from the k^{th} tree
  - F: Space of all possible decision tree

- $Obj = \sum_{i=1}^{n} l(y_i, \widehat{y}_i) + \sum_{k=1}^{K} \Omega(f_k)$

  This is the XGBoost model's equation, where

  - $l(y_i, \widehat{y}_i)$: Loss function

- $\Omega(f_k)$: Regularization terms for controlling model complexity

2.2.2.3 LSTM (Long Short-Term Memory):

LSTM was proposed by Sepp Hochreiter and Jürgen Schmidhuber in 1997. It is a special recurrent neural network (RNN) that can reduce the problem of gradient vanishing and solve the problem that the original RNN has a weak ability to handle long-term dependencies, such as time series and language. LSTM usually uses gates to determine whether to keep or discard the data information. The core operations of LSTM are divided into four gate types, which are the forget gate, input gate, cell state update, and output gate. In this project, LSTM can capture historical inventory changes, reduce the problems caused by incomplete data and missing values on memory calculations, improve the accuracy of predicting inventory shortage, and optimize supply chain efficiency. The following are the explanations and equations of the LSTM gates:

- Forget Gate: Determine the amount of past information to forget $f_t = \sigma(W_f \bullet [h_{t-1}, x_t] + b_f)$
- Input Gate: Decide the amount of new information added

  $i_t = \sigma(W_i \bullet [h_{t-1}, x_t] + b_i), \overline{C}_t = tanh(W_c \bullet [h_{t-1}, x_t])$

- Cell State Update: Capture long-term memory for the sequence

  $C_t = f_t \odot C_{t-1} + i_t \odot \overline{C}_t$

- Output Gate: Outputs relevant information for the next step

  $o_t = \sigma(W_o \bullet [h_{t-1}, x_t] + b_o), h_t = o_t \odot tanh(C_t)$

  This is the LSTM model's equation, where

  - x_t: Input at time
  - h_t: Hidden state at the time t

- C_t: Memory cell value
- W_f, W_i, W_c, W_o: Weight matrices
- σ: Sigmoid activation function
- ⊙: Element-wise multiplication

2.2.2.4 Hybrid Model (ARIMA + XGBoost):

The hybrid model is combined with the ARIMA model and XGBoost model. The ARIMA model forecasts may have errors and the residuals may be related to order volume, shipping time, and weather. Hence, the XGBoost model will help to handle these nonlinear effects. By combining the two models, the inventory forecasting issues can be captured for robust prediction whether both linear and nonlinear patterns. The following are the equations and explanations of the Hybrid model:

- $\widehat{y}_t = \widehat{y}_{ARIMA} + \widehat{y}_{XGBoost}$

  This is the Combine Model model's equation, where

  - $\widehat{y}_{ARIMA}$: Forecast from the ARIMA model
  - $\widehat{y}_{XGBoost}$: Forecast from the XGBoost model

2.2.3 Real-Time and Predictive Analytics in Supply Chain

This section will explain how all of the above methods can be applied to real-time and predictive analytics in the supply chain. First, the real-time analytics is processing the data immediately when it is generated. It allows the enterprises to monitor operational changes, track inventory situations, and make respond or take action once there is any changes. On the other hand, the predictive analytics uses historical and real-time data to predict future events, providing data evidence for demand forecasting, risk management, and logistics optimization such as the possibility of potential demand disruptions or fluctuations. On a technical level, both real-time and predictive analytics are crucial. For example, predictive analytics can predict inventory shortages and propose the risks reduction solution, while real-time analytics can ensure the accuracy and timeliness of predictions. Although both have advantages, there are still a lot of challenges that have to be addressed in the aspects of data collection, data integration, data authenticity, and model complexity. In summary, the combination of the above methods can provide solutions to overcome these challenges and improve supply chain performance.

2.3 Limitation

Despite the advancements in predictive analytics and real-time monitoring, several limitations hinder their effectiveness in supply chain management. Traditional methods, such as ARIMA and SARIMA, are constrained by their reliance on linear and stationary assumptions, making them unsuitable for handling the non-linear and complex patterns observed in modern supply chains. Furthermore, these methods lack the capability to process high-dimensional datasets with multiple influencing factors.

Machine learning techniques, including XGBoost and LSTM, offer more flexibility in capturing non-linear relationships and long-term dependencies. However, they come with their own set of challenges. These models require large datasets and significant computational resources, which can limit their applicability for small- to medium-sized enterprises with resource constraints. Moreover, the interpretability of machine learning models often remains a concern, making it difficult for decision-makers to trust and adopt their predictions fully.

Hybrid approaches, which combine traditional statistical methods with machine learning models, have shown promise in improving forecasting accuracy. However, their implementation can be complex and time-intensive, often requiring specialized expertise. Additionally, the integration of these models into existing supply chain systems is not seamless, posing challenges in real-world deployment.

Data-related issues also pose significant limitations. Supply chain data is often noisy, incomplete, or sparsely distributed, reducing the reliability of predictions. Real-time analytics further amplifies these challenges, as it demands high-speed processing and accuracy under time-sensitive conditions.

2.4 Research Gap

While substantial progress has been made in applying predictive analytics and real-time modeling to supply chain management, several research gaps remain. One significant gap is the limited focus on holistic solutions that integrate multiple facets of the supply chain, such as demand forecasting, inventory management, and logistics optimization. Current research often addresses these components in isolation, resulting in fragmented solutions that fail to optimize the supply chain as a whole.

Another gap lies in the real-time adaptability of predictive models. Many studies develop models under static or controlled conditions, which do not adequately account for the dynamic and uncertain nature of real-world supply chains. There is a need for predictive systems that can adapt to evolving patterns and disruptions in real time, ensuring their relevance and accuracy.

The issue of data quality also presents a critical research gap. Existing studies have not sufficiently addressed the challenges posed by noisy, sparse, or incomplete datasets, nor have they explored robust preprocessing techniques to improve data usability. Additionally, there is limited research on how to handle data integration from multiple sources in real-time, which is essential for creating a unified and efficient predictive system.

Lastly, while hybrid approaches have shown potential, there is little research on standardizing their implementation or evaluating their performance under diverse scenarios. Future studies could explore how to simplify the integration of hybrid models and enhance their scalability for widespread adoption in the supply chain industry.

2.5 Definition of Terms

Term	Definition
Lag observation	The relationship between a current data point and its past values (lagged observations). It represents how the past influences the present or future in a data sequence. The "lag" refers to the number of time steps back from the current observation.
Partial autocorrelation function	The measurement of correlation between time series and its lagged values after removing the effects of intermediate lags
Autocorrelation function	The measurement of correlation between time series and its lagged values

2.6 Summary

The literature review highlights the evolution of methodologies used for predictive analytics and real-time decision-making in supply chain management. Traditional methods, such as ARIMA and SARIMA, have been widely applied due to their simplicity and interpretability. However, their inability to handle non-linear patterns and high-dimensional data limits their effectiveness in complex supply chain scenarios. Machine learning techniques, including XGBoost and LSTM, have emerged as powerful alternatives, offering flexibility in capturing intricate relationships and long-term dependencies. Despite their advantages, these methods require extensive data, computational resources, and expertise, which can hinder their broader adoption.

Hybrid approaches that combine traditional and machine learning techniques provide a promising avenue for addressing the limitations of individual models. These methods leverage the strengths of each approach to improve forecasting accuracy and adaptability. However, challenges such as integration complexity, data sparsity, and real-time implementation remain significant barriers to their practical use.

The review also identifies critical research gaps, including the need for holistic solutions that integrate multiple supply chain components, the development of models that can adapt dynamically to real-world conditions, and strategies for addressing data quality issues. Addressing these gaps through innovative methodologies and robust system designs could enhance the efficiency and resilience of supply chain management systems.