Vaasan yliopisto
UNIVERSITY OF VAASA

Shazid Hasan

Integrating Machine Learning and Scenario Planning for Enhanced Demand Forecasting in Bangladesh's Garment Industry

School of Technology and Innovation
Master's Thesis in Industrial Management programme

Vaasa 2026

UNIVERSITY OF VAASA
School Of Technology and Innovation

Author:	Shazid Hasan
Title of the thesis:	Integrating Machine Learning and Scenario Planning for Enhanced Demand Forecasting in Bangladesh's Garment Industry
Degree:	Master of Sciences in Technology and Innovation
Degree Programme:	Industrial Management
Supervisor:	Emmanuel Ndzibah
Year:	2026 **Pages:** 98

ABSTRACT:
The garment industry in Bangladesh is highly unstable and volatile in terms of exporting due to the cyclic nature of demand based on seasonal order, concentration of buyers, and disruption in the supply chain. The traditional forecasting models like the ARIMA and the Holt-Winters are very popular due to their interpretability and comparative ease of use though they cannot capture non-linear and non-regular order behaviour. Simultaneously, machine learning models have been brought up as a helpful instrument in enhancing forecast accuracy, and scenario planning has been identified as a means of aiding in decision making when there is uncertainty. This thesis studies how machine learning can be integrated with scenario planning to forecast the demand in the garment industry in Bangladesh. Based on this, the research question of the study is as follows: How can machine learning models, combined with scenario planning, improve the accuracy and adaptability of demand forecasting in Bangladesh's garment industry compared to traditional forecasting methods?

The theoretical basis of the thesis is based on forecasting theory, Systems Theory, the Dynamic Capabilities View, and literature on supply-chain resilience. The literature discusses that conventional statistical models can be applied to predict stable and seasonal demand trends but are less effective in situations where demand is influenced by sudden changes, uncertainty on the buyer side and structural volatility. Machine learning models, especially tree-based ensemble models, like Random Forest and XGBoost, are more flexible in the modelling of non-linear demand behaviour. The accuracy of forecasts is, however, not enough to make operational decisions in uncertain supply-chain conditions. Scenario planning is therefore a complementary decision-support layer, which transforms results of forecasts into scenarios of the stress test in the form of demand declines, demand peaks and buyer cancellation exposure.

Quantitative case-study approach is applied under the internal company-order data of a garment exporter in Bangladesh. The forecasting target is the monthly order quantity, and the traditional statistical models are compared with the Random Forest and XGBoost using a chronological holdout test period. The lagged order quantities, rolling demand momentum, and calendar-based indicators are used to form machine learning models. The best model's baseline forecast is then subjected to scenario planning via deterministic stress tests of a demand drop, a demand spike, and the cancellation of the largest buyer order exposure. The results suggest that machine learning can be used together with scenario planning to enhance accuracy of forecasting and flexibility in management. The paper concludes that a unified forecasting-scenario planning framework may aid in production planning, capacity calculation, buyer-risk evaluation, and information-intensive production planning in the garment sector of Bangladesh.

Keywords: Demand Forecasting, Machine Learning (ML), Scenario Planning, Garment Industry, Supply Chain Optimization

DISCLAIMER AND DECLARATION OF INDEPENDENT AUTHORSHIP

I hereby declare that this Master of Science (MSc) thesis has been written independently and represents my own original work carried out in accordance with the academic rules, ethical standards, and research integrity principles of University of Vaasa.

I confirm that, to the best of my knowledge and intention, I have not engaged in any form of academic misconduct, including but not limited to plagiarism, falsification, ghostwriting, unauthorized collaboration, or the misuse of Artificial Intelligence (AI), Large Language Models (LLMs), or other automated content-generation tools in a manner that violates university regulations or academic integrity standards.

Furthermore, I declare that this thesis has not been subcontracted, outsourced, purchased, or produced in whole or in part by any paid service provider, third party, commercial writing agency, or external individual. All analysis, interpretation, writing, and presentation contained in this thesis are the result of my own independent academic effort unless explicitly and properly referenced.

Where AI-assisted tools or digital technologies may have been used for limited support purposes permitted under university guidelines (such as language refinement, grammar checking, formatting assistance, or idea organization), such usage has been conducted responsibly, transparently, and without compromising the originality, intellectual ownership, or academic integrity of the work.

I fully acknowledge and accept responsibility for the contents of this thesis. I understand that if evidence of plagiarism, unauthorized AI misuse, academic dishonesty, or third-party authorship is discovered at any stage before or after submission, the University reserves the right to take appropriate disciplinary and academic actions in accordance with its regulations and policies. Such actions may include rejection of the thesis, annulment of the degree, disciplinary sanctions, or any other measures deemed necessary by the University.

By signing this declaration, I affirm my commitment to honesty, transparency, and ethical academic conduct.

Student Name: Shazid Hasan

Student Number: 2403806

Programme: Industrial Management

Title of Thesis: Integrating Machine Learning and Scenario Planning for Enhanced Demand Forecasting in Bangladesh’s Garment Industry

Content

Figures

Tables

1 Introduction

1.1 Background of the Study

Bangladesh's garment industry is increasingly known as a world-class manufacturing hub with an important role in economic development and exports. Bangladesh is now the second largest exporter of ready-made garments in the world, with more than four million workers, and around 80% of its workforce is women (Siddiqi, 2025). The garment industry is the largest export earner with about USD 47.39 billion in 2023 (Saif & Islam, 2025). In Bangladesh, export earnings play an enormous role in foreign exchange, employment, and industrialization, both in Bangladesh and abroad (Mondal et al., 2025).

Nevertheless, there is an increasing number of operational issues within the industry, especially in the forecasting of demand on orders, production schedules, and aligning of supply-chain responsiveness. The forecasting models that have been used in most garment-manufacturing enterprises have been developed based on the traditional statistical tools such as moving averages, exponential smoothing, Holt-Winters and autoregressive integrated moving average models, ARIMA. These approaches commonly rely on historic time series data, and anticipate predictable trends and seasonality (Douaioui et al., 2024). However, in a global clothing market, which is overwhelmed by the rapid shift in the consumer preferences, the rapid lead times and the frequent occurrence of external disturbances, such as buyer cancellations, logistic delays and raw material shortages, these traditional models do not adapt rapidly. Or, conversely, when suddenly demand has either risen or fallen, simple forecasts may become inaccurate or outdated, such as the sudden spike or drop of demand brought about by fashion trends or world events (Castañeda-Navarrete et al., 2021).

To address these complexities, more advanced approaches based on artificial intelligence (AI) and machine learning have become increasingly popular within the supply-chain space. Non-linear relationships, interaction between many explanatory variables (seasonality, holidays, lead times, cancellations), and combined endogenous and exogenous signals are all included in ML algorithms, especially tree-based models (Random Forest and XGBoost) (Douaioui et al., 2024). And on top of that, the practice of building “what-if” scenario predictions of future demand, which is known as scenario planning, can be used to complement predictive models with strategic preparedness and supply-chain resiliency. For example, scenario planning can be incorporated as a tool for firm stress testing and the ability to adapt capacity, inventory, and production plans accordingly.

More research on using ML forecasting models and scenario-planning in manufacturing supply chains, specifically in high-volume, export-oriented sectors such as Bangladesh’s garment sector, is in line with increasing research interest. But, when in practice many firms depend on traditional methods, or in some cases only partially use advanced techniques, this indicates a gap between research advancement and industry adoption.

1.2 Research Gap, Questions, and Objectives

Regardless of the ongoing development in machine learning and scenario-planning methods, there are still several gaps in the research of demand-forecasting.

First, most of the existing literature is focused on industries, or on supply-chain settings, most ML forecasting studies focus on large companies or on specific sectors within developed economies that presume good data infrastructure, large historical data, and ready access to technical talent (Ameen et al., 2025). Consideration of the Bangladesh ready-made garment industry is also very limited, as it is characterized by the export-driven nature of the sector, the changing multinational demands, and the low-cost production with labor intensive nature (Mondal et al., 2025). The empirical study of 43 Bangladeshi SMEs in the textiles and garments sector by Ameen et al. (2025) recorded the presence of severe barriers like data scarcity which affected 70% of firms, technical capacity constraints which occurred in 60% of them, and infrastructure constraints as manifested in power cuts in 46% of the SMEs and highlighted the scope to which the forecasting research disregards the unique reality of the developing countries. Shuvo et al. (2025) found poor demand forecasting to be a recurrent problem in the RMG industry in Bangladesh and observed that although artificial intelligence and machine learning have a lot of potential in improving operational efficiency, little empirical evidence on its implementation in the RMG sector has been found (Shuvo et al., 2025). Therefore, there is an evident gap in regarding developing and validating ML based demand forecasting models that would be more specific to operational constraints, data constraints and industry specifics of the ready-made garment industry in Bangladesh.

Second, although conventional time-series models like ARIMA and Holt-Winters have been widely investigated, they do not work well when there is a sudden change in demand or structural changes, such as buyer cancellations, global shocks, and logistic bottlenecks. This issue was further validated by Cooper et al. (2025), who showed that most state-of-the-art forecasting systems are easily broken when there is a shift in regime, with their experimental findings showing that the error of the black-box ML models grew by several times during regime shifts relative to normal times (Cooper et al., 2025). This result validates the assertion that predictive accuracy is not enough to support operational performance in case structural volatility is present. As a result, there is still a substantial research gap on disruption-sensitive forecasting structures capable of preserving predictive accuracy in response to events of abrupt demand fluctuation, structural turbulence, and supply chain turbulence, especially in the domain of the Bangladesh RMG sector where such disruptions are very common and extreme.

Third, most of the studies on machine learning models used in demand forecasting don't include the integration of scenario planning, which constrains their practical value in predicting extreme or non-historical events. In their review of 119 studies, Douaioui et al. (2024) found that despite the fact that supply chains are susceptible to unexpected shocks and sudden disruptions, the current literature on the use of models based on ML and deep learning is mainly focused on addressing the challenges of data quality, model interpretability, and system integration, but it suggests that the relevant literature does not address how these models are supposed to be tested or adjusted in response to the rapidly emerging disruption. Cooper et al. (2025) also mentioned that not all forecasting systems with disruption-free architecture can be well correlated with the inventory decision making process in the presence of uncertainty, and that solutions based on predictive modelling along with a systematic analysis of scenarios should be suggested.

Thus, integrating ML based demand forecasting and non-historical disruptive event driven scenario planning methodologies that can simulate extreme disruption events remains a gap in the literature that can be used to support proactive decision making amid volatile supply chain conditions.

Research question: How can machine-learning models, combined with scenario planning, improve the accuracy and adaptability of demand forecasting in Bangladesh's garment industry compared to traditional forecasting methods?

Research objectives:

1. To compare the forecasting accuracy of traditional statistical methods (e.g., ARIMA, Holt-Winters) with machine-learning models (e.g., Random Forest, XGBoost) in predicting garment-industry demand in Bangladesh.
2. To evaluate the role of scenario planning in anticipating demand fluctuations (e.g., demand spikes, drops, buyer cancellations, logistic delays) and enhancing strategic decision-making in the garment supply chain.
3. To analyze the combined impact of machine-learning forecasting and scenario-planning methods on production planning, inventory control, and overall supply-chain efficiency in the Bangladeshi garment sector.

By meeting these objectives, the study aims to bridge the theoretical and practical divide in forecasting approaches and offer actionable guidance for garment-manufacturing firms in Bangladesh.

1.3 Definitions and Scope of the Study

This section has the keywords that will create the theoretical foundation which is the basis of this study and shows the limits of each concept to be studied.

Demand Forecasting: It involves the systematic process of examining past data, trends, and either statistical or computational models to estimate future customer demand to make the production, inventory, and supply chain decisions. Thippur Manjunath (2025) compared various methods for demand forecasting discovered Random Forest and XGBoost are particularly useful in case we are using the historical demand data scouring at the product level. The scope of this research is, the demand forecasting is limited to the company level, where the order quantity is aggregated to a monthly level with the Latest Ex-Fty Date as temporal index. Product-category information is employed to describe the segmentation and the forecasting comparison is made on the entire order-quantity series in the months. The forecasting methods used are two classical time series methods, the autoregressive integrated moving average (ARIMA) and the Holt-Winters methods, and two machine learning methods, Random Forest and XGBoost.

Machine Learning: ML is a subset of artificial intelligence where the algorithms are based on the concept of learning from training data to figure out the complex non-linear patterns of input variables and their corresponding output data and thus providing better prediction accuracy than traditional statistical approaches. Ma et al. (2025) found volatile SKUs Forecasting performance with exogenous factors such as holidays and promotions with rich error analysis validate the model performances of XGBoost (Ma et al., 2025). Gaertner et al. (2024) again tested decision trees, Random Forest and XGBoost for products level prediction with certain practical considerations including short history and high variance in demand patterns. This study is limited to tree-based ensemble models, i.e., Random Forest and XGBoost (Gaertner et al., 2024). For this study, we only consider two machine learning models (Random Forest and XGBoost) with lagged order quantities, and with indicator variables for month, quarter, and a rolling three-month average. The models are tested on a chronological holdout test period, and typical forecast-error measures are employed.

Scenario Planning: It is a strategic approach that includes the creation of alternative future scenarios and assessing the potential impact of those on operations and decision-making that enables contingency planning and resilient responses. Arora et al. (2025)

proposed disruption awareness forecasting model which preserves resilience in diversity of disruption phases, which covers direct the sudden demand pattern change in crisis (Arora et al., 2025). von der Gracht and Darkow (2010) demonstrated systematic estimation of the impacts of logistics delays and structural changes using Delphi based scenario planning as a methodological base for anticipating future disruptions (von der Gracht & Darkow, 2010). This study uses scenario planning as a deterministic stress testing overlay on the XGBoost baseline forecast. The empirical scenarios center on the exposure of cancellation of the largest buyer's order, a drop in demand and a spike in demand, with a logistics delay as a broader contextual risk in the garment industry's supply chains. The scope is limited to disruption scenarios at the demand side and is not expanded to include geopolitical modelling or macroeconomic simulation.

Garment Industry: It can be defined as the sector of manufacturing which is concerned with the production of clothing and textiles for the domestic and international markets. This study is limited to the ready-made garment sector of Bangladesh, which is the largest export sector of the country providing job to millions of workers and serves on global buyers. The scope is further reduced to Tier-1 factories working business-to-business export orders that focus on the knitwear and woven product categories such as tops and bottoms. Tier-2 and Tier-3 subcontracting units, domestic retail operations and non-apparel textile manufacturing are not included in this study.

Supply Chain Optimization: It is the process of optimizing supply chain for efficiency, cost-effectiveness, and responsiveness by better forecasting and planning. Zahraei and Teo (2017) modelled how production smoothing and safety stocks can be optimised together to combat demand variability in supply networks, thus providing a theoretical foundation for making operational adjustments based on forecasts. In this study, optimisation of supply chain is studied under the perspective of application of forecast outputs in three distinct areas of operation: safety stock tuning, smoothing of production and basic capacity alignment (Zahraei & Teo, 2017). Broader supply chain functions like procurement strategy, supplier relationship management, logistics network designing and transportation optimisation, are not discussed. The study uses company-level data of their demand over a period from January 2023 to January 2026. Analysis is restricted to internal company data and secondary information with additional broader contextual knowledge of the Bangladesh garment industry. The results are not meant to be generalisable in any way beyond the Bangladeshi RMG context or for the types of factories considered.

1.4 Structure of the Study

Chapter 1 discusses the background of the study, including the Operational theoretical facility and economic issues of the garment industry in Bangladesh. The research objectives and questions are developed. Finally, definitions of key terms, as well as limitations of the study are presented.

Chapter 2 will present an extensive literature review about demand forecasting methods, covering both traditional statistical methods and machine learning techniques. This will also involve considering scenario planning as a strategic tool, and the theoretical and conceptual foundations which provide a basis for the research.

Chapter 3 explains the research methodology that is used in this study in which the data collection process, data preparation steps, and development of both traditional and machine learning forecasting models are explained. The scenario-planning approach as well as model validation techniques are discussed, as well.

Chapter 4 will present the data analysis results and findings. This includes descriptive statistics of the data set, forecasting results from the different models, scenario planning results and comparing forecasting accuracy and adaptability. Key findings interpretation is also provided.

Chapter 5 will draw conclusions to the study by discussing the results in relation to theory and practice. Limitations of the research are acknowledged, and recommendations are suggested for future research as well as for practitioners in the garment industry in Bangladesh.

2 Literature Review

This chapter is about the analysis of the literature on demand forecasting in the garment and textile industry has been conducted critically. It discusses how traditional statistical tools have developed into modern machine learning tools, how scenario planning plays a strategic role in overcoming global supply chain uncertainty, and the theoretical background that is the basis of this study, Systems Theory and Dynamic Capabilities. Finally, these ideas are put in context with the structural limitations of the Bangladesh garment industry.

2.1 Demand Forecasting in the Garment Industry

A major aspect of supply chain planning in the global garment and textile sector is demand forecasting where firms are engaged in an environment dominated by high volume, short product life cycle and volatile consumer tastes. The industry can be described as dynamic due to the presence of such features as a fast fashion cycle, variable orders, and ongoing globalisation of manufacturing makes it inevitable and difficult to forecast by nature (Bruzzone et al., 2022). The researchers have indicated that the accuracy of the forecasts in the apparel supply chains is more considerable than in most other manufacturing industries because the stockouts, discounted losses, and long lead times can make an immediate impact on the financial performance and competitive advantage (Wren, 2022). This section is a critical appraisal of the forecasting techniques, the flaws of the time-series models applied traditionally, the volatility and external interference, and puts the challenges in perspective in the context of the Bangladesh garment industry. The summary of key industry features that complicate garment demand forecasting, such as short product lifecycles, volatility, buyer dependency, and macro-level disruptions are summarised in Table 1.

Table 1 Apparel Industry Characteristics and Forecasting Implications.

Industry Characteristic	**Bangladesh/Global Context Impact**	**Forecasting Implication**
Short Product Lifecycles	Fast fashion collections last only weeks.	Insufficient historical data for traditional time-series models (underfitting).
High Volatility	Sudden shifts in global consumer tastes.	Traditional linear models (ARIMA) fail to capture non-linear demand spikes.
Buyer Dependency	Bangladesh factories must meet tight global brand lead times.	Forecast errors lead to immediate stockouts or costly emergency airfreight.
Macro Disruptions	Geopolitical instability and logistical delays.	Need for scenario planning and stress testing beyond pure statistical outputs.

2.1.1 Overview of Forecasting Methods in Textile and Apparel Supply Chains

The classical statistical methods used to make forecasts in the textile and apparel industry, including Moving Average, Exponential Smoothing, Holt-Winters, and ARIMA, assume that past trends can adequately reflect demand trends in the future (Agathangelou et al., 2020). The models are highly utilized because of their interpretability and minimal data needs. As an illustration, the ARIMA models are known to be effective in capturing autocorrelations, and the Holt-Winters model is known to be effective in modelling the seasonality of apparel sales (Alam et al., 2016). But modern research also tends to suggest that these models are less effective in the areas that are defined by a considerable level of product proliferation, as well as the variability in promotional approaches.

With the frenzied pace of fashion cycles, investigations started incorporating causal designs, including regression with macroeconomic variables, yet these as well failed to address nonlinearities introduced by the new era apparel market (Bruzzone et al., 2022). More recently, machine learning (ML) algorithms, including Random Forest, Gradient Boosting and LSTM neural networks have gained popularity since they can model more intricate interactions and non-linear dynamics (Wren, 2022). However, ML techniques usually demand hefty data sets, computer infrastructure, and cautious feature engineering, which smaller companies might not have. Cross-study analyses suggest that hybrid methods that combine classical and ML methods are more productive than individual methods, especially in the short-life-cycle clothing (Mansur et al., 2025). It is this nature that is developed that brings into focus the need of data-intensive forecasting systems in the garment industry.

Figure 1. The evolution of demand forecasting methodologies. Author's illustration based on Agathangelou et al. (2020), Goel et al. (2024), Makridakis et al. (2022b), and Mansur et al. (2025).

To visualize such a methodological change, Figure 1 illustrates how the traditional linear foundations were developed into the hybrid and AI-integrated architecture that is used in contemporary research.

2.1.2 Challenges in Traditional Forecasting (ARIMA, Holt–Winters)

To start with, models like ARIMA presuppose the linearity, stationarity, and constant pattern of demand -all factors that are uncommon to the clothes market (Agathangelou et al., 2020). It has been found that ARIMA assumptions may be invalidated by sudden fashion trends, promotional campaigns, or geopolitical disruptions, or may result in large forecast errors (Wren, 2022). Holt-Winters is more sensitive to seasonality, although several studies indicate that its performance decreases as seasonality itself becomes volatile due to changing consumer sentiment or digital retailing cycles (Bruzzone et al., 2022). The second challenge is associated with the low lifecycle of apparel products. The demand horizons of many clothing items can be only a matter of weeks, and hence, there is little historical information available to tune the models. ARIMA and HoltWinters have long data series requirements; thus, when applied to new or fast-fashion SKUs, they tend to underfit (Mansur et al., 2025).

Third, the traditional models cannot include exogenous variables, including economic indicators, weather changes, international trade policies, and retailer promotions. Even though there are ARIMAX variants, empirical studies indicate that they continue to be subpar compared with machine learning models in terms of accountability to multicollinearity and nonlinear interactions (Wren, 2022).

Lastly, a comparative analysis of prediction accuracy reveals that forecast errors in the clothing industry by applying classical approaches may exceed a high percent at turbulent times (Agathangelou et al., 2020). This has encouraged companies to shift to ML-based and hybrid forecasting models. Nevertheless, critics cautiously note that ML techniques provide overfitting risks and demand high-quality data, which most developing-country manufacturers, such as those in Bangladesh, might not be able to ensure (Alam et al., 2016). Therefore, it is not just a technological problem, but also an organisational and infrastructural one. Table 2 contrasts ARIMA and the Holt-Winters as conventional forecasting models and outlines their applicability and shortcomings in the setting of the Bangladesh garment-industry.

Table 2 Critical Comparison of Traditional Forecasting Methodologies (adapted from Atesongun & Gulsen, 2024; Douaioui et al., 2024; Swaminathan & Venkitasubramony, 2024).

Model	**Parameters**	**Bangladesh Industry Relevance**	**Critical Limitations (Reviewer Requirement)**
ARIMA	Autocorrelation, Differencing, Moving Average.	Used for stable, bulk-order commodities where patterns repeat.	**Stationarity Constraint:** Fails to account for non-linear demand shocks and volatility clusters typical of fast-fashion cycles.
Holt-Winters	Level, Trend, and Seasonality (Smoothing).	Effective for basic seasonal wear with predictable peaks.	**Static Nature:** Cannot incorporate exogenous variables (e.g., social media trends or logistical disruptions); purely backward-looking.

2.1.3 Demand Volatility, Seasonality, and External Disruptions

The garment industry is characterized by demand volatility. The presence of fashion trends, cultural events, varying weather conditions, and social media contribute to the creation of sharp changes in sales, making forecasting a risky practice (Bruzzone et al., 2022). Seasonality also increases this complexity; apparel demand is multi-seasonal (e.g., Eid, Ramadan, Christmas, Winter collections), in that it varies across regions. Though classical models try to explain seasonality with an additive or multiplicative factor, the patterns in the real-world change year-to-year, making seasonal patterns volatile and hard to model (Mansur et al., 2025).

The COVID-19 pandemic, supply chain logistical constraints, transportation delays, costs of energy, and geopolitical conflicts have caused previously unknown forecasting challenges. Research indicates that forecast accuracy in the pandemic period failed in the entire world as lockdowns overnight modified purchasing behaviours (Wren, 2022). Table 3 singles out some demand exogenous drivers that can disrupt apparel forecasting such as seasonality, trade-policy, supply, and macroeconomic conditions.

Table 3. Key external factors impacting apparel demand (adapted from Milewska, 2022; Swaminathan & Venkitasubramony, 2024.)

Factor	Example Impact	Supporting Source
Seasonality/Fashion cycles	Spike in Q3–Q4	Industry studies
Trade policies	Reduced buyer orders	WTO/WB reports
Supply disruptions	Delays occur lost orders	COVID-19 research
Economic conditions	Consumer spending goes down	Macro research

The recent literature highlights the importance of using scenario planning to supplement forecasting to enable managers to simulate the best-case, worst-case, and moderate-case demand paths depending on the assumed disruption (Agathangelou et al., 2020). Scenario planning, as opposed to the use of pure statistical models, brings forward qualitative judgement, strategic foresight, and multi-factor stress testing. The hybrid is becoming more popular as a necessity in industries such as apparel, in which uncertainty is not an exceptional phenomenon but a systemic feature (Bruzzone et al., 2022).

2.1.4 Comparative Perspectives & Bangladesh Context

Bangladesh, being the second-largest garment exporter in the world, has the same forecasting issues witnessed in the global market, but with further structural constraints. The manufacturer supply chain is also highly buyer-oriented and, hence, the precision of predictions is not only critical to internal planning but also to meet high lead-time requirements of global brands (Alam et al., 2016). The use of advanced ML would not be easy since the factories in Bangladesh have little data, digital infrastructure, and analytics potential, compared to developed economies.

It has been shown that the cross-country analysis shows that the use of ML-driven forecasting has been more rapid in the companies in China and Turkey where inventory turnover and stockout reduction are high (Mansur et al., 2025). Although historically averages and simple time-series analysis have been closer to the truth in volatile environments, these methods and previous information still play a more important role in Bangladeshi companies (Wren, 2022). The gap identified indicates that it is necessary to incorporate the latest forecasting and scenario planning to develop resilience. Table 4 compares the data infrastructure, adoption of technology, the role of supply-chain and the analytical capacity of the garment sector in Bangladesh with other more digitally advanced markets of garment production.

Table 4. Comparative analysis of forecasting capabilities in Bangladesh and developed garment-export markets. Author's synthesis based on Alam et al. (2016), Wren (2022), and Mansur et al. (2025).

Feature	Bangladesh Garment Sector	Developed Markets (e.g., China, Turkey)	Research Gap / Contextual Need
Data Infrastructure	Predominantly manual or fragmented ERP systems; historical data often lacks granularity.	High-integrated IoT and real-time cloud-based POS data sharing.	Need for models that perform well with "thin" or "noisy" data sets.
Technology Adoption	Heavy reliance on simple moving averages and expert intuition (Wren, 2022).	Rapid integration of ML-driven inventory turnover and automated replenishment (Mansur et al., 2025).	Significant opportunity for ML-hybrid model testing in local contexts.
Supply Chain Role	Primary focus on cut-make-trim (CMT); limited visibility into end-consumer behavior.	End-to-end integration; high visibility from raw material to retail shelf.	Need for "Scenario Planning" to buffer against upstream buyer-side shocks.
Analytical Expertise	Limited availability of in-house data scientists within manufacturing units.	High availability of specialized analytics teams and AI-enabled software.	Need for user-friendly, robust tools that don't require PhD-level maintenance.

2.2 Machine Learning Applications in Demand Forecasting

Machine learning (ML) has naturally revolutionized the demand forecasting in all industries, as well as the garment industry where the classical statistical models may not be as complex as the nonlinear demand patterns, short product life cycle and market dynamics. The shift to ML methods is a more universal global trend of resorting to data-driven supply chain management, particularly in industries with a high level of uncertainty (Makridakis et al, 2022a). The apparel industry is also extremely susceptible to the changing nature of the consumer population, marketing, and seasonal fluctuations, and, therefore, it is among the industries that can be successfully implemented into the ML-based prediction. This section will critically examine the development of ML techniques, how tree-based ensemble models work, the developments in hybrid modelling, and accuracy metrics to approximate predictive power.

2.2.1 Evolution of AI/ML Techniques in Forecasting

The initial forecasting models in the supply chains mainly depended on statistical tools like ARIMA, SARIMA, and the exponential smoothing and regression models. Successful with regular patterns, the approaches do not support the non-linear, high-dimensional, and discontinuous forms of modern apparel demand (Makridakis et al, 2022a). This gap has been filled with the early development of machine learning in the early 2000s, which provides methods that will be able to model intricate relationships without them being explicitly specified.

Among the early ML algorithms in prediction, support vector regression (SVR), artificial neural networks (ANNs), and k-nearest neighbours (kNN) were used. Empirical research indicates that ANNs are more effective than ARIMA in cases where there is nonlinearity in the data or when the data are affected by various other variables (Atesongun & Gulsen, 2024). Nonetheless, initial neural networks were large datasets and highly tuned and would not be readily adopted by smaller and medium-sized apparel companies.

The development of deep learning and long Short-Term Memory (LSTM) networks represented an important step towards being able to capture long temporal relationships (Kontopoulou et al., 2023). LSTMs are also particularly useful to predict the demand of apparel under the influence of consecutive promotions, repeating patterns of seasonality, and sudden shifts in trends. However, as several scientists state, deep learning does not always perform better; in the case of small or noisy data, generic models such as the Random Forest often coincide with LSTM as sources of lower overfitting (Makridakis et al, 2022a). So, the history of ML as a forecaster has seen the transition to data-driven methods of forecasting replacing the model-driven methods of forecasting, although its uptake severely depends on the availability of data and organisational capacity. Figure 2 is a summary of the progression of early statistical and machine-learning methods, to the more advanced ensemble and deep-learning methods applied to forecasting research.

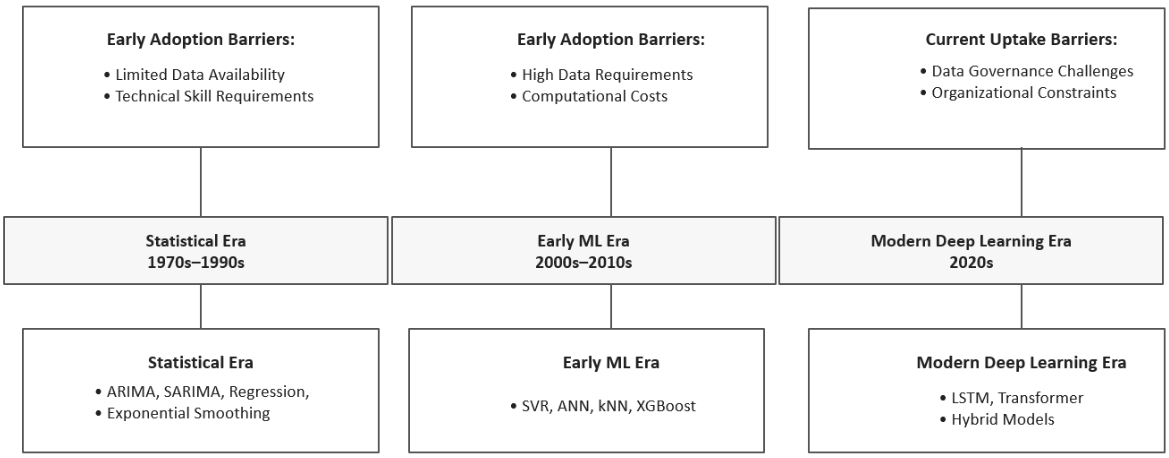

Figure 2. Evolution of AI/ML techniques in forecasting. Author's illustration based on Makridakis et al. (2020), Chen and Guestrin (2016), Kontopoulou et al. (2023), Douaioui et al. (2024), and Ameen et al. (2025).

2.2.2 Tree-Based & Ensemble Models (Random Forest, XGBoost)

Random Forest (RF) and Extreme Gradient Boosting (XGBoost) are tree-based ensemble models that have become prevalent in demand forecasting because of their strengths and have been shown to work well in capturing nonlinear interactions (Chen & Guestrin, 2016). Random Forest uses bagging, the method of combining several decision trees haphazardly to minimize variance and enhance generalisation. XGBoost employs the concept of boosting, which is a sequence of corrections to past mistakes, allowing it to perform better predictive tasks in complicated conditions.

Comparative analysis proves that random forests can often be a more efficient predictor of apparel than classical models such as ARIMA and Holt-Winters because it can use categorical variables, promotion factors, and weather data (Hobor et al., 2025). RF does not presuppose a stationary situation, and it is formulated to suit apparel that has a volatile demand. Nevertheless, other researchers posit that Random Forest may be inefficient with data that is extremely seasonal without any lags and rolling windows that were constructed manually (Atesongun & Gulsen, 2024). This drawback indicates that the performance of RF is very sensitive to the quality of feature engineering.

XGBoost has received much attention in terms of accuracy and computation efficiency. It is also one of the most successful in forecast competitions, and the M-series standards (Makridakis et al, 2022a). In the context of retail forecasting, XGBoost remains more successful compared to LSTM or Random Forest in the prediction of fashion, given that the datasets comprise several exogenous variables like marketing expenditure, discount rates, and macroeconomic factors (Hobor et al., 2025). It has regularisation mechanisms that help to keep overfitting at bay, a vital attribute when handling short life cycle apparel items.

Among all the advantages, critics note that ensemble models require extensive data pre-processing. Any gaps in values, inaccurate timestamps, and broken SKU life cycles can negatively impact performance unless treated with great care. Other researchers also suggest that, although XGBoost provides high accuracy, its inability to be understood as a simple tree model can be a limiting factor to the adoption of the algorithm among supply chain managers (Bruzzone et al., 2022). These contradictory results explain that, despite the high power of the ensemble models, their applicability relies on the quality of the data, the representation of features, as well as organisational willingness to use more sophisticated analytics. Table 5 compares three large families of machine-learning models regarding their accuracy, interpretability, data needs, and feature engineering requirements.

Table 5. Comparative analysis of Random Forest, XGBoost, and LSTM for demand forecasting. Author's synthesis based on Chen and Guestrin (2016), Makridakis et al. (2022a), Kontopoulou et al. (2023), and Douaioui et al. (2024).

Criteria	**Random Forest (RF)**	**XGBoost (Ensemble)**	**LSTM (Deep Learning)**
Prediction Accuracy	High; robust against outliers and noisy data.	Very High; often superior in forecasting competitions.	Exceptional for long-term sequences and complex patterns.
Interpretability	**Moderate:** "Feature Importance" shows which factors matter most.	**Low/Complex:** Harder for managers to trace the decision logic.	**Black Box:** Highly complex; requires specialized visualization tools.
Data Requirements	Flexible; performs well with smaller or "thin" datasets.	Requires structured data and careful hyperparameter tuning.	Needs massive, high-frequency datasets to avoid overfitting.
Feature Engineering	Low; can handle raw categorical data effectively.	Moderate; requires manual creation of lags and rolling windows.	High; requires intensive data normalization and sequence preparation.

2.2.3 Hybrid & Integrative Forecasting Models

The goal of hybrid forecasting models is to leverage the advantages of other modelling methods, such as statistical, machine learning, and deep learning models, to enhance the accuracy of the predictions, especially in highly uncertain industries. The literature continues to demonstrate that all prediction models cannot be considered universal and optimal; hybrid models are frequently better than monolithic methods in usage due to their ability to reflect both linear and nonlinear forms (Atesongun & Gulsen, 2024). This is especially applicable in the case of garment demand forecasting, where trends can be linear growth, whereas promotion or season spikes are nonlinear.

A hybrid model is a typical combination of ARIMA (linear components) and neural networks (nonlinear dynamics). Research suggests that in cases where demand dynamics are both stable in terms of their growth at the base and unstable with regular short-term variations, ARIMA-ANN hybrids will outperform either ARIMA or ANN models Zhang (2003). Shorter SKU life cycles are also in this category because the ARIMA would consider the seasonality in the baseline, whereas the ML would be capable of managing unexpected changes.

A different body of work studies ensemble-hybrid structures, which use weighted averaging or stacking of predictions of Random Forest, XGBoost, and LSTM. Based on empirical studies, there appears to be a reduction in the error rates by up to 15 to 20 per cent in stacked hybrid models than in individual ML models (Hobor et al., 2025). This improvement can be explained by the fact that hybrid ensembles can capture various characteristics of RF models' interactions, XGBoost captures boosted patterns, and LSTM models temporal dependencies.

Nevertheless, hybrid models have their drawbacks. Researchers warn that model combinations complicate the process, make them more expensive to calculate, and are prone to overfitting, particularly when they are being trained on small quantities of data, a situation that is prevalent in the garment industry of Bangladesh (Alam et al., 2016). Data scientists, sophisticated software tools, and powerful data pipelines are also needed in hybrid systems and are not always attainable by small and medium-sized manufacturers. Hence, as much as hybrid models have good accuracy advantages, they should be applied with resource limitation in mind. Figure 3 illustrates the possibility of combining different forecasting structures (statistical, machine-learning, and hybrid) to extract both linear and non-linear demand patterns.

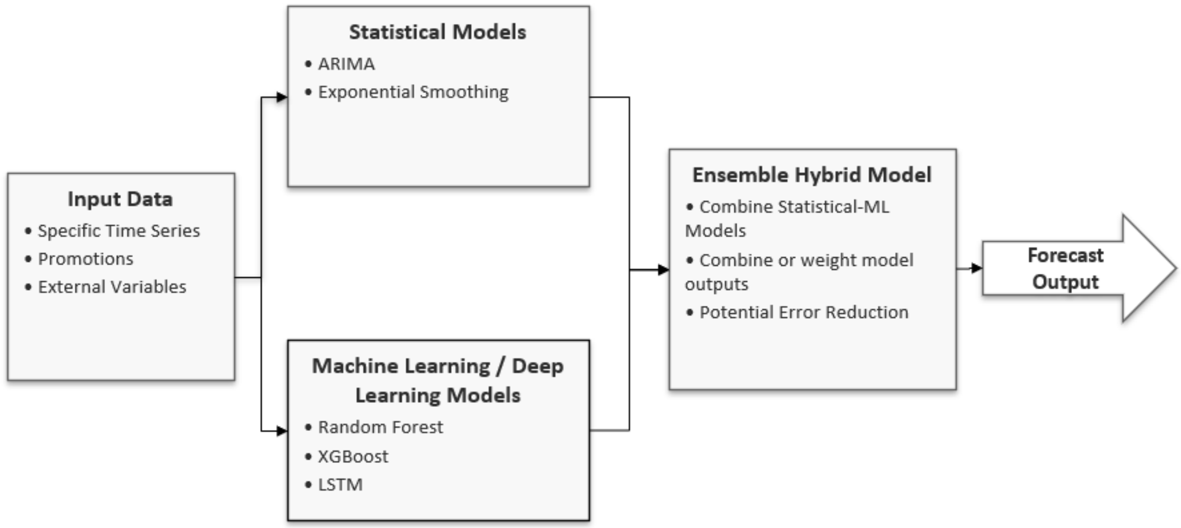

Figure 3. Framework of hybrid and integrative forecasting models. Author's illustration based on Zhang (2003), Atesongun and Gulsen (2024), and Hobor et al. (2025).

2.2.4 Performance Metrics & Model Evaluation

Model assessment requires performance metrics that reflect the unique characteristics of apparel demand. Mean Absolute Error is defined as MAE, Root mean squared error (RMSE), Mean Absolute Percentage Error (MAPE), Symmetric MAPE (sMAPE). All the metrics represent various facets of the accuracy of the forecast. MAE measures average error magnitude, RMSE penalises large errors, and MAPE presents error as a percentage, useful for comparison across SKUs (Makridakis et al., 2022b).

But many argue that traditional measures, like the measurement of apparel performance, can be misleading in forecasting. For example, MAPE becomes unstable when demand reaches zero—a common phenomenon in fashion items in off-season periods (Atesongun & Gulsen, 2024). In both cases, RMSE penalizes outliers, which may be the result of external disruptions, not model failure. To address this, researchers recommend scale-independent metrics such as Mean Absolute Scaled Error (MASE), which will enable fair comparison between models and product categories (Makridakis et al., 2022b).

Table 6. Performance metrics applied in forecasting literature (adapted from Goel et al., 2024; Makridakis et al., 2022b; Mansur et al., 2025).

Metric	**Meaning**	**Pros**	**Cons**
MAE	Avg absolute error	Easy to interpret	Ignores error scale
RMSE	Penalizes large errors	Good sensitivity	Harder to explain to managers
MAPE	% error measure	Useful for business	Problematic when volumes low
R^2	Model explanatory power	Widely used	Can be misleading

Table 6. summarizes commonly used forecast evaluation metrics and their general strengths and limitations based on forecasting literature. These variations underscore the importance of multi-metric evaluation and assessment against naive models that will often be overlooked in industry practice.

Table 7. Strategic Metric Selection Matrix based on Demand Characteristics. Author’s synthesis based on Makridakis et al. (2022b) and Syntetos et al. (2015).

Demand Profile	Typical Garment Category	Recommended Primary Metric	Analytical Justification
High-Volume / Seasonal	Basic Essentials (T-shirts, Underwear)	MAPE / WMAPE	Provides percentage-based error relative to volume; WMAPE ensures high-volume errors are prioritized.
Intermittent / Lumpy	Niche Fashion, High-End Outerwear	MASE	Overcomes the "division by zero" issue; compares the model against a "Naive" forecast to prove added value.
High-Volatility / Trend-Driven	Fast-Fashion / Viral Items	RMSE	Penalizing large misses heavily; essential for managing risk when overstocking leads to high costs.
Cross-Category Portfolio	Mixed Retail Assortments	sMAPE	Symmetric properties prevent the metric from over-penalizing high forecasts more than low ones.

2.3 Scenario Planning and Strategic Forecasting Integration

2.3.1 Concept and Approaches to Scenario Planning

Scenario planning refers to a strategic foresight framework that is meant to guide organisations through uncertainty by building various valid futures rather than basing them on one deterministic view. Schoemaker (1995) asserts that scenario planning was conceived due to constraints of the classic models of forecasting that often assume linearity and steady environments that are not always the case within the global supply chain. Derbyshire and Wright (2017) also support this perspective in that scenario planning motivates decision-makers to investigate structural uncertainties that include geopolitical risk, input-price variations, and technological interruptions.

Strategies of scenario planning generally exist in two general categories: normative scenarios (describing desired futures) and exploratory ones (painting a range of plausible outcomes) (Bishop et al., 2007). The tradition of exploration (led by Royal Dutch Shell) is still dominant in the context of supply chains because the influence of external forces is rather high and unmanageable. Nevertheless, critics state that traditional methods of scenario are too dependent on expert opinion and thus on subjectivity and bias (Amer et al., 2013).

The recent advances implement data-founded approaches to scenario construction, as Ramirez et al. (2015) observe that hybrid systems incorporating both qualitative descriptions and quantitative modelling can be more credible and usable. This trend indicates a greater academic shift towards foresight methods that complement managerial intuition with evidence-based analytics. However, on the one hand, scenario planning broadens the strategic imagination, and on the other, inconsiderate scenarios can block decision-makers or foster a misleading sense of security in improbable scenarios Csaszar et al. (2024). These criticisms highlight exigencies of systematic, empirically grounded scenario development, especially those of the volatile sector such as textiles and apparel.

2.3.2 Scenario Planning in Supply Chain and Demand Management

The emerging global uncertainty has made scenario planning more appropriate in the field of supply chain management. Christopher and Peck (2004) detail that new supply chains are vulnerable to systemic weaknesses such as supplier concentration risks and unexpected demand shocks, which are not constrained by conventional methods of deterministic forecasting. Scenario-based planning, in turn, allows companies to stress-test their operations when the macroeconomic situation, climatic conditions, and geopolitical situation are different.

Scenario planning used in demand management aids organisations to prepare against the various demand paths. According to Syntetos et al. (2015), deterministic forecasts tend to be ineffective in periods of extreme volatility, hence making scenario-based tests of elasticity essential in industries with seasonal and trend changes that prevail in the apparel industry. Other researchers, though, refute it by the fact that scenario planning can be too time-consuming or resource-consuming to be used as a fast process in sectors that have to be moving very fast (De Angelis et al., 2018). Recent studies, however, indicate that scenario planning promotes resilience. As an example, Ivanov (2021) shows that scenario-based modelling improves the supply chain recovery strategies on outbreaks, including COVID-19, and discloses how the alternative demand/supply circumstances can transform capacity planning.

In the garment industry it is observed that due to the globalisation of production networks, scenario-based risk assessment becomes invaluable (Christopher & Peck , 2004). The example of the apparel supply chain in Bangladesh, where long lead times and sensitivity to changes in global demand are typical indicators, is apt for this argument. Although statistical models continue to be employed in most forecasting studies of the textile industry, scenario planning provides a complement to them in terms of providing the ability to incorporate low-probability, high-impact events. Supply chain resilience scholars are increasingly promoting this integration, which indicates they are becoming more relevant strategically (Ivanov, 2022).

2.3.3 Integrating ML Forecasts with Scenario-Based Stress Testing

Combining machine learning (ML) forecasting with scenario-based analysis is an important methodological innovation. ML models operate well in non-linear change, multi-dimensional interactions, and dynamism in behavioural shifts of demand (Douaioui et al., 2024). Nevertheless, several researchers note that ML forecasts remain inherently backward-looking in nature, as they are trained on historical events (Bengio et al., 2021). This missing information explains why scenario planning is used to explore structural uncertainties that are absent from the training data.

Some of the scholars have mentioned the existence of conceptual synergy between the two strategies. As an example, Badakhshan et al. (2024) demonstrate that a combination of ML predictions and scenario-based simulations can be utilized to make a more solid inventory and capacity decision-making instead of using either of the techniques to complete this task. Their results are consistent with the reports by Hosseini et al. (2019), who state that ML offers accuracy granularity, whereas scenario planning is framed to provide strategic breadth. This bilaterality is particularly critical in those businesses that have long production cycles like apparel, whereby the error in forecasting is passed on through the supply chain.

When the tests on ML forecasts in alteration scenarios are done, it helps companies know how the models respond to extreme conditions. Klibi and Martel (2012) uphold that even very precise models can do badly with exogenous shocks that alter the underlying data-generating process. Therefore, the robustness evaluation based on scenarios is critical in terms of assessing model stability. Similar studies by Ivanov and Dolgui (2022) validate that combining ML predictions and disruption situations improves the insight into possible bottlenecks and escalation of costs.

Although, critics warn of the possibility of overwhelming firms with model complexity because of ML scenario integration (Makridakis et al., 2022). In a bid to cope with this objection, Ivanov and Dolgui (2022) support simplified hybrid frameworks in which ML offers baseline predictions, and scenario planning analysis assesses the deviations of such. This method is both analytical and practical. Such integrated forecasting architectures will be of great benefit to the apparel industry in Bangladesh, where customers have become more demanding in terms of rapid responsiveness.

2.3.4 Global Case Study Insights

International case studies showcase the resilience of the supply chain based on scenario-integrated forecasting. Royal Dutch Shell is the most notable one; its scenario planning practices have enabled the company to long-term predict geopolitical and economic turbulence (Wack, 1985). Although Shell is not a manufacturing company, its practices give a picture of how structured narratives can be used to supplement quantitative models in highly uncertain settings.

Nike has received broad mention in the apparel segments as having orchestrated data-driven forecasting with risk assessments through scenarios, especially in the face of the COVID-19 pandemic. In the fast fashion supply chain of Zara, store closures in case of a pandemic and variable demand have been simulated with the help of scenario analysis (Milewska, 2022). In the case of the Bangladesh garment industry, such models may help it a great deal to lessen its susceptibility to international shocks like changes in demand, logistical upsets, and macroeconomic changes. Table 8 presents some examples of international experience of scenario-based integrated planning and show how they can be applied to address demand uncertainty in the export-oriented garment sector in Bangladesh.

Table 8. Comparative Case Insights: Strategic Approach and Relevance to Bangladesh. Author's synthesis based on Wack (1985) and Milewska (2022)

Organization	**Strategic Approach**	**Key Outcome**	**Relevance to Bangladesh Context**
Royal Dutch Shell	Narrative Scenario Planning (Long-term)	Transitioned from "predicting" to "preparing" for oil price volatility.	Establishes a framework for navigating macroeconomic and geopolitical trade policy shifts.
Nike	Data-Driven Risk Assessment (Mid-term)	Orchestrated inventory agility during pandemic-related retail disruptions.	Demonstrates how manufacturers can buffer against "Buyer-side" order cancellations.
Zara (Inditex)	Simulation of Store/Logistics Closures (Short-term)	Optimized stock distribution and reduced liquidation losses via simulation.	Provides a model for managing "Just-in-Time" production during supply chain interference.

2.4 Theoretical and Conceptual Framework

In this section, the theoretical backgrounds underpinning the research are made, which moves the discussion from how forecasting models work out to why they are applied strategically. It starts by assessing Forecasting Theory and the movement towards Data-

Driven Decision Making (DDDM). It then examines the Systems Theory perspective to get a feel of the interdependence of the supply chain and then the Dynamic Capabilities View (DCV) perspective, which explains why analytical tools such as ML and scenario planning can be a source of competitive advantage. Lastly, the theories are integrated into a Conceptual Framework that depicts the association between technological inputs and organizational resilience in the Bangladesh garment sector.

2.4.1 Forecasting Theory and Data-Driven Decision Making

The theory of forecasting gives the theory of demand prediction its analytical basis as it focuses on the analysis of past occurrences, statistical tendencies, and model-based predictions of future trends. Classical forecasting theory, which is based on the writings of Armstrong (2001), is that to create an accurate forecast, systematic processes, parsimony, and validations over time horizons are necessary. But recent researchers question whether the classical statistical methods are appropriate in volatile and globalised sectors like apparel. But, when faced with non-linear dynamics, the potential for sudden disruptive events or structural discontinuity even in theory our traditional models seem to fail (Arend & Bromiley, 2009). This limitation substantiates the case of data based adaptive forecasting paradigms.

The data-driven decision-making theory suggests that a higher prediction quality can be achieved through the deployment of high-frequency data, machine learning and real-time analytics in operational planning (Brynjolfsson & McElheran, 2016). However, critics say that even advanced algorithms can be affected by bias because of the quality and representativeness of training data Bengio et al. (2021). In very turbulent contexts, past patterns might not be sufficient to describe the risk of rare but high impacts.

This way, literature is likely to lean towards an intermediary stance of forecasting between rigour of statistics and flexibility of data driven. According to Veiga et al. (2016), the current forecasting is not a one-method discipline but a complex system of decision-support using analytics, managerial judgement, and experimenting with the scenarios. This transformation is quite relevant in the situation of the garment industry of Bangladesh where business uncertainty, seasonality, and geopolitics necessitate those changeable and empirically confirmed forecasting models be implemented.

2.4.2 Systems Theory and Dynamic Capabilities View

Systems Theory assumes that organisations and supply chains are systems that are interdependent and cannot be analysed based on the behaviours of the components independently. The general argument of Kalaitzi and Tsolakis (2022) is that systems must be studied in their entirety, and this has been extensively used in the supply chain literature. Related to forecasting and scenario planning, this theory points to the fact that demand trends, production processes, logistics bottlenecks, and external shocks all interact dynamically. Ivanov (2022) builds on Systems Theory and applies it to a new model of viable supply chain and proves that disruption tends to diffuse non-linearly across the networks. This view questions forecasting methods that only use sales or historical time series with no regard to system interactions at large.

The Dynamic Capabilities View (DCV), a complement to Systems Theory, is a behavioural explanation of the adaptability of firms to forecasting, planning, and decision-making in uncertain situations. Arguing on the same, Munir et al. (2025) assert that organisations need to create the capacity to sense, seize, and restructure resources based on volatile market conditions. Accuracy of forecasting, therefore, is not merely a technical output, but also a strategic capability. Researchers, including Teece (2023), argue that organisational learning and analytical skills, such as the capability to combine machine learning applications and stress tests based on scenarios, play a central role in dynamic capabilities.

Skeptics, however, challenge the operational distinctness of the concept of DCV and indicate that the constructions are still conceptually unclear and immeasurable (Arend & Bromiley , 2009) empirical research on supply chain resilience (Chowdhury et al., 2021) indicates that companies that have stronger analytical and adaptive skills are quicker to recover after a disruption. These two theories together emphasize the fact that the forecast process should be system-wide and capability-based, instead of being a small, statistical task. Figure 4 connects Systems Theory and the Dynamic Capabilities View by showing the connections between forecasting, scenario planning, organisational learning and resilience in a dynamic supply-chain context.

Figure 4. Systems Theory and Dynamic Capabilities in Forecasting (Source: Author's illustration based on Kalaitzi & Tsolakis (2022); Ivanov (2022); Munir et al. (2025); Teece (2023))

2.4.3 Strategic Management and Supply Chain Resilience

The theories of strategic management focus on long-term placement, resource management, and environmental scanning. The application is a popular strategic foresight tool, which is called scenario planning and is heavily influenced by this tradition. According to Schoemaker (1995), the dynamics of a situation contribute to strategic thinking, broadening managerial thinking and enabling companies to envisage non-linear changes. This is an addition to forecasting techniques in that it considers some sorts of alternative futures that are plausible instead of historical trends only.

This strategic view is developed further by the supply chain resilience theory. According to Christopher and Peck (2004), resilience is exposed as the capacity to endure, absorb, and recuperate from disruptants- a feature that is becoming progressively significant in international apparel supply networks. Ivanov and Dolgui (2022) emphasize that the concepts of resilience are founded on redundancy, flexibility, and collaboration, which demand forward-thinking information. Resilient supply chains, therefore, rely on uncertainty-based forecasting systems that include stress testing and adaptive planning.

However, critics caution that resilience efforts may turn out to be expensive and wasteful when overemphasised (Pettit et al., 2010). Hence, strategic management models suggest a striking equilibrium between performance and strength. This trade-off is embodied in the collaboration of ML-based forecasting and scenario planning: high-frequency operational information is given by machine learning, and scenario planning offers strategic resilience to an uncertain future.

Table 9. Predicting strategies, efficiency impacts, and resilience consequences. The synthesis of the author based on Christopher and Peck (2004), Ivanov and Dolgui (2022), and Teece (2023).

Forecasting Strategy	**Primary Goal**	**Efficiency Impact (Cost/Speed)**	**Resilience Impact (Flexibility/Robustness)**
Lean/Statistical (ARIMA/Naive)	Cost Minimization	High: Low computational and data storage costs.	Low: Vulnerable to sudden demand spikes or supply shocks.
Agile/ML-Driven (XGBoost/LSTM)	Accuracy & Speed	Moderate: Requires investment in data pipelines and talent.	Moderate: Adapts quickly to trend shifts but remains backward-looking.
Resilient/Scenario-Integrated	Risk Mitigation	Low: High resource demand for qualitative narrative building.	High: Prepares the firm for "black swan" events and disruptions.
Hybrid (ML + Scenario Planning)	Balanced Optimization	Optimized: High initial cost but reduces long-term "mis-forecast" losses.	Very High: Combines operational precision with strategic foresight.

Resilience-oriented forecasting is required in emerging economies like Bangladesh, where the export of apparel in the global market is experiencing volatility in demand and political unpredictability. This perspective is supported by empirical evidence in the developing markets: it is demonstrated that companies with functional systems of strategic foresight perform better than their counterparts in times of crisis (Sidorkin & Srholec, 2014). All in all, predicting is a strategic, rather than an analytic, resource, as characterized by strategic management and resilience theories.

2.4.4 Synthesis of Theoretical Foundations

Forecasting Theory, Data-Driven Decision Making, Systems Theory, the Dynamic Capabilities View, and Strategic Management/Resilience, provide a strong foundation for an attractive blend of machine learning forecasting with scenario-based planning. Forecasting Theory is about authentic methodology, and data-guided theories result in better adaptability. Systems Theory stresses interdependency, while DCV emphasizes organisational learning and responsiveness. Long-term robustness is offered by resilience and strategic management frameworks. They fall in place to demonstrate the applicability of a hybrid forecasting-scenario planning model which may be beneficial for the apparel industry in Bangladesh dealing with complicated and unpredictable global scenarios. Figure 5 shows the theoretical basis of the thesis, which links forecasting theory, machine learning methods, scenario planning and supply chain resilience.

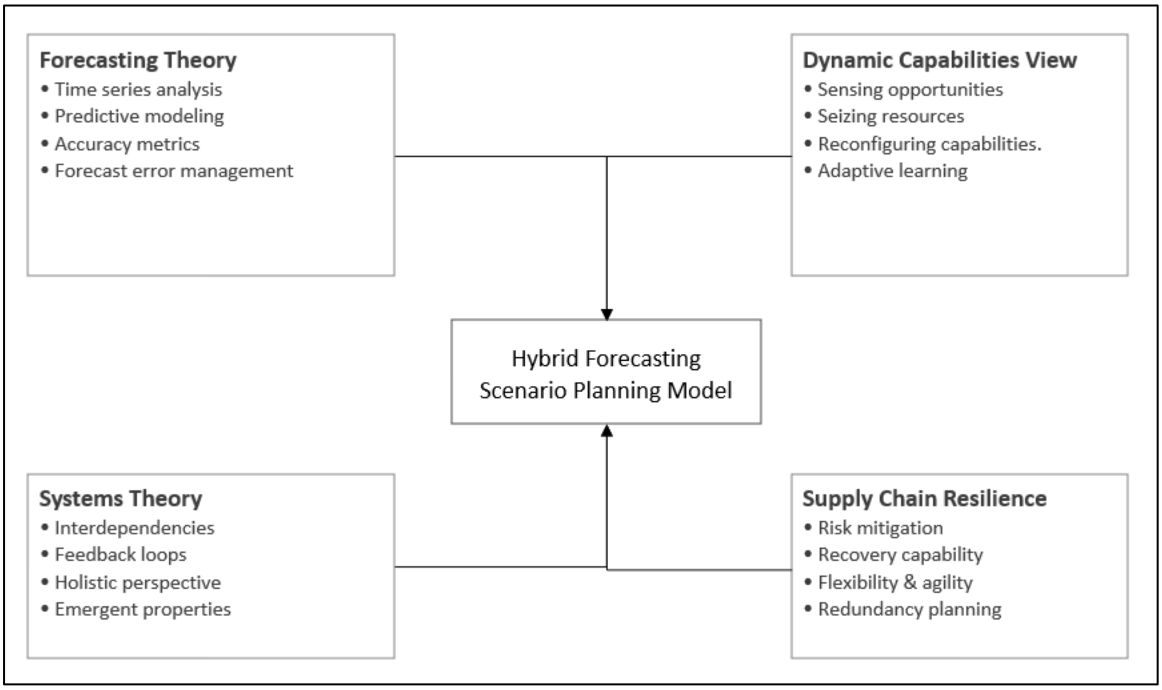

Figure 5. Conceptual/Theoretical Framework (Author's own illustration based on Goel et al., 2024 and Chowdhury & Quaddus, 2017)

2.5 Summary of Literature Review

This chapter deals with the fast development of demand forecasting, machine learning applications, scenario planning, and theoretical underpinnings of forecasting in the garment industry. However, both ARIMA and Holt-Winters are historically useful forecasting models that do not reflect the volatility, seasonality and non-linear demand patterns of apparel markets. As a reaction to this, machine learning methods, in particular tree-based ensembles, neural networks, and hybrid statistical-ML models, have been shown to gain advantages by incorporating complex interaction and large-scale data. However, their performance is largely determined by the quality of data, computational power, and the organizational preparedness. Figure 6 indicates the way the empirical methodology is informed by the literature review through the connection of forecasting models, logic of scenario-planning, evaluation metrics, and operational implications.

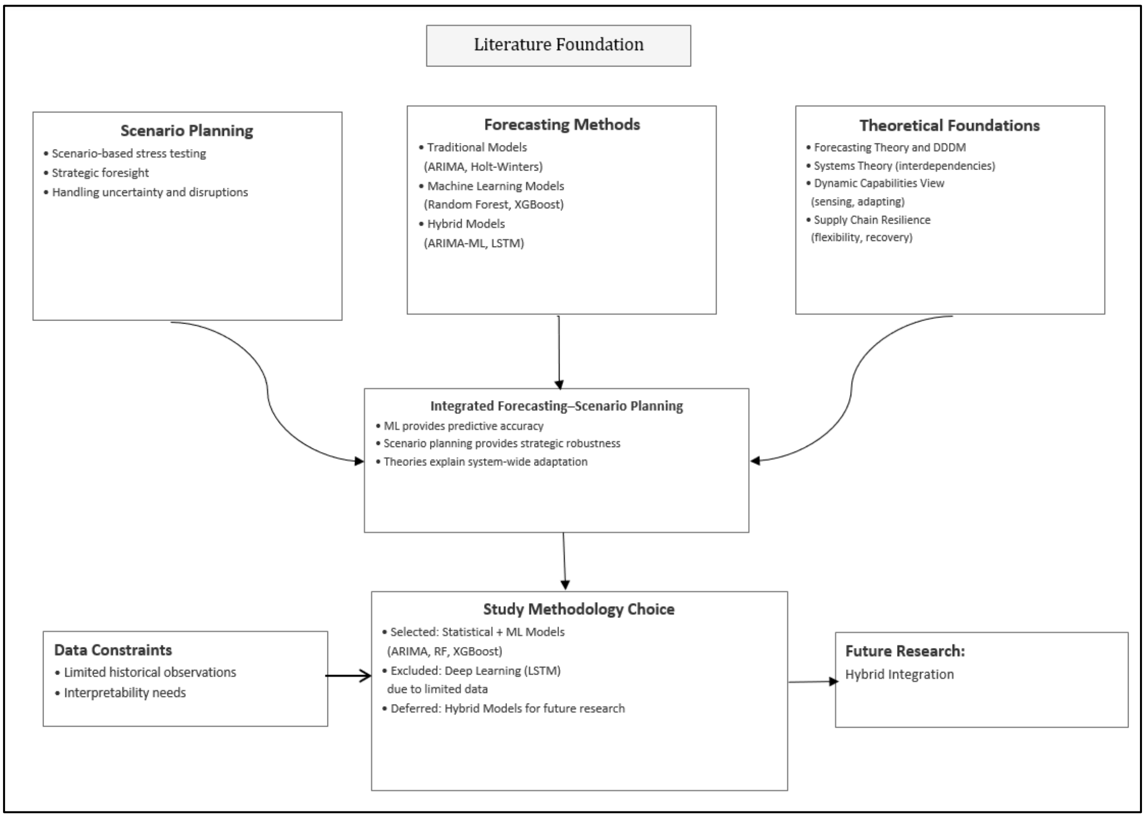

Figure 6. Integrated conceptual framework linking the literature review to the methodology.

Another significant aspect of scenario planning as a strategic instrument as noted in this chapter is its increasing significance as a strategic instrument. Having prediction is not enough in highly uncertain situations in the medium of uncertainty and combined with scenario-based stress testing provides resilience and support the planning decision-making. The integration of forecasting-scenario planning systems by firms also shows that in the context of disruption, they are more resilient to change.

Lastly, it synthesized forecasting theory, information-based decision-making, Systems Theory, Dynamic Capabilities View and strategic resilience literature. These foundations are consistent in their argument of a hybrid forecasting and scenario-planning model to the industries such as the garment industry in Bangladesh where uncertainties, global dependency and a shifting market demand require superior and expected forecasting technologies. Nevertheless, in practice, the existing study priorities are independent statistical and machine learning models rather than hybrid systems due to such practical considerations as the length of historical data and the need that these models are interpretable in manufacturing environments that are constrained by resources. The use of hybrid models is recognized as a positive trend to be improved in the future. Also, although the capabilities of Long Short-Term Memory (LSTM) neural networks are strong in terms of providing sequential time-series predictions, the number of Monthly data points available in this scenario is limited, which constrains their applicability and poses a risk of unstable training; thus, in this study deep learning models are not implemented.

3 Methodology

The chapter introduces the quantitative methodology used for the analysis of the potential impact of machine learning models and scenario planning on the demand forecasting process in the Bangladesh's garment industry. The data used for the analysis is of the company's internal orders collected from January 2023 to January 2026. The forecast comparison is on a monthly basis based on latest ex-fty date. Chronological holdout test period is used to compare traditional models (ARIMA and Holt-Winters) with the new ones (Random Forest and XGBoost). The XGBoost baseline forecast is then stressed using scenario planning (Saunders et al., 2009; Makridakis et al., 2020). Figure 7 gives a summary of the entire research design by linking the methodological decisions and forecasting model comparison to the stage of stress-testing by scenarios.

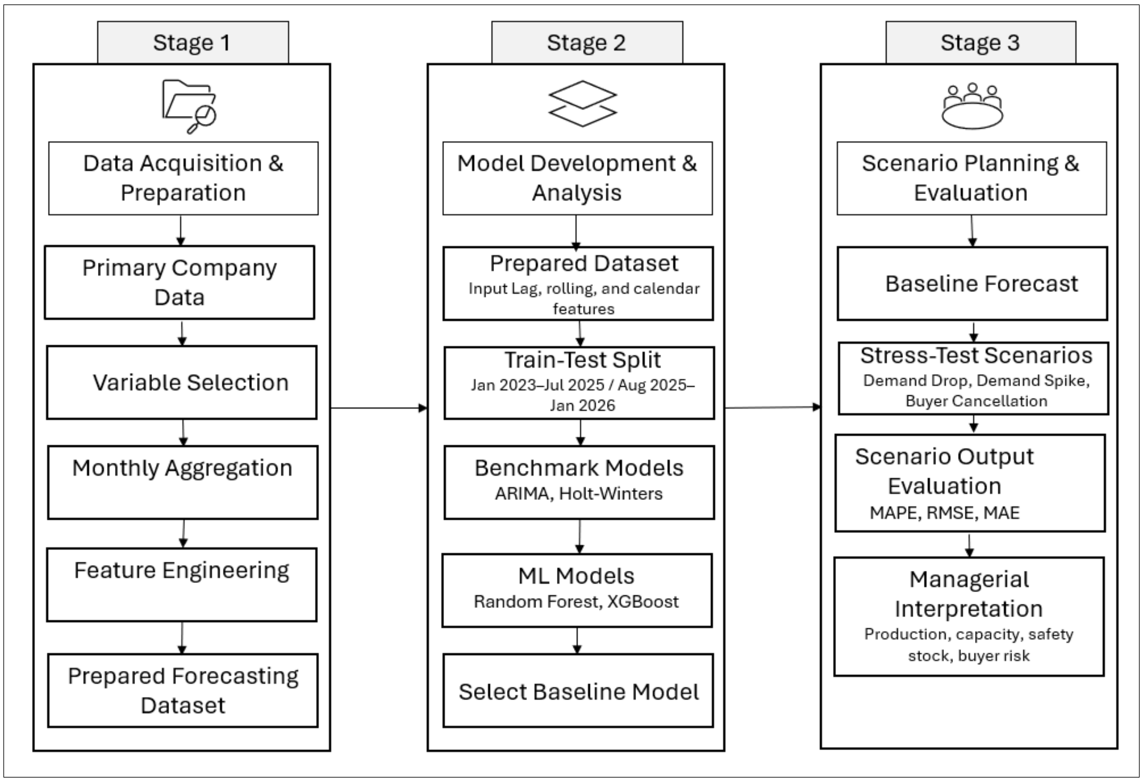

Figure 7. Research Design Framework

3.1 Data Collection

The empirical data is collected from order data of the companies from January 2023 till January 2026. Order Qty is the main forecasting variable with Latest Ex-Fty Date as the temporal index. Customer information is only used for the buyer-concentration scenario, product-category information is only used for descriptive understanding of the order structure. External market disturbances are referred to as contextual risks and are included through scenario interpretation (Yasir et al., 2024; Swaminathan & Venkitasubramony, 2024).

3.1.1 Primary Data: Monthly Order Records

The main source of data is internal order history of a Bangladesh based Garment exporter (Company X) for the period of January 2023– January 2026. This anonymity ensures the principles of privacy, while keeping the manufacturing context intact. The company participating is a medium-sized knit garment manufacturer based in Dhaka, Bangladesh, who provide their products to retailers in EU and UK. Its product portfolio is intimate apparel and related products. The firm was chosen because it had clear data set available for the past and faced varying orders from buyers in different seasons. Order Qty, Latest Ex-Fty Date, Customer ID, Product category and Schedule Status are data elements utilized in the empirical model. Order Qty is summed up into a monthly time series on the company level, based on Latest Ex-Fty Date, for forecasting. Customer information is used to estimate buyer-concentration exposure in the scenario-planning stage. Product category and schedule status are saved for descriptive interpretation, but do not directly used in the final forecasting models (Saunders et al., 2009).

Forecasting analysis is done at the Company level monthly order quantity. This aggregation delivers data of 37 months, while avoiding issues of instability from very intermittent SKU level observations. Product-category information is applied to descriptive segmentation and interpretation of the demand structure (Syntetos et al., 2015; Swaminathan & Venkitasubramony, 2024).

Sampling Logic and Representativeness- This study used a purposive sampling strategy. The study sample, which is Company X, is one company, but as a medium-sized manufacturer directly impacted by the EU/UK retail cycles (major players in Bangladesh's export market), it represents the Dhaka-based knit garment industry. Wider industry-wide averages can miss the more non-linear interactions, so a single-firm operational reality would be analytically scrutiny in detail (Saunders et al., 2009).

The final forecasting comparison is statistically valid at the monthly company level aggregation, as the aggregation helps to reduce the intermittency at the SKU level and thus produces a full time series. The product category detail is kept for purposes of descriptive interpretation

Descriptive Statistics for Monthly Order Quantity: Table 10. provides information on the central tendency and variability of the Order Qty series for the months of January 2023 to January 2026.

Table 10. Descriptive Statistics for Monthly Order Quantity (Jan 2023–Jan 2026)

Metric	Monthly Order Quantity
Number of monthly observations	37
Period	Jan 2023–Jan 2026
Mean	3,154,136
Median	2,836,902
Standard deviation	1,172,802
Coefficient of variation	0.372
Minimum	1,139,760
Maximum	5,821,660

3.1.2 Secondary Data Sources & Industry Reports

The demand situation of Bangladesh's Garment Industry is set in the broader picture through the various secondary sources like industry report, academic literature etc. These sources provide background information on export volatility, uncertainty of buyers, logistic disruption, seasonality and relevance of scenario planning for operations.

They also help interpret the empirical findings, by providing a connection between firm-level forecasting outcomes and broader industry conditions. Hence, in this study, secondary data fortify the industry background, theoretical underpinning and a rationale for scenario-planning (Chowdhury et al., 2021; Milewska, 2022). Table 11 summarizes the key data types, factors, frequency and role of each item of data in the forecasting and scenario-planning analysis

Table 11. Data Description and Source Summary Compiled by Author (2026)

Data Type	Variables Used	Period	Frequency	Role in Analysis
Primary order data	Order Qty, Latest Ex-Fty Date	Jan 2023–Jan 2026	Monthly after aggregation	Forecasting target
Customer data	Customer Short Name, Order Qty	Jan 2023–Jan 2026	Aggregated	Buyer-cancellation scenario
Product data	Product Cat, Schedule Status	Jan 2023–Jan 2026	Descriptive	Contextual segmentation
Calendar features	Month, quarter	Jan 2023–Jan 2026	Monthly	ML predictors
Lagged demand features	Lag-1, Lag-2, Lag-3, rolling 3-month average	Jan 2023–Jan 2026	Monthly	ML predictors

Table 12. Variable Selection Used in the Final Empirical Model

Variable Category	Variable	Role in Model	Analytical Justification
Target variable	Monthly Order Qty	Forecasting target	Represents company-level demand volume
Lagged demand	Lag-1, Lag-2, Lag-3	ML predictors	Captures short-term temporal dependence
Rolling demand	Rolling 3-month average	ML predictor	Captures recent demand momentum
Calendar feature	Month	ML predictor	Captures recurring monthly seasonality
Calendar feature	Quarter	ML predictor	Captures broad seasonal structure
Customer concentration	Largest buyer share	Scenario input	Captures buyer-cancellation exposure

Table 12 indicates the target variable, machine-learning predictors, and scenario input that will be used in the empirical analysis.

3.1.3 Variable Selection and Feature Engineering

The transaction-level data were aggregated to the monthly level with Latest Ex-Fty Date as the time index, as a first step of data preprocessing. The forecasting target is Order Qty which is a company level aggregation for the forecasting period from January 2023 to January 2026. This aggregation generates a continuous 37-month demand series which can be compared to the traditional statistical models and machine learning models. The emphasis on order quantity captures the demand that can be found in the company's order book, and information on the product category is preserved to facilitate descriptive interpretation of the order structure (Yasir et al., 2024; Gaertner et al., 2024).

The dataset captures 37 monthly observations from January 2023 to January 2026 for statistical reliability. The sample consists of three sequential seasonal cycles and makes an adequate foundation to train exploratory models with Random Forest and XGBoost. This enables the models to detect recurrent temporal and seasonal behaviors that would otherwise not be detectable in smaller datasets.; when aggregated, 37 months data span is not a big data set but still enough right grounds for making an exploratory analysis between the classical statistical models vs. Machine Learning forecasting models performance comparison (Gaertner et al., 2024; Kontopoulou et al., 2023).

Empirical Justification for Lags and Features

Lag-1, Lag-2 and Lag-3 have been incorporated to incorporate the short-time dependence in the monthly order series. Lag-1 indicates the current persistence of orders, that is, whether order quantity in one month might affect order quantity in the next. Lag-2 and Lag-3 add more time dimension and record the order effects in the last two and three months, respectively, which is applicable in the garment industry where the change of order occurs may not be immediate from one month to another. In addition to these lagged variables, a three-month moving average of the variables was added as a proxy for the demand momentum in the latest three months and to smooth short-term fluctuations in the order series. These features together give the machine learning models structured information around the recent demand behaviour and are still directly derived from the historical order data (Yasir et al., 2024; Ma et al., 2025).

Justification of Sample Size and Overfitting Mitigation

Although the 37-month data (January 2023 to January 2026) is not a large dataset to train a deep learning algorithm, it can be used to conduct an exploratory comparison of the classical and ensemble machine learning models. To reduce the risk of overfitting:

Model Selection: Deep learning models such as LSTM, were not used since they need much larger datasets; rather, Random Forest and XGBoost models were selected due to their ability to be trained on smaller longitudinal samples (Gaertner et al., 2024; Douaioui et al., 2024).

Holdout validation: The data was split chronologically between the training period and holdout test period to assess out-of-sample forecasting performance. This is a way of maintaining the sequence of the order series in the months and to mimic what is actually done in forecasting: predicting the next demand based on data at hand. Due to the small sample size of 37 months, a simple chronological holdout design is used to compare the performance of the ARIMA, Holt-Winters, Random Forest, and XGBoost without adding an unnecessary model complexity (Makridakis et al., 2020; Swaminathan & Venkitasubramony, 2024).

Scope & Value: The 37-month window gives just three full seasonal cycles, giving the models a chance to learn the recurrent fashion surges without the noise that much older, pre-disruption historical data would have. The lagged, rolling and calendar-based variables are described in Table 13 to explain the monthly order-quantity series used in the Random Forest and XGBoost modelling.

Table 13. Feature Engineering Variables

Feature Name	Type	Description	Transformation Applied
Lag-1	Numerical	Previous month’s order quantity	One-month shift
Lag-2	Numerical	Order quantity two months earlier	Two-month shift
Lag-3	Numerical	Order quantity three months earlier	Three-month shift
Rolling 3-month average	Numerical	Average of the previous three months	Rolling mean using lagged values
Month	Calendar	Month of the year	Integer 1–12
Quarter	Calendar	Quarter of the year	Integer 1–4

Feature Selection and Importance Ranking

This research does not consider all the variables equally.

Feature importance: Following the training of the model, this technique was applied to determine which predictors were most important for XGBoost's prediction. The interpretation is centered around the time and calendar variables that were observed in the final model: lags of order quantities, three-month rolling demand momentum, month and quarter. This feature-importance output also aids in the interpretability of the machine learning output, by providing information on how the model applied recent demand behaviour and recurring seasonal patterns to forecasts (Chen & Guestrin, 2016; Douaioui et al., 2024).

Dimensionality Reduction: The data is aggregated into a monthly order series at the company level for the purpose of the study. This aggregation offers a steady forecasting target and represents a solution to statistical instability of highly intermittent observations at SKU level being modelled separately. It also allows easier comparison between the traditional statistical models, and the machine learning models in the current data set of 37 months.

3.1.4 Data Cleaning and Preparation

Data preparation started with the transformation of the time series Latest Ex-Fty Date to a monthly time index. Records that had an Order Qty from January 2023 up to January 2026 were kept and Order Qty was summed up at a monthly level to create the primary forecasting series. Historical order quantities were used to create the monthly series and then the lagged variables and rolling three-month average. Reproducible Preprocessing Pipeline. In order to ensure that all models would be trained and tested on complete observations, rows where one or more of the lags were not available due to the construction of the features were omitted prior to estimation.

Reproducible Preprocessing Pipeline

The last month of the monthly series was divided into a training period and a holdout test period. Training will be conducted from January 2023 to July 2025 and holdout testing will be conducted between August 2025 and January 2026. This structure helps to maintain the chronological order of time and is representative of the actual forecasting context, where future demand is to be estimated from past observations. Backward looking lag and rolling variables were used as features, meaning that the predictors for each month were chosen based on the previous observations in the order series. All tree-based machine learning models were estimated using the original feature scales and the same chronological split was kept in the entire set of models in order to be able to compare them (Makridakis et al., 2020; Kontopoulou et al., 2023).

Preprocessing Pseudo-code

A high-level logic flow enhances scientific transparency:

Algorithm: Data Preparation Pipeline

LOAD company order dataset.

CONVERT Latest Ex-Fty Date to datetime format and filter records from January 2023 to January 2026.

AGGREGATE Order Qty by month.

CREATE Lag-1, Lag-2, Lag-3, rolling 3-month average, month, and quarter features.

SPLIT data chronologically into training and holdout test periods.

TRAIN ARIMA, Holt-Winters, Random Forest, and XGBoost; EVALUATE models using MAE, RMSE, and MAPE.

3.2 Data Analysis

The analysis will contrast the conventional time-series forecasting models against machine learning models and then perform stress testing under uncertainty involving scenarios. To maintain the time series of the monthly order series and prevent data leakage, a chronological train-test split was employed. The training period will be January 2023 to July 2025, and the holdout test period will be August 2025 to January 2026. This design is indicative of a real-world forecasting environment, where future quantities of orders must be forecasted using past available data. The final six months of monthly order quantity are assessed on model performance based on MAE, RMSE and MAPE (Makridakis et al., 2020; Goel et al., 2024).

Overall, the analysis is structured around two main hypotheses:

H1: Machine learning models, especially Random Forest and XGBoost, will have a lower forecast error than traditional statistical models, ARIMA and Holt-Winters, when used to predict the monthly order-quantity series.

H2: Scenario planning implemented on the XGBoost base forecast will improve its operational adaptability and will show the impact that the assumptions of demand-drop, demand-spike, and buyer-cancellation have on the predicted volume of orders.

Defining "Improvement"

The multi-dimensional definition of improvement in this research is:

Accuracy: Get a lower Mean Absolute Percent Error (MAPE) on the holdout.

Robustness: Robustness is understood as the capacity of the forecasting framework to test alternative demand results in a systematic fashion by the use of scenario analysis.

This is evaluated in this research by imposing the assumptions of demand-drop, demand-spike and buyer-cancellation to the XGBoost baseline prediction.

Interpretability: The model should be able to provide actionable knowledge through Feature Importance rankings.

3.2.1 Traditional Forecasting Models (ARIMA, Holt-Winters)

Autoregressive Integrated Moving Average (ARIMA) and Holt-Winters exponential smoothing have been utilized as benchmarks since they continue to be extensively utilized to forecast apparel (Swaminathan & Venkitasubramony, 2024). ARIMA can effectively represent autocorrelation and linear temporal patterns, but those characteristics that change the demand suddenly and long-term seasonality are not as well-covered (Makridakis et al., 2020). Holt-Winters can handle seasonality but cannot cope with non-linearity and structural break sensitivity.

ARIMA(1,1,1) and Holt-Winters exponential smoothing were used as the traditional forecasting benchmarks. ARIMA(1,1,1) is the autoregressive time-series model, and Holt-Winters is a seasonal smoothing model. Their incorporation enables the machine learning models to be compared with existing and interpretable forecasting models that are usually applied in demand forecasting research (Makridakis et al., 2020; Swaminathan & Venkitasubramony, 2024).

Technical Specification and Diagnostics:

In the case of ARIMA benchmark, the stationarity was evaluated with the help of the Augmented Dickey-Fuller test, and the existence of residual autocorrelation was determined with the help of the Ljung-Box test. Holt-Winters exponential smoothing was approximated and additive trend and multiplicative seasonality where estimation of the model was possible. The statistical foundation of the comparison of traditional forecasting performance with the Random Forest and XGBoost is offered by these diagnostics and benchmark specifications. Figure 8 shows the analytical process of data preparation and benchmark modelling followed using the ensemble forecasts, model assessment and operational interpretation of each scenario.

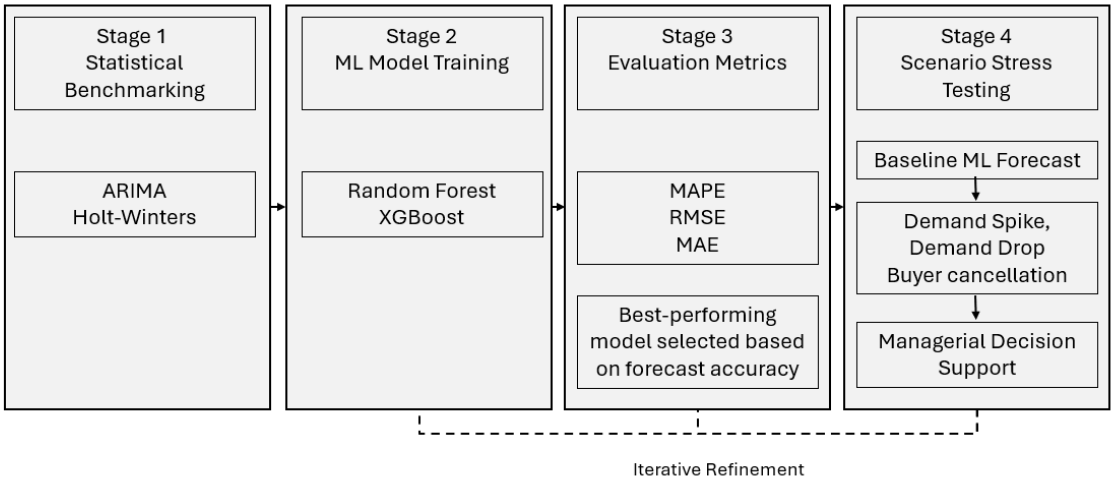

Figure 8. Structured Analytical Workflow – Sequential integration of statistical benchmarking, machine learning development, and scenario-based stress testing.

3.2.2 Machine Learning Models (Random Forest, XGBoost)

Random Forest and XGBoost were used as supervised regression models with time related and calendar-based predictors based on the monthly order series. The input characteristics are Lag-1, Lag-2, Lag-3, rolling three-month average, month, and quarter. These variables enable the models to acquire recent momentum of the order, delayed demand effects, and repetitive seasonal patterns in a more flexible manner compared to the traditional linear time-series models (Douaioui et al., 2024; Yasir et al., 2024).

The choice of Random Forest and XGBoost was based on the fact that these are the most appropriate models to be used to model non-linear relationships between engineered features and monthly order quantity. Random Forest employs multiple decision trees to minimize variance and enhance stability, whereas XGBoost employs sequential boosting to address the previous projection mistakes and enhance the accuracy of the forecast (Chen & Guestrin, 2016; Douaioui et al., 2024). The tree-based ensemble models are appropriate since they offer a proper balance between modelling flexibility and methodological transparency with the availability of the 37-month dataset.

In the same vein, despite having been shown to exhibit better forecasting accuracy in numerous studies, hybrid statistical ML models (e.g., ARIMA-NNA, ARIMA-ML, XGBoost, etc.) require longer historical series, higher computational resources and introduce interpretability challenges. Thus, the proposed study covers standalone models to guarantee methodological transparency and adherence to the characteristics of the available dataset. Hybrid modelling has a way to be developed and compared in the future.

Table 14. Forecasting Models Used Compiled by Author (2026).

Model Category	**Model Name**	**Key Assumptions**	**Strengths**	**Limitations**
Traditional Statistical	ARIMA	Linear trends, stat. stationarity	Interpretability	Poor performance with shocks
Traditional Statistical	Holt-Winters	Strong seasonal patterns	Simple, fast	Struggles with complex data
Machine Learning	Random Forest	Non-linear relationships	Robust to overfitting	Requires extensive data
Machine Learning	XGBoost	Boosting for accuracy	Top-tier performance	High computation cost

Model Parameter Settings

Fixed parameter settings were used to estimate the machine learning models so that consistency and reproducibility was achieved throughout the model comparison. Random Forest has been estimated using 500 trees, square-root feature selection, and minimum leaf size of 4. XGBoost was trained using 200 estimators, with a learning rate of 0.05, a maximum depth of 6, a subsampling rate of 0.8, and using a squared-error regression goal. The machine learning models used a fixed random seed to allow reproducible results. Table 14. provides a summary of the four forecasting models that are used in the study, its assumptions, strengths and limitations.

The six-month chronological holdout test period was used to assess model performance. Along with the metrics of the accuracy, the XGBoost feature importance was utilized to explain the relative importance of the lagged order quantities, rolling demand momentum, and calendar-based predictors to the final forecasts.

Predictive results obtained by these algorithms can be dynamically adjusted to new patterns, and they must comply with trends towards digitalization of supply chains and Industry 4.0 (Goel et al., 2024). The rationales behind the choice of these models are fully consistent with the Systems Theory and Data-Driven Decision-Making concepts presented in Chapter 2, and which ensures that the methodology is based on established forecasting principles.

3.2.3 Scenario Planning Framework (Demand Shocks, Spikes, Cancellations)

As a deterministic stress-testing layer on top of the XGBoost baseline forecast, the scenario-planning stage is used. Once the baseline forecast is generated by the holdout period, three other demand conditions are tested: a 25% demand downward, a 25% demand upwards, and order exposure cancellation of the biggest buyer. This structure allows the forecasting analysis to be further extended beyond the analysis of single point prediction by depicting how the anticipated order volume varies with varying stress conditions of operations.

Cordova-Pozo and Rouwette (2023) justify the application of scenario planning as a predictive tool in the analysis of uncertainty and enhancing strategic responsiveness. The XGBoost forecast will provide the base demand estimate in this analysis and the scenario layer will convert the estimate to realistic stress-test results. The scenarios are modeled according to the demand-side uncertainty, exposure to buyer-concentration, which is directly applicable to the export-based garment production (Klibi & Martel, 2012; Ivanov & Dolgui, 2022).

The adjustment of the baseline forecast is done in the following manner:

a) The scenario of demand-drop: a 25 percent decrease of the XGBoost baseline forecast.

b) Demand-spike: 25% rise over the XGBoost baseline forecast.

c) Buyer-cancellation: A cut of the biggest buyer of the total order book. To keep the information confidential the buyer is anonymised with Buyer A.

The scenarios stress tests are applied to analyze the impact of other demand conditions on production planning, safety-stock evaluation, and capacity alignment. Using these assumptions of stress to the XGBoost baseline forecast, the analysis links forecast accuracy to realistic planning issues in an export-manufacturing setting. This step thus further builds on the model comparison by demonstrating how the chosen forecasting output can aid in operational decision-making amidst demand uncertainty (Ivanov & Dolgui, 2022; Zahraei & Teo, 2017).

Table 15. Scenario Planning and Stress-Test Design (Cordova-Pozo and Rouwette, 2023)

Scenario	**Adjustment**	**Variable Adjusted**	**Purpose**
Baseline	0%	XGBoost forecast	Expected order-volume forecast
Demand drop	-25%	Baseline forecast	Tests downside demand risk
Demand spike	+25%	Baseline forecast	Tests upside capacity pressure
Buyer cancellation	-largest buyer share	Baseline forecast	Tests buyer-concentration exposure

Scenario Assumption Basis

The demand-drop and demand-spike cases are based on symmetric 25% changes to the XGBoost baseline forecast. The buyer-cancellation scenario is based on customer concentration in the observed order book. The total Order Qty is used to determine the share of the largest buyer that is then used as a stress factor on the basis forecast. This will enable the scenario analysis to focus on buyer-concentration exposure but maintain confidentiality with the anonymised label Buyer A. The baseline, demand-drop, demand-spike, and buyer-cancellation scenarios are summarised in Table 15

3.2.4 Evaluation Metrics & Validation Strategy

The models are tested on a multi-metric basis to measure various aspects of error, ensuring that the chosen models are both operationally useful and statistically sound.

Rationale for Metric Selection

The outlined metrics were selected to overcome the high variance of the Bangladesh garment export sector:

MAPE (Mean Absolute Percentage Error): MAPE was chosen as the main measure of accuracy since it represents the error in forecasting an order as a percentage of the actual order. This simplifies the interpretation of the result at various levels of demand and offers a clear measure of proportional forecasting error using which the managers can make decisions.

RMSE (Root Mean Square Error): RMSE was used since it assigns more importance to huge forecasting errors. In garment production planning, a large forecasting error will have more operational pressure than a series of smaller errors particularly where the volume of orders influences capacity allocation, the preparation of materials and delivery planning.

MAE (Mean Absolute Error): MAE was employed to estimate average error in the forecasts in actual order units. This measure is advantageous as it converts model performance into the same unit as production and order planning, and the degree of error is an easier measure to interpret on the basis of operational concerns. The forecast-error measures that were employed to measure the models are shown in Table 16 and the analytical purpose of each measure is explained.

Business-Relevant Evaluation

The forecast accuracy of the forecasting models is evaluated against the interpretation of production-planning. The smaller forecasting error makes it easier to conduct a more informed debrief on the production scheduling, safety-stock analysis, material preparation and capacity alignment. MAE, RMSE, and MAPE usage thus enable the statistical outcomes to be viewed as not only indicators of model-performance, but also as planning-relevant indicators to the garment manufacturing environment (Chopra & Meindl, 2016; Zahraei & Teo, 2017).

Table 16. Evaluation Metrics for Forecast Accuracy (Makridakis et al., 2020)

Metric	Analytical Purpose
MAE	Average absolute error in order units
RMSE	Penalises large forecast errors
MAPE	Percentage error relative to actual demand

3.3 Validity, Reliability, and Ethical Considerations

Ensuring methodological integrity was central given the commercial sensitivity and decision-critical nature of forecasting in manufacturing.

The approach taken in this research focuses on methodological integrity, as it is important to guarantee the findings are scientifically sound and operationally feasible in the high-stakes setting of garment manufacturing.

3.3.1 Data Confidentiality & Ethical Use of Company Data

Data used in the company was anonymized and stored safely so that the identifiers would not be sensitive to the business's confidentiality. There was limited access and ethical approval was also in line with existing governance practices in industrial research. Moral management helps in fostering the principles of credibility and adherence to data responsibility in corporate partnerships

To enhance the validity of the study, the ethical and compliance guidelines outlined below were strictly adhered to:

Institutional Review: The research was performed in accordance with formal academic ethics rules, with the emphasis on the non-interventional data analysis.

Data Handling Protocol: MIS records were transferred as part of a signed Non-Disclosure Agreement (NDA) with Company X. Information was exchanged over an encrypted channel and the data itself was stored in a confidential password protected medium.

Anonymization Techniques: all buyers' names, specific factory unit addresses, and unit prices were redacted or replaced with a generic name (e.g., Buyer A, Category 1).

3.3.2 Analytical Validity and Reliability

All four forecasting models were supported by the same training and test period in terms of chronology, to help with the analytical validity. ARIMA, Holt-Winters, Random Forest and XGBoost were thus tested on similar terms with the same monthly order-quantity series and holdout period. A well-documented modelling process that encompassed the feature-engineering process, fixed parameter settings, and random seed of the machine learning models increased its reliability (Saunders et al., 2009; Makridakis et al., 2020).

3.3.3 Reproducibility & Transparency in Data Processing

The empirical workflow was coded in Python with pandas, statsmodels, scikit-learn, XGBoost and matplotlib. The data workflow includes the monthly data aggregation, feature building, model estimation, performance test, stress testing scenarios and figure generation. This documentation can be used to reproduce the analysis path used in the original company order records to the end model comparison and scenario planning products.

3.3.4 Limitations of Methodological Approach

Several limitations were recognized. Our internal data sets cannot fully represent the underlying demand dynamics of the sector, as these vary by factory levels and market categories. Another recognised challenge is the interpretability of ML (Chopra and Meindl, 2016), although it is reconciled to use features as costs for apparent mitigation. The estimates in the scenarios could be updated as environmental conditions change over time around the globe. Notwithstanding, the hybrid approach should create a functional and viable forecasting tool that can be adapted to meet the distinctive characteristics of the Bangladesh garment export market.

The research recognizes several limitations that guide the interpretation of the research results:

Sample Size & Generalizability:

The 37-month dataset is a machine learning small-data situation. Although it is enough to use the ensemble methods, it restricts the application of deep learning. To alleviate this, the research concentrates on the robustness of the model rather than crude complexity.

Single-Case Design:

Although the study is able to analyse the operational reality in depth when the focus is on a single manufacturer, it can restrict the generalizability of the results to the whole sector of Bangladesh. This framework needs to be extended to a multi-firm comparison in future research.

Model Bias: It is possible to have a bias towards stable periods. In response to this, the Scenario Planning Framework was explicitly included to experiment on the performance of the models when the tail events (shocks) that a 37-month history may under-represent occur.

The approach to the study is defined by the organization of the company data at hand. The final forecasting models incorporate the temporal and calendar-based demand items such as lagged order quantities, rolling demand momentum, month and quarter. The contextual discussion and scenario-planning rationale include wider industry influences like freight costs, buyer-market conditions, inflation and logistics disruption to aid in the interpretation of the forecasting results in the Bangladesh garment-industry environment.

4 Data Analysis and Results

4.1 Introduction

The chapter presents computational findings and diagnostic outputs from predictive models implemented in the Bangladesh garment industry. Whereas the previous chapters identified the theoretical need of the Systems Theory and Dynamic Capabilities Teece (2023), Chapter 4 switches to the quantitative validation of these frameworks. In line with the Pragmatism philosophy of research, this analysis will focus on the practical usefulness of machine learning (ML) architecture rather than on purely theoretical constructs, gauged by their ability to reduce forecasting errors in a high-volatility manufacturing environment.

The monthly order quantity at the company is operationalised to empirically represent the theoretical framework in the analysis as the empirical output of the volatile garment supply chain system. The forecasting models apply historical patterns of orders, lagged demand, rolling momentum of demand and calendar-based indicators, to analyse the responses of various modelling styles to irregular demand behaviour. This ties the empirical research to the wider explanation of the forecasting theory, evidence-based decision making and flexibility of supply-chain. A deterministic layer of stress-testing, scenario planning is then applied to investigate the effects of alternative demand conditions on the interpretation of the baseline forecast.

The reason why the ensemble learning is used together with traditional benchmarks is because of the forecasting literature which suggests that linear models like ARIMA can be constrained in the case of demand lumpy, seasonal and non-linear. The analysis compares ARIMA and Holt-Winters to the random forest as well as XGBoost to determine the forecasting accuracy at the same holdout test period. Scenario based stress testing is then used to extend the best-performing machine learning forecast to enable the results to include both predictive accuracy and operational adaptability. The month-to-month change in the company level order quantity is presented in figure 9, and it is against this that the forecast will be compared.

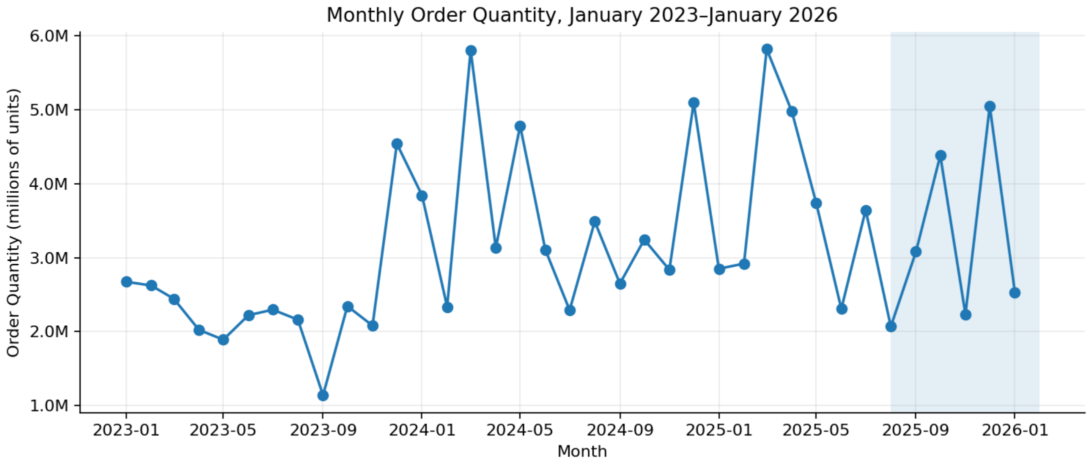

Figure 9. Monthly Order Quantity per Month (January 2023-January 2026). Compiled by the author using company order records from the subject organisation (2026).

4.2 Descriptive Analysis of the Dataset (Order Quantity, Product Structure, and Seasonality)

This section describes the longitudinal company data which were used in the empirical analysis. The dataset consists of 37 monthly observations January 2023 and January 2026 and represents the trend of the works of Company X, a medium-size knit garment manufacturer that is headquartered in Dhaka. The empirical forecasting series will be developed by summing up company level Order Qty based on Latest Ex-Fty Date creating a continuity of monthly series of order demand over the entire study period. According to Figure 10, the most significant predictors in the XGBoost model were the recent order momentum and monthly seasonality. The descriptive analysis, comparison of models and scenario-based stress testing in this chapter are based on this structure.

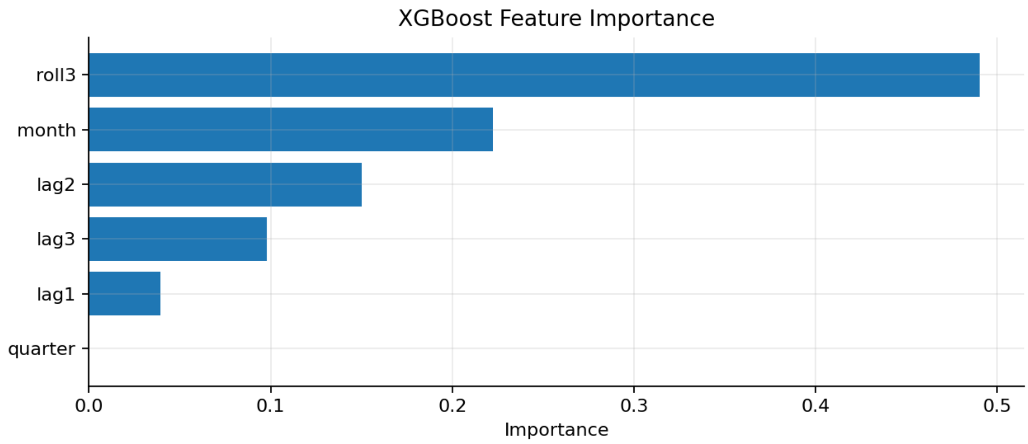

Figure 10. Relative Importance of Temporal and Calendar-Based Features in the XGBoost Forecasting Model. Author's calculation using Python (Scikit-Learn/XGBoost library).

4.2.1 Statistical Properties and Stationarity Testing

Descriptive diagnostics and stationarity tests were carried out in advance of the estimation of the model in line with Forecasting Theory. The series of monthly Order Qty includes 37 observations between January 2023 and January 2026. The average amount ordered monthly is 3,154,136 units, the median is 2,836,902 units and the standard deviation is 1,172,802 units. The coefficient of variation stands at 0.372, which means that there is a lot of volatility in terms of the order series on a monthly basis.

Stationarity Testing: The Augmented Dickey-Fuller test was used to train series to determine whether the order series was stationary prior to benchmark modelling. The test generated a statistic of about -4.193 and a p-value of about 0.0007, which was an indication of stationarity in the level series in the period of training. First difference also gave an outcome of a stationary series with p-value of about 0.0331.

ARIMA Benchmarking: ARIMA(1,1,1) was considered the traditional autoregressive benchmark and tested against the same holdout period as the other forecasting models.

This gave it a uniform statistical foundation on which to compare it with Holt-Winters, Random Forest, and XGBoost.

Features of ML Models: Lag-1, Lag-2, Lag-3, rolling average of three months, month, and quarter were used as input features in the machine learning models. These variables are short run temporal dependence, recent demand momentum and calendar effects in the monthly order series.

4.2.2 Product Segmentation and Seasonality

The descriptive explanation of the order structure of the company was supported with the help of product-category information and the forecasting comparison was made on the company-level monthly series of Order Qty. This combination offers a 37 month period of continuous demand and a consistent foundation to compare ARIMA, Holt-Winters, Random Forest and XGBoost. It also minimises the instability which can arise when highly intermittent SKU- or category-level observations are fitted separately.

The Systems Theory of Seasons and Systems:

The seasonality analysis reveals that the calendar effects in the order series are monthly. This is evident in machine learning set of features, which added month and quarter indicators to represent the repetitive seasonal trends. As a Systems Theory part of the broader production and demand environment, these calendar effects are manifestations of where garment orders are placed, processed, and delivered as a part of seasonal cycles.

Volatility and Strategic Resilience:

The volatility in the monthly order series is observed and is in line with the uncertainty that is depicted in the literature on garment supply chains This kind of volatility justifies the application of the overlay of scenario planning to the underlying forecast. The XGBoost forecast in this study is stress-tested in three demand-side conditions: a 25% drop in demand, a 25% spike in demand, and a scenario where the biggest buyer cancels.

4.3 Model Results and Forecast Comparisons

This section compares the four selected models for predictive accuracy. All models were evaluated using the same chronological holdout test period, covering August 2025 to January 2026, against monthly Order Qty values that were not used during model. The comparison of the actual holdout-period order quantities and the predicted quantities is shown in figure 11

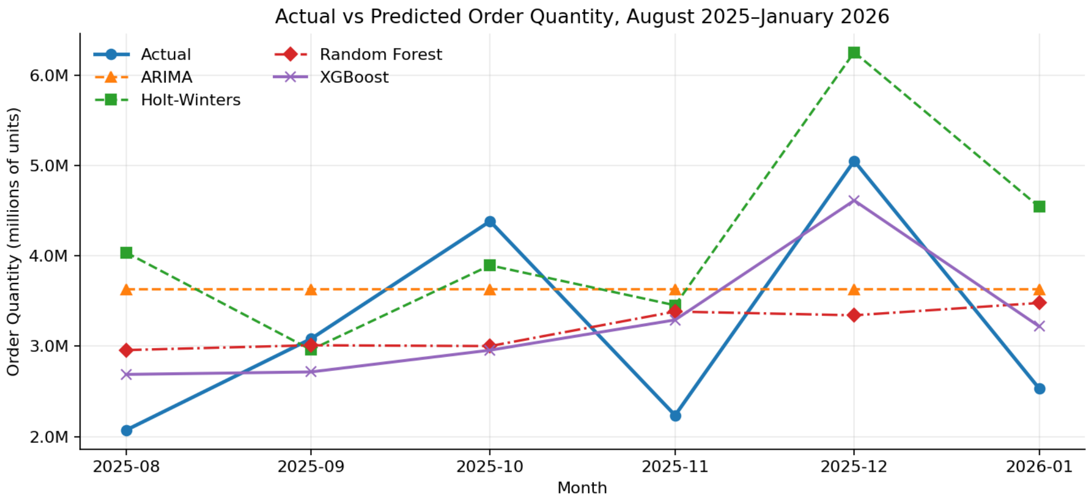

Figure 11. Comparison between Actual Order Quantity and Predicted Values During the Holdout Test Period. Produced by the author (2026).

4.3.1 Evaluation of Traditional Benchmarks (ARIMA and Holt-Winters)

The modelling design pursued is consistent with the traditional forecasting models which were initially estimated as a baseline of assessing the monthly order-quantity series. The models offer an explanatory point of reference and then the machine learning models are compared to the performance of these models.

ARIMA(1,1,1) Performance and Diagnostics: ARIMA(1,1,1) was applied as the classical autoregressive model. For the August 2025–January 2026 holdout period, the model produced a MAPE of 40.75%, RMSE of 1,189,013 units, and MAE of 1,129,948 units. The p-value of the Ljung-Box test at lag 10 was about 0.4938, which shows that there was no significant residual autocorrelation left. Despite the good residual diagnostics, the scale of month-to-month variation in the order series is not well followed by ARIMA as indicated by the forecast errors.

Holt-Winters Exponential Smoothing: Holt-Winters was used as the seasonal statistical benchmark. In the holdout test period, it produced a MAPE of 44.68%, RMSE of 1,360,989 units, and MAE of 1,168,825 units. It was not as good as ARIMA and both machine learning models, which means that seasonal smoothing did not fully explain the volatility that appeared in the monthly order-quantity series.

H1 is supported by the benchmark results since the forecast errors of both the Random Forest and XGBoost were lower than that of ARIMA and Holt-Winters in the holdout period. This makes the traditional models useful as points of comparison and validates the more predictive ability of the machine learning models in this dataset. As indicated in Table 17, the most significant predictors in the XGBoost model include the rolling three-month average and the month variable, which means that the momentum of demand in the recent past and monthly seasonality are important.

Table 17. Relative Feature Importance in the XGBoost Model.

Feature Rank	**Variable Name**	**Importance Score**	**Description**
1	Rolling 3-month average	0.490	Captures recent order momentum
2	Month	0.222	Captures recurring monthly seasonality
3	Lag-2	0.150	Captures delayed short-term order dependence
4	Lag-3	0.098	Captures longer short-term memory
5	Lag-1	0.039	Captures immediate prior-month persistence
6	Quarter	0.000	Adds little explanatory value after monthly effects

4.3.2 Comparative Performance of Ensemble Machine Learning Models

The ensemble machine learning models were applied to test whether it was possible to capture non-linear temporal patterns in the garment order series using the ensemble machine learning models more effectively than the traditional statistical benchmarks. The supervised regression models estimated the random Forest and XGBoost models were based on lagged order quantities, rolling demand momentum, and calendar indicators. These characteristics enable the models to take into consideration the recent behaviour of demands and repetitive monthly occurrences in the order sequence.

Random Forest was better than the traditional benchmarks. It achieved a MAPE of 33.24%, RMSE of 1,143,463 units, and MAE of 1,024,590 units. This suggests that the bagging-based ensemble framework could acquire more non-linear data out of the engineered time elements. Nevertheless, Random Forest was still not as accurate as XGBoost in the holdout period.

Table 18. Fixed Machine Learning Model Parameter Settings

Model	**Hyperparameter**	**Model Setting**
XGBoost	n_estimators	200
	learning_rate	0.05
	max_depth	6
	subsample	0.8
Random Forest	n_estimators	500
	max_features	sqrt
	min_samples_leaf	4

XGBoost worked the best of the four tested models. It achieved a MAPE of 26.25%, RMSE of 850,095 units, and MAE of 765,816 units. The outcome indicates that the rolling three months average, monthly seasonality, and lagged order effects were better utilized with XGBoost as compared to the other models. This observation justifies the application of boosted tree-based learning as the chosen baseline forecast to use in the stage of scenario-planning. Table 18 reports the fixed setting of the parameters of the Random Forest and XGBoost settings, which can be used to facilitate the reproducibility and consistency of the empirical model comparison

Table 19 makes a comparison of ARIMA, Holt-Winters, Random Forest, and XGBoost based on MAPE, RMSE, and MAE and indicates that XGBoost had the lowest forecast error.

Table 19. Comparison of Model Performance Metrics.

Model Type	**MAPE (%)**	**RMSE (Units)**	**MAE (Units)**
ARIMA(1,1,1)	40.75%	1,189,013	1,129,948
Holt-Winters	44.68%	1,360,989	1,168,825
Random Forest	33.24%	1,143,463	1,024,590
XGBoost	26.25%	850,095	765,816

Python Model Execution Summary

The analysis was implemented empirically in Python with the help of pandas, statsmodels, scikit-learn, XGBoost, and matplotlib. The monthly Order Qty series were separated by time, where January 2023-July 2025 was the model training set and August 2025-January 2026 was the holdout test set. ARIMA and Holt-Winters were estimated using statsmodels, whereas Random Forest and XGBoost were estimated using XGBoost and scikit-learn libraries. The XGBoost had been configured as a squared-error regression model with 200 estimators, learning rate of 0.05, maximum depth of 6 and subsampling rate of 0.8. The results of this implementation environment were the model-performance output and forecast comparison results shown in this chapter.

4.3.3 Scenario Planning and Operational Stress Testing

The operation stress testing in form of scenarios was based on the baseline XGBoost forecast. The six-month holdout between August 2025 and January 2026 was explored on the three demand conditions: a 25 percent demand decline, a 25 percent demand burst, and the cancellation of the biggest buyer order exposure. Such scenarios are an extension of the base forecast to the other operating conditions to ensure the analysis examines the risk of the downside of the demand, the volume of the upside pressure and the exposure associated with the concentration of buyers.

Scenario A: Reduction in demand (-25%).

The demand-drop scenario uses a 25% weighted down XGBoost baseline forecast. The six-month forecast was 19,482,392 units. In this case, the volume forecasted is lower to 14,611,794 units, a change of -4,870,598 units. This finding gives an idea of the magnitude of order-volume exposure that can occur when buyer demand decreases significantly. This adjustment is arrived at by multiplying the baseline forecast by a stress factor of 0.75 that gives a clear picture of the anticipated order volume in the case of a pessimistic demand environment.

Scenario B: Demand Spike (+25%).

The demand-spike scenario will increase the XGBoost baseline forecast by 25 percent. Based on this assumption, the six-month projection will be 19,482,392 units, instead of 24,352,990 units, which will be an increment of 4,870,598 units. This outcome demonstrates the possible order-volume pressure that can be obtained in case of a high-demand situation. As an operation, this case would apply to material preparation, production schedule, capacity planning, and short-term planning discussions, where the increased demand would necessitate the factory to determine whether it had the capacity to meet the increased order quantity.

Scenario C: Buyer A Cancellation.

The buyer-cancellation scenario is based on the largest buyer's share of the observed order book. To protect confidentiality, this buyer is anonymised as Buyer A. Buyer A had about 29.44% of total Order Qty in the data. When this share is used as a cancellation stress factor, the six months baseline projection becomes 13,746,206 units as compared to 19,482,392 units, a difference of 5,736,186 units. The case reveals the operational importance of buyer concentration in the production of garments that are exported. The three scenarios presented jointly illustrate how the XGBoost baseline forecast can be converted into other demand conditions and H2 is supported by showing the use of scenario planning as a formal decision-support layer.

4.3.4 Comparative Evaluation: Accuracy and Operational Adaptability

The empirical results show that there is an evident performance hierarchy among the four forecasting models. The comparison of model performance in MAPE, RMSE, and MAE was over the same six-month chronological holdout period, and operational adaptability was tested by stress testing the XGBoost baseline forecast with scenarios. Table 20 demonstrates the changes in the XGBoost baseline forecast during demand-drop, demand-spike, and Buyer A cancellation conditions.

Forecast Accuracy Comparison

The multi-metric analysis indicates that XGBoost did the best in this data. XGBoost had the lowest MAPE of 26.25 against ARIMA with 40.75, Holt-Winters with 44.68 and Random Forest with 33.24. This shows that XGBoost was able to learn the monthly order trend better than the traditional statistical thresholds as well as Random Forest.

The XGBoost minimized MAPE by 14.50 percentage points, 40.75 to 26.25, compared to ARIMA. ARIMA was also outperformed by random Forest with a MAPE of 33.24. These outcomes indicate that the machine learning models offered better forecast accuracy during the holdout period with XGBoost offering the most appropriate baseline of the scenario-planning phase.

Measurement of Adaptability- The adaptability in this research is evaluated based on the scenario-planning layer and not on the re-estimation of recursive models. XGBoost forecast gives a baseline demand estimation, where the demand-drop, demand-spike, and buyer cancellation scenarios indicate how the forecasted volume of order varies in case of alternative operating conditions. This offers a viable foundation for managerial discussion of capacity alignment, smoothing of production and buyer-risk exposure. Responsiveness is explained based on the results of the scenario, indicating how the XGBoost baseline forecast is sensitive to other demand assumptions.

Forecasting-Scenario Planning Integration- The findings indicate that XGBoost has the highest baseline forecast out of the four models that were tested. Scenario planning then builds on this prediction by analyzing the changes in total order volume when there is a 25% drop in demand, a 25% demand spike, and Buyer A cancellation scenario. This integration bridges the gap between the predictive accuracy and operational interpretation whereby the forecast can facilitate production planning, capacity discussion, and buyer-risk estimation to react to the alternative demand conditions.

Table 20. Scenario Sensitivity Analysis Based on XGBoost Baseline Forecast, August 2025-January 2026.

Scenario Type	**Scenario Assumption**	**Six-Month Forecast Volume**	**Change vs Baseline**	**Operational Interpretation**
Baseline	No stress adjustment	19,482,392	0	Expected XGBoost order-volume forecast
Demand drop	-25%	14,611,794	-4,870,598	Downside demand exposure
Demand spike	+25%	24,352,990	+4,870,598	Upside volume and capacity pressure
Buyer cancellation	-29.44%	13,746,206	-5,736,186	Largest-buyer concentration exposure

4.4 Synthesis of Implications and Strategic Recommendations

The shift in classical statistical forecasting to a machine-learning baseline with assistance of scenario planning has practical consequences on forecasting governance, planning and operational risk in the garment industry in Bangladesh. It demonstrates how forecasting can shift away to become more of a historical and reactive process toward a more structured decision support process that links model results to production and buyer-risk factors.

4.4.1 Policy and Governance Implications for the Garment Industry

The empirical results show that the issue of using intuition is still one of the challenges in predicting governance in the garment sector in Bangladesh. To have a more robust data-driven forecasting culture, more regular data gathering, better defined planning practices, and institutional reinforcement by industry associations like the BGMEA are needed. The comparison of the models indicates that the XGBoost minimized the MAPE of ARIMA (40.75) to 26.25 indicating the usefulness of structured forecasting models despite having limited data in the form of company level monthly order records.

4.4.2 Operational and Design Implications for Supply Chain Management

Machine learning and scenario planning can help shift to a data-driven planning process instead of reactive production planning. The XGBoost base forecast gives an approximate of the expected volume of orders and the scenario-planning layer indicates how the volume will vary due to demand drop, demand spike and Buyer A cancellation. This provides managers with better ground upon which they can talk about material planning, smoothing of production and exposing a buyer to risk prior to the operational pressure being harder to handle.

4.4.3 Economic and Environmental Efficiency Considerations

The accuracy of the forecast and testing of the scenarios in the production planning and operational risk assessment is practically applicable as demonstrated in section 4.3.4. Within an export-based garment environment, enhanced forecasting facilitates a more accurate match between forecasted volume of orders, material forming, production schedules and capacity planning.

Economic efficiency: Better forecast accuracy may help minimize the misalignment of plans and enable more reliable production choices. The XGBoost model enables managers to predict order volume needs more effectively by giving them a more robust baseline forecast, with the scenario testing introducing a more formal perspective on the possible demand change.

Environmental sustainability: More precise forecasting can also contribute to more effective resource utilization, as more accurate forecasting helps firms produce much more precisely based on anticipated demand. The forecasting scenario planning framework, in this aspect, is not only applicable to operational efficiency but also in minimizing unnecessary overproduction and waste in the manufacturing of garments.

4.4.4 Stakeholder Perspectives and Strategic Readiness

Although ensemble models are technologically superior, the major challenge in implementing such models lies in the preparedness of humans for strategy implementation. There is usually skepticism among practitioners about "black box" models, preferring simple linear trends even though their predictive capabilities are inferior. For success, it is essential that companies close the gap between math and the art of garment production by creating "Translators" and breaking down data silos. The success of ML-Scenario architecture requires a paradigm shift to Data Driven Decision Making (DDDM).

4.5 Chapter Summary

Chapter 4 showed the empirical findings of the forecasting and the scenario-planning analysis. Descriptive analysis indicated a high volatility in the series of monthly order-quantity which supports the relevance of the proposed comparison of the traditional forecasting models and machine learning methods. The model comparison revealed that the XGBoost model performed the best with a MAPE of 26.25 percent as compared to ARIMA with 40.75 percent, Holt-Winters with 44.68 percent, and Random Forest with 33.24 percent. This finding suggests that the XGBoost model better reflected the temporal structure of the monthly order series, compared to the traditional benchmarks and the Random Forest.

The XGBoost baseline forecast was then extrapolated to the scenario-planning analysis using three deterministic stress tests, namely a 25 percent drop in demand, a 25 percent demand spike and a Buyer A cancellation scenario to the largest share of orders per buyer. These scenario results illustrate how the base forecast can be translated into alternative to demand conditions which can be utilized in production planning, capacity discourse, and buyer risk assessment. Broadly, the chapter shows that machine learning can offer a more robust forecasting base, and scenario planning can offer an operational decision-support layer to interpret demand uncertainty in the garment industry of Bangladesh.

4.6 Hypothesis Testing Results

The results of the hypotheses are based on the empirical evidence and simulation testing presented in this chapter.

H1: Machine learning based forecasting, especially XGBoost and Random Forest is more precise than traditional statistical forecasting, ARIMA and Holt-Winters, in the garment order-demand setting.

Verdict: Supported.

The model-comparison results of the holdout support H1. XGBoost recorded the lowest forecasting error amongst the four models that were tested and produced a MAPE of 26.25%. Random Forest also took the lead over the traditional benchmarks, with a MAPE of 33.24, versus 40.75 with ARIMA and 44.68 with Holt-Winters. These findings indicate that the machine learning models and in particular XGBoost were more successful in capturing the monthly order-quantity pattern as compared to the traditional statistical models.

H2: Scenario planning that uses machine learning to predict operations is more adaptable in operations than forecast-only evaluation. Decision: Decision supported

The findings confirm H2 by demonstrating the ability to extend the XGBoost baseline forecast to other operational contexts. The scenario planning layer converted the six-month baseline forecast to three stress tests: a 25% drop in demand, a 25% spike in demand, and a Buyer A cancellation scenario. This enabled the systematic basis in the consideration of the downside demand exposure, upside volume pressure and buyer-concentration risk and increased the operational utility of the forecasting output.

5 Discussion and Conclusion

5.1 Discussion of Key Findings

The primary objective of the study was to investigate the possibility of integrating machine learning with scenario planning to enhance the ability to predict demand in the unstable setting of the garment industry in Bangladesh. Chapter 4 results indicate that the integrated framework gives forecasting and decision-support value. The XGBoost had the highest forecasting performance among the four models tested, with the scenario-planning layer extrapolating the baseline forecast to structured stress-test results. These findings collectively demonstrate that machine learning can reinforce accuracy of the baseline forecast, and scenario planning can enhance the operational interpretation of the uncertainty of demand.

Comparative Performance of Ensemble Learning: The findings indicate that XGBoost was more successful in dealing with the monthly order-quantity series in comparison with the conventional statistical benchmarks. The model demonstrated a MAPE of 26.25% as opposed to 40.75% by ARIMA, 44.68% by Holt-Winters and 33.24% by Random Forest. This has been attributed to the fact that the model uses rolling demand momentum, monthly seasonality and lagged order effects. These results hence confirm the usefulness of tree-based machine learning to identify non-linear time patterns in company-level data on garment orders.

The Scenario-Planning Contribution to Resilience: A key finding is that the XGBoost base forecast can be more helpful to the managerial decision-making process when integrated with scenario planning. The scenario tests demonstrate the changes in the baseline forecast when there are three demand conditions: 25% drop in demand, 25% spike in demand and cancellation of the order exposure of the largest buyer. This changes the focus of analysis to a one-point forecast to a more decision-support perspective of downside demand risk, upside volume pressure and buyer-concentration exposure.

Scenario planning can so complement the accuracy of forecasts by providing the managers with a systematic foundation on which to formulate the other conditions of operation.

Connection with theoretical frameworks: The results are aligned with the Dynamic Capabilities View since the integrated forecasting and scenario planning framework aids in sensing and seizing demand uncertainty. XGBoost and scenario planning can be used to supplement the sensing and seizing functions, respectively, through enhancing the base order forecast and converting the outputs of the forecast into alternative operational specifications. This association implies how routine practices based on analysis might be used to facilitate more flexible planning in an export-oriented garment-manufacturing setting.

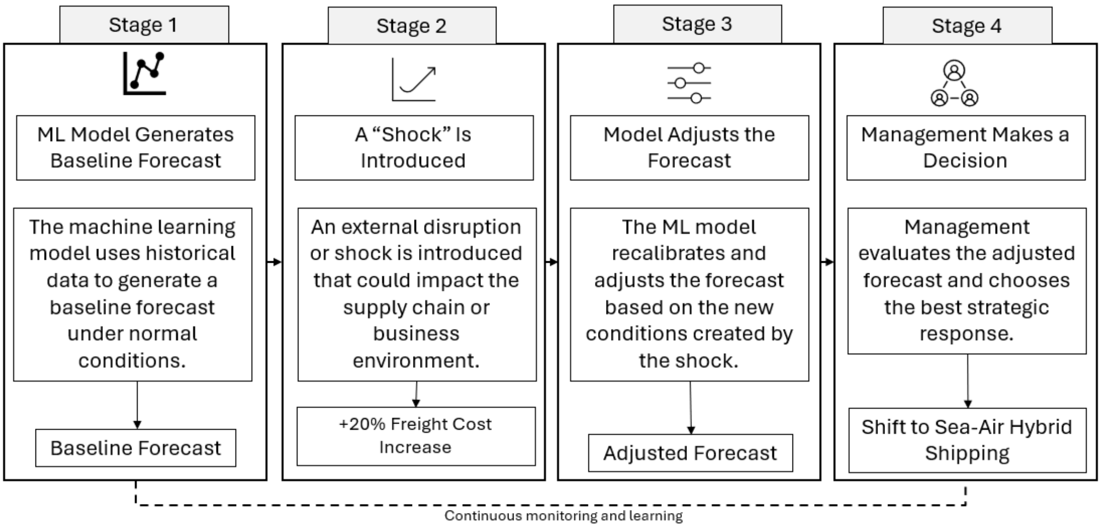

Figure 12. Proposed Integrated Framework for Data-Driven Decision-Making: Linking XGBoost Baseline Forecasts to Scenario-Based Stress Testing.

5.2 Contributions to Theory and Practice

This research paper fits the relationship between forecasting analytics and real-world industrial management within an environment of developing economy manufacturing.

It demonstrates that company-level order data can be converted into a model comparison and operational interpretation-capable forecasting and scenario-planning framework.

Theory: The theoretical contribution to the study is that it relates machine learning-based demand forecasting with a scenario-based stress test. The thesis illustrates the forecasting scenario planning format whereby the machine learning model will give the baseline forecast, and scenario planning will extrapolate the forecast to other operating conditions. This upholds the argument that forecasting cannot be seen as a mere statistical process, but as an element of a larger decision-supporting process in dynamic supply chain settings.

Combination of Systems Thinking and Machine Learning: The research also adds to the literature by aligning machine learning as an applicable computational process to explain volatile demand trends. The identification of non-linear temporal patterns of the monthly order data is performed by ensemble-based learning, specifically XGBoost, and based on the rolling demand momentum, lagged order quantities, and calendar-based variables. This is similar to Systems Theory where demand forecasting can be seen as an element of a larger operational system where order behaviour, seasonality, production planning and buyer exposure are interdependent.

Continuing the Dynamic Capabilities Perspective: The research expands the Dynamic Capabilities perspective by demonstrating how analytical routines could facilitate the sensing and seizing in demand planning. Machine learning also adds to the sensing part, where patterns in the monthly order data are found, and scenario planning to the seizing part, where the baseline forecast is converted into differing operational conditions. The framework, using a Human-in-the-Loop methodology, integrates model-based forecasting and managerial interpretation, which is significant in the garment manufacturing setting where operational decisions remain at the discretion of the planner, interaction with buyers, and production limitations.

The results also apply to the stakeholders of the Bangladeshi garment industry. In terms of digital transformation, the study proposes a replicable roadmap in shifting the expert-judgement forecasting approach to a more data-informed planning pattern. It involves consolidating internal order data, building lagged and calendar-based forecasting capability, comparing machine learning and traditional models, picking the best underlying forecast, and stress-testing using scenario-based stress testing to aid operational planning.

Economic and Operational Relevance: The analysis indicates that the XGBoost minimized the MAPE of 40.75% with ARIMA by 26.25% during the holdout test period. The operationally relevant improvement of this is that more accurate predictions will be able to assist in better production planning, preparation of materials and capacity discussions. A more robust baseline forecast in the garment export environment where the changes in order may impact on the purchasing of fabrics and the line planning and delivery obligations offers a more accurate foundation to coordinate the planning decisions.

Sustainability by Accuracy: The suggested approach can also aid in the more efficient use of resources by enabling companies to coordinate their forecasted demand for orders with their production plans. Greater precision in the forecasting would minimize unnecessary planning errors and aid in enhancing the coordination of the demand expectations and volume of production. In this regard, the forecasting-scenario planning model can be potentially relevant to sustainability-based planning, especially when excess production and inefficient utilization of resources are issues of operational concern.

5.3 Limitations of the Study

Although the model comparison results are useful, a number of limitations must be taken into consideration when explaining the results.

Data Size and Technical Limits: The research is founded on a 37-month longitudinal dataset, which offers an adequate foundation of comparing the traditional forecasting models with the tree-based machine learning models. Nevertheless, the data size is still relatively small to use more data and intensive methods like LSTM or Transformer based models. This is why the analysis is dedicated to ARIMA, Holt-Winters, Random Forest, and XGBoost which are more suitable to the data structure available. The forecasting model is also estimated at the company-level monthly Order Qty level as opposed to SKU or product-category level. This enhances the persistence of the time series, but prevents the generation of product-specific predictions.

Scope and Generalisability: The case study is concerned with one medium sized export firm in Dhaka. As such, the results should be applied in terms of the case company and not the entire Bangladesh garment industry. The ranking of features by importance, especially the importance of rolling demand momentum and monthly seasonality, might differ between firms with diverse buyer demands, product build, data quality, and planning parameters.

Model Scope: The empirical model is based on the company order data temporal and calendar features. More industry-wide influences like freight prices, buyer-market inflation, exchange rates and logistics delays are also applicable to garment demand, but are handled under the industry context, literature discussion and scenario-planning interpretation. The model may be further developed in future research by adding regular external and operational indicators in case these data are known.

Scenario-Planning Scope: The scenario-planning component is based on deterministic stress tests. The 25% demand-drop scenario, 25% demand-spike scenario and Buyer A cancellation scenario demonstrate how the baseline forecast varies given different demand conditions. These are some situations that give useful exposure estimates to plan discussion and future research might create probabilistic scenario models or cost-based simulations to determine the probability and monetary impact of such events. By doing that a more realistic scenario will be achievable to get more accurate plan to handle the effects of different shocks.

5.4 Recommendations and Future Research

Future Research directions

Even though in the current research it is demonstrated that XGBoost performed better than the chosen statistical benchmarks in this dataset, the following methodological directions should be considered in future studies:

Development of Hybrid Statistical-Machine Learning Models: Future research needs to explore the formalisation of ARIMA-XGBoost or ARIMA-ANN hybrids. The primary research questions would be: How can residual-based hybrid modelling be used to predict better than standalone ensemble methods in the intermittency, low-volume, SKU categories? This type of research would need more intensive and detailed product level data, which was not available in the current research.

Inclusion of External and Real-Time Data: There was no external variable that was taken as a direct predictor in this study, and the internal order monthly data were incorporated. Exogenous regressors that can be included in future studies are buyer-market indicators, freight cost indices, exchange rates, macroeconomic variables and retail sentiment indicators. This would enable a more rigorous empirical examination of the influence of external market forces on Bangladesh garment demand prior to its showing up in order records.

Scalability to Small and Medium Enterprises (SMEs): As in this study it is a single firm, it is recommended that a multi-firm comparative study be used. Research should consider how the ranking of feature importance varies when the forecasting model is used on SMEs with low bargaining power and high exposure to cancellations by buyers.

5.5 Conclusion

This paper has demonstrated that the use of XGBoost forecasting and scenario-based stress testing can be useful in providing a decision-support structure in demand forecasting in the garment industry of Bangladesh. XGBoost performed the best in the empirical analysis and had the lowest MAPE of 26.25% when compared to the four tested models. This is compared to 40.75% of ARIMA, 44.68% of Holt-Winters and 33.24% of Random Forest. The finding favors the thesis that tree-based machine learning can boost the precision of monthly order forecasts within a company-level monthly order sample, especially when forecasting variables are recent demand momentum, lagging order quantities, and calendar-based predictors. The scenario analysis also indicates how the XGBoost baseline forecast may be transformed into the operational stress-test results. A 25 percent demand spike raised the six-month baseline forecast of 19,482,392 units to 24,352,990 units and a 25 percent demand drop lowered it to 14,611,794 units. The Buyer A cancellation scenario decreased the forecast base of 13,746,206 units, which indicates the magnitude of buyer-concentration risk in the order book of the company in question. The result supports the hypothesis that tree-based machine learning can enhance the accuracy of monthly order predictions in a monthly order sample at the company-level, particularly in predicting variables that are recent demand momentum, lagging order quantities, and calendar-based predictors. The thesis finds that machine learning will enhance the accuracy of the baseline forecast, and scenario planning enhances the managerial interpretation of uncertainty. The integrated framework thus helps with a more organized way of planning demand in the Bangladesh garment industry. If data standardisation, forecasting practices, and human-in-the-loop control are enhanced throughout the value chain, garment manufacturers will be able to achieve more plausible steps toward data-driven planning, resilience in operations, and more economical resource utilisation.

References

Agathangelou, P., Trihinas, D., & Katakis, I. (2020). A Multi-Factor Analysis of Forecasting Methods: A Study on the M4 Competition. Data (Basel), 5(2), 41. https://doi.org/10.3390/data5020041

Alam, Sadman & Alam, Md. Shafiul & Uzzaman, Md & Rahman, Mohammad. (2016). A Research on Supply Chain Network of RMG Sector of Bangladesh. International Journal of Scientific and Engineering Research. 7. 2. 10.14299/ijser.2016.10.004.

Ameen, Sidra & Aslam, Muhammad & Ahmad, Dr & Ali, Imad. (2025). Enhancing Supply Chain Resilience in Developing Country SMEs through Machine Learning-Based Demand Forecasting: An Empirical Study. The Critical Review of Social Sciences Studies. 3. 3050-3088. 10.59075/tr48sj71.

Amer, M., Daim, T. U., & Jetter, A. (2013). A review of scenario planning. Futures : the journal of policy, planning and futures studies, 46, 23-40. https://doi.org/10.1016/j.futures.2012.10.003

Arend, R. J., & Bromiley, P. (2009). Assessing the dynamic capabilities view: Spare change, everyone? Strategic organization, 7(1), 75-90. https://doi.org/10.1177/1476127008100132

Armstrong, J. S. (2001). Principles of forecasting: A handbook for researchers and practitioners. Boston, MA: Kluwer Academic. https://doi.org/10.1007/978-0-306-47630-3

Arora, R., Sachs, A., Svetunkov, I., Muth, M., & Boylan, J. E. (2025). Forecasting demand during and after supply chain disruptions using a shock smoother ETS. International journal of production research, 1-27. https://doi.org/10.1080/00207543.2025.2577158

Atesongun, A., & Gulsen, M. (2024). A Hybrid Forecasting Structure Based on Arima and Artificial Neural Network Models. Applied sciences, 14(16), 7122. https://doi.org/10.3390/app14167122

Badakhshan, E., Mustafee, N., & Bahadori, R. (2024). Application of simulation and machine learning in supply chain management: A synthesis of the literature using the Sim-ML literature classification framework. Computers & industrial engineering, 198, 110649. https://doi.org/10.1016/j.cie.2024.110649

Bengio, Y., Lodi, A., & Prouvost, A. (2021). Machine learning for combinatorial optimization: A methodological tour d'horizon. European journal of operational research, 290(2), 405-421. https://doi.org/10.1016/j.ejor.2020.07.063

Bishop, P., Hines, A., & Collins, T. (2007). The current state of scenario development: An overview of techniques. Foresight (Cambridge), 9(1), 5-25. https://doi.org/10.1108/14636680710727516

Bruzzone, A. G., Massei, M., & Frosolini, M. (2022). Redesign of Supply Chain in Fashion Industry based on Strategic Engineering. Procedia computer science, 200, 1913-1918. https://doi.org/10.1016/j.procs.2022.01.392

Brynjolfsson, Erik & Mcelheran, Kristina. (2016). Data in Action: Data-Driven Decision Making in U.S. Manufacturing. SSRN Electronic Journal. 10.2139/ssrn.2722502.

Castañeda-Navarrete, J., Hauge, J., & López-Gómez, C. (2021). COVID-19's impacts on global value chains, as seen in the apparel industry. Development policy review, 39(6), 953-970. https://doi.org/10.1111/dpr.12539

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. New York, NY, USA: ACM. https://doi.org/10.1145/2939672.2939785

Chopra, S., & Meindl, P. (2016). Supply chain management: Strategy, planning, and operation (6th ed., Global ed.). Pearson.

Chowdhury, M. M. H., & Quaddus, M. (2017). Supply chain resilience: Conceptualization and scale development using dynamic capability theory. International journal of production economics, 188, 185-204. https://doi.org/10.1016/j.ijpe.2017.03.020

Chowdhury, P., Paul, S. K., Kaisar, S., & Moktadir, M. A. (2021). COVID-19 pandemic related supply chain studies: A systematic review. Transportation research. Part E, Logistics and transportation review, 148, 102271. https://doi.org/10.1016/j.tre.2021.102271

Christopher, M., & Peck, H. (2004). Building the Resilient Supply Chain. The international journal of logistics management, 15(2), 1-14. https://doi.org/10.1108/09574090410700275

Cooper, P. E., Walker, C. M., & Babb, L. C. (2025). Explainable Demand Forecasting for Retail Inventory Systems Under Supply Disruptions. International Journal Management and Organizational Research, 4(6), 121-125. https://doi.org/10.54660/IJMOR.2025.4.6.121-125

Cordova-Pozo, K., & Rouwette, E. A. (2023). Types of scenario planning and their effectiveness: A review of reviews. Futures : the journal of policy, planning and futures studies, 149, 103153. https://doi.org/10.1016/j.futures.2023.103153

Csaszar, F. A., Ketkar, H., & Kim, H. (2024). Artificial Intelligence and Strategic Decision-Making: Evidence from Entrepreneurs and Investors. Strategy science, 9(4), 322-345. https://doi.org/10.1287/stsc.2024.0190

De Angelis, R., Howard, M., & Miemczyk, J. (2018). Supply chain management and the circular economy: Towards the circular supply chain. Production planning & control, 29(6), 425-437. https://doi.org/10.1080/09537287.2018.1449244

Derbyshire, J., & Wright, G. (2017). Augmenting the intuitive logics scenario planning method for a more comprehensive analysis of causation. International journal of forecasting, 33(1), 254-266. https://doi.org/10.1016/j.ijforecast.2016.01.004

Douaioui, K., Oucheikh, R., Benmoussa, O., & Mabrouki, C. (2024). Machine Learning and Deep Learning Models for Demand Forecasting in Supply Chain Management: A Critical Review. Applied system innovation, 7(5), 93. https://doi.org/10.3390/asi7050093

Gaertner, T., Lippert, C., & Konigorski, S. (2024). Automated Demand Forecasting in small to medium-sized enterprises. https://doi.org/10.48550/arxiv.2412.20420

Goel, L., Nandal, N., Gupta, S., Karanam, M., Prasanna Yeluri, L., Pandey, A. K., . . . Grabovy, P. (2024). Revealing the dynamics of demand forecasting in supply chain management: A holistic investigation. Cogent engineering, 11(1), . https://doi.org/10.1080/23311916.2024.2368104

Hobor, L., Brcic, M., Polutnik, L., & Kapetanovic, A. (2025). Comparative Analysis of Modern Machine Learning Models for Retail Sales Forecasting. https://doi.org/10.48550/arxiv.2506.05941

Hosseini, S., Ivanov, D., & Dolgui, A. (2019). Review of quantitative methods for supply chain resilience analysis. Transportation research. Part E, Logistics and transportation review, 125, 285-307. https://doi.org/10.1016/j.tre.2019.03.001

Ivanov, D. (2021). Supply Chain Viability and the COVID-19 pandemic: A conceptual and formal generalisation of four major adaptation strategies. International journal of production research, 59(12), 3535-3552. https://doi.org/10.1080/00207543.2021.1890852

Ivanov, D. (2022). Viable supply chain model: Integrating agility, resilience and sustainability perspectives—lessons from and thinking beyond the COVID-19 pandemic. Annals of operations research, 319(1), 1411-1431. https://doi.org/10.1007/s10479-020-03640-6

Ivanov, D., & Dolgui, A. (2022). Stress testing supply chains and creating viable ecosystems. Operations management research, 15(1-2), 475-486. https://doi.org/10.1007/s12063-021-00194-z

Kalaitzi, D., & Tsolakis, N. (2022). Supply chain analytics adoption: Determinants and impacts on organisational performance and competitive advantage. International journal of production economics, 248, 108466. https://doi.org/10.1016/j.ijpe.2022.108466

Klibi, W., & Martel, A. (2012). Scenario-based Supply Chain Network risk modeling. European journal of operational research, 223(3), 644-658. https://doi.org/10.1016/j.ejor.2012.06.027

Kontopoulou, V. I., Panagopoulos, A. D., Kakkos, I., & Matsopoulos, G. K. (2023). A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks. Future internet, 15(8), 255. https://doi.org/10.3390/fi15080255

Ma, B. J., Jackson, I., Huang, M., Villegas, S., & Macias-Aguayo, J. (2025). A data-driven and context-aware approach for demand forecasting in the beverage industry.

International journal of logistics, 1-28. https://doi.org/10.1080/13675567.2025.2566806

Makridakis, S., Fry, C., Petropoulos, F., & Spiliotis, E. (2022). The Future of Forecasting Competitions: Design Attributes and Principles. INFORMS journal on data science, 1(1), 96-113. https://doi.org/10.1287/ijds.2021.0003

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). The M4 Competition: 100,000 time series and 61 forecasting methods. International journal of forecasting, 36(1), 54-74. https://doi.org/10.1016/j.ijforecast.2019.04.014

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2022). M5 accuracy competition: Results, findings, and conclusions. International journal of forecasting, 38(4), 1346-1364. https://doi.org/10.1016/j.ijforecast.2021.11.013

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2022). The M5 competition: Background, organization, and implementation. International journal of forecasting, 38(4), 1325-1336. https://doi.org/10.1016/j.ijforecast.2021.07.007

Mansur, S., Sattar, K., Hosseini, S. E., Pervez, S., Ahmad, I., Saleem, K., & Zohier Elhendi, A. (2025). Sales forecasting for retail stores using hybrid neural networks and sales-affecting variables. PeerJ. Computer science, 11, e3058. https://doi.org/10.7717/peerj-cs.3058

Milewska, B. (2022). The Impact of the COVID-19 Pandemic on Supply Chains in the Example of Polish Clothing Companies in the Context of Sustainable Development. Sustainability, 14(3), 1899. https://doi.org/10.3390/su14031899

Mondal, Md & Uddin, Mohammad & Akter, Nazma. (2025). The Textile Industry in Bangladesh: Growth Trends, Challenges, and Future Prospects. Global Disclosure of Economics and Business. 13. 73-86. 10.18034/gdeb.v13i2.789.

Munir, M. A., Abbas, M. M., Qamar, A., Farooq, M., & Cohen, V. (2025). Managing volatile markets: A dynamic capability approach to analytics-driven performance in fast-moving supply chains for sustainable development. Sustainable futures, 10, 101338. https://doi.org/10.1016/j.sftr.2025.101338

Pettit, T. J., Fiksel, J., & Croxton, K. L. (2010). ENSURING SUPPLY CHAIN RESILIENCE: DEVELOPMENT OF A CONCEPTUAL FRAMEWORK. Journal of business logistics, 31(1), 1-21. https://doi.org/10.1002/j.2158-1592.2010.tb00125.x

Ramirez, R., Mukherjee, M., Vezzoli, S., & Kramer, A. M. (2015). Scenarios as a scholarly methodology to produce "interesting research". Futures : the journal of policy, planning and futures studies, 71, 70-87. https://doi.org/10.1016/j.futures.2015.06.006

Saif, S., & Islam, A. M. (2025). Growth, development and selected social sustainability challenges facing the Bangladesh export garment industry. Businesses, 5(1), 1-22. https://doi.org/10.3390/businesses5010015

Saunders, M. N. K., Lewis, P., & Thornhill, A. (2009). Research methods for business students (5th ed.). Prentice Hall.

Schoemaker, P. J. (1995). Scenario planning: A tool for strategic thinking. Sloan management review, 36(2), 25

Shuvo, Tamim Forhad & Habib, Md & Raisa, Rodoshi. (2025). Innovative Technologies in Supply Chain Management for Bangladesh's Readymade Garments Sector -A Comprehensive Review. Social Politics. 1. 1-9. 10.69971/ek445y02.

Siddiqi, D. M. (2025, March 26). What's happening in Bangladesh's garment industry? Economics Observatory. https://www.economicsobservatory.com/whats-happening-in-bangladeshs-garment-industry

Sidorkin, O., & Srholec, M. (2014). Surviving the times of crisis: Does innovation make a difference? International journal of technological learning, innovation and development, 7(2), 124-146. https://doi.org/10.1504/IJTLID.2014.065881

Swaminathan, K., & Venkitasubramony, R. (2024). Demand forecasting for fashion products: A systematic review. International journal of forecasting, 40(1), 247-267. https://doi.org/10.1016/j.ijforecast.2023.02.005

Syntetos, A. A., Zied Babai, M., & Gardner, E. S. (2015). Forecasting intermittent inventory demands: Simple parametric methods vs. bootstrapping. Journal of business research, 68(8), 1746-1752. https://doi.org/10.1016/j.jbusres.2015.03.034

Teece, D.J. (2023). The Evolution of the Dynamic Capabilities Framework. In: Adams, R., Grichnik, D., Pundziene, A., Volkmann, C. (eds) Artificiality and Sustainability in Entrepreneurship. FGF Studies in Small Business and Entrepreneurship. Springer, Cham. https://doi.org/10.1007/978-3-031-11371-0_6

Thippur Manjunath, S. (2025). Demand forecast optimization using machine learning algorithms [Master's thesis, University of Skovde]. DiVA Portal. https://www.diva-portal.org/smash/record.jsf?pid=diva2:2004158

Veiga, C. P. d., Veiga, C. R. P. d., Puchalski, W., Coelho, L. d. S., & Tortato, U. (2016). Demand forecasting based on natural computing approaches applied to the foodstuff retail segment. Journal of retailing and consumer services, 31, 174-181. https://doi.org/10.1016/j.jretconser.2016.03.008

von Bertalanffy, L. (1968). General system theory: Foundations, development, applications. George Brazille

von der Gracht, H. A., & Darkow, I. (2010). Scenarios for the logistics services industry: A Delphi-based analysis for 2025. International journal of production economics, 127(1), 46-59. https://doi.org/10.1016/j.ijpe.2010.04.013

Wack, P. (1985). Scenarios: Uncharted waters ahead. Harvard business review, 63(5), 73-89.

Wren, B. (2022). Sustainable supply chain management in the fast fashion Industry: A comparative study of current efforts and best practices to address the climate crisis. Cleaner Logistics and Supply Chain, 4, 100032. https://doi.org/10.1016/j.clscn.2022.100032

Yasir, M., Ansari, Y., Latif, K., Maqsood, H., Habib, A., Moon, J., & Rho, S. (2024). Machine learning-assisted efficient demand forecasting using endogenous and exogenous indicators for the textile industry. International journal of logistics, 27(12), 2867-2886. https://doi.org/10.1080/13675567.2022.2100334

Zahraei, S. M., & Teo, C. (2017). Optimizing a supply network with production smoothing, freight expediting and safety stocks: An analysis of tactical trade-offs. European journal of operational research, 262(1), 75-88. https://doi.org/10.1016/j.ejor.2017.02.045

Zhang, G. (2003). Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing (Amsterdam), 50, 159-175. https://doi.org/10.1016/S0925-2312(01)00702-0

Appendix: Python Model Execution Summary

This appendix summarises the corrected Python execution used for empirical analysis. The analysis was carried out using the final Google Colab-compatible workflow based on the company order dataset. The target variable was monthly Order Qty aggregated by Latest Ex-Fty Date. The training period covered January 2023 to July 2025, and the holdout test period covered August 2025 to January 2026. The older appendix entries referring to log-loss, early stopping, and XGBoost classification-style convergence were removed because the final model is a regression model.

A.1 Dataset and Preprocessing Scope

Total observations: 37 monthly observations from January 2023 to January 2026.

Target variable: monthly Order Qty aggregated by Latest Ex-Fty Date.

Training period: January 2023 to July 2025.

Test period: August 2025 to January 2026.

Input features for machine learning: Lag-1, Lag-2, Lag-3, rolling three-month average, month, and quarter.

Buyer names were anonymised in the scenario-planning interpretation.

A.2 Descriptive Statistics Output

Mean monthly order quantity: 3,154,136 units.

Median monthly order quantity: 2,836,902 units.

Standard deviation: 1,172,802 units.

Coefficient of variation: 0.372.

Minimum monthly order quantity: 1,139,760 units.

Maximum monthly order quantity: 5,821,660 units.

A.3 Stationarity and Residual Diagnostics

ADF test on training level series: statistic = -4.1926; p-value = 0.0007.

ADF test on first-differenced training series: statistic = -3.0194; p-value = 0.0331.

Ljung-Box test on ARIMA residuals at lag 10: statistic = 9.4084; p-value = 0.4938.

The Ljung-Box result indicates no statistically significant residual autocorrelation at lag 10.

A.4 Forecasting Model Performance

ARIMA(1,1,1): MAPE = 40.75%; RMSE = 1,189,013; MAE = 1,129,948.

Holt-Winters: MAPE = 44.68%; RMSE = 1,360,989; MAE = 1,168,825.

Random Forest: MAPE = 33.24%; RMSE = 1,143,463; MAE = 1,024,590.

XGBoost: MAPE = 26.25%; RMSE = 850,095; MAE = 765,816.

XGBoost achieved the lowest forecast error among the four tested models in the holdout period.

A.5 XGBoost Feature Importance Output

Rolling three-month average: 0.490.

Month: 0.222.

Lag-2: 0.150.

Lag-3: 0.098.

Lag-1: 0.039.

Quarter: 0.000.

The feature-importance output confirms that recent order momentum and monthly seasonality were the dominant predictors in the final empirical model.

A.6 Scenario Planning Output

Actual six-month order quantity in the holdout period: 19,351,214 units.

Baseline XGBoost six-month forecast: 19,482,392 units.

Demand-drop scenario (-25%): 14,611,794 units; change versus baseline = -4,870,598 units.

Demand-spike scenario (+25%): 24,352,990 units; change versus baseline = +4,870,598 units.

Buyer A cancellation scenario (-29.44%): 13,746,206 units; change versus baseline = -5,736,186 units. The scenario results were used as deterministic decision-support stress tests and were not interpreted as probabilistic forecasts.

A.7 Software Environment

Python was used for all empirical calculations.

Primary libraries: pandas, numpy, statsmodels, scikit-learn, XGBoost, and matplotlib.

The final workflow generated descriptive statistics, model-performance tables, forecast-comparison figures, feature-importance output, and scenario-planning outputs.

All file paths were environment-specific and are therefore not reported in the thesis text.