applied sciences

Article

Sustainable Cold Chain Management: An Evaluation of Predictive Waste Management Models

Hajar Fatorachian [1,*] and Kulwant Pawar [2]

[1] Leeds Business School, Leeds Beckett University, Leeds LS1 3HB, UK
[2] Nottingham University Business School, University of Nottingham, Nottingham NG8 1BB, UK; kul.pawar@nottingham.ac.uk
* Correspondence: h.fatorachian@leedsbeckett.ac.uk

Abstract: The integration of advanced predictive models is pivotal for optimizing demand forecasting and inventory management in cold chain logistics. This study evaluates the application of machine learning techniques—ARIMA (Auto-Regressive Integrated Moving Average) and Multiple Linear Regression (MLR)—to forecast demand trends and analyze key drivers in a mid-sized cold chain operation. Trained on a multi-year sales dataset, the ARIMA model excelled in capturing seasonal patterns, while the MLR model effectively incorporated multivariable factors such as temperature, product type, and promotional activity. Both models demonstrated strong predictive accuracy, with low Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), offering reliable and computationally efficient solutions for mid-sized operations. The findings underscore the novelty of combining ARIMA's time-series capabilities with MLR's multivariable analysis to address complex demand drivers. By aligning with Resource-Based View (RBV) and Supply Chain Resilience Theory, this research advances the understanding of AI-driven predictive models as strategic tools for enhancing operational efficiency, reducing waste, and promoting sustainability in cold chain logistics. This work sets the stage for future innovations in AI-driven supply chain optimization.

Keywords: cold chain logistics; Artificial Intelligence (AI); demand forecasting; sustainable waste management; Internet of Things (IoT) integration

Academic Editors: Giorgia Casella, Letizia Tebaldi and Eleonora Bottani

Received: 17 November 2024
Revised: 24 December 2024
Accepted: 27 December 2024
Published: 14 January 2025

Citation: Fatorachian, H.; Pawar, K. Sustainable Cold Chain Management: An Evaluation of Predictive Waste Management Models. *Appl. Sci.* **2025**, *15*, 770. https://doi.org/10.3390/app15020770

1. Introduction

Cold chain systems are critical for preserving the quality and safety of perishable goods, particularly in the food and pharmaceutical industries. By controlling temperature throughout the storage and distribution processes, cold chains prevent spoilage and maintain product integrity. However, these systems face significant challenges in sustainability due to waste generated from inventory mismanagement, temperature abuses, and inaccurate demand forecasting [1]. In the food industry alone, cold chain inefficiencies contribute to approximately 15% of food loss globally, which has serious environmental, economic, and public health impacts [2].

Recent technological advancements, particularly in Artificial Intelligence (AI) and the Internet of Things (IoT), offer promising solutions for overcoming these challenges. IoT devices, like temperature and humidity sensors, enable real-time monitoring, offering logistical managers unprecedented insight into environmental conditions that can affect product quality [3]. These devices also produce vast amounts of data that, when analyzed through AI models, can enable predictive capabilities essential for inventory optimization and waste reduction. Predictive models, such as ARIMA for time-series analysis and

Multiple Linear Regression (MLR) for multivariable influence assessment, are proving instrumental in cold chain logistics by forecasting demand, optimizing stock levels, and reducing spoilage risk [4]. Industry 4.0 technologies, including AI and IoT, have the potential to address cold chain inefficiencies by enhancing operational and waste management processes, particularly in high-stakes environments like food logistics.

Despite their benefits, cold chain logistics generate approximately 2.5 billion metric tons of CO_2 emissions annually, exacerbating environmental concerns [2]. Furthermore, waste generated from overproduction and spoilage, much of it preventable, leads to significant financial losses and intensifies the environmental footprint of cold chains. For instance, it is estimated that temperature fluctuations alone, often due to gaps in monitoring and response, can result in substantial losses in both the food and pharmaceutical sectors [5]. Predictive analytics, when paired with real-time IoT data, has the potential to address these issues by dynamically aligning inventory with demand and mitigating waste.

This study explores the application of AI models to enhance sustainability in cold chains by reducing waste and emissions. Employing predictive models in conjunction with IoT data, this research aims to bridge the gap between real-time monitoring and proactive waste management, providing a scalable solution for the industry. The use of AI for demand forecasting and real-time inventory adjustments represents a proactive approach to addressing waste, benefiting not only logistics operations but also aligning with global sustainability goals [6].

While the literature has extensively explored the benefits of IoT and AI independently in cold chains, few studies address their combined application for predictive waste management. IoT-based solutions are well-documented in enabling environmental monitoring; however, there remains a limited exploration of how IoT data can be fully integrated with AI-driven predictive models to address waste from an end-to-end perspective [7]. Specifically, gaps exist in the literature on the use of ARIMA and MLR models to optimize waste reduction by forecasting demand and enhancing inventory control [8]. Maheshwari et al. [4] emphasize the importance of a comprehensive AI-IoT framework for sustainable logistics, noting that most studies either lack real-time data integration or fail to utilize the full potential of predictive models.

This study aims to address these gaps by implementing ARIMA and MLR models and examining their impact on cold chain logistics in the presence of IoT data limitations. Through a case study of Company A, this research evaluates how machine learning can reduce waste and enhance efficiency, contributing to the field of sustainable cold chain management by highlighting the operational and environmental benefits of integrating IoT and AI.

The main aim of this study is to assess the potential of AI-driven predictive waste management models within cold chain logistics, with a particular focus on demand forecasting and inventory optimization. The study's objectives are as follows:

1. Assess the availability and quality of IoT sensor data, including environmental metrics like temperature and humidity, to support AI-based waste management solutions.
2. Identify primary waste drivers in cold chains, using historical data trends to pinpoint factors like inventory misalignment and temperature inconsistencies.
3. Develop and evaluate ARIMA and MLR models, leveraging historical and environmental data to enhance demand forecasting accuracy and mitigate waste.

The paper begins with a literature review detailing current knowledge of AI and IoT applications in cold chain management. Following this, the methodology section outlines the CRISP-DM framework and describes the data collection and model selection process. The Section 5 present model performance outcomes and practical implications, highlighting the role of IoT data in predictive accuracy. The paper concludes by summarizing key insights, discussing limitations, and offering recommendations for future research.

2. Literature Review

2.1. Leveraging AI, IoT, and Predictive Models in Cold Chain Logistics

This literature review examines the integration of Artificial Intelligence (AI), the Internet of Things (IoT), and predictive modeling techniques in cold chain logistics, emphasizing their roles in demand forecasting, inventory management, waste reduction, and sustainable operations. The Resource-Based View (RBV) and Supply Chain Resilience Theory provide a theoretical foundation, contextualizing these technologies as strategic resources that enhance efficiency, adaptability, and resilience in cold chain systems.

AI and machine learning (ML) applications have demonstrated significant value in optimizing cold chain logistics by forecasting demand, managing inventory, and ensuring product quality. Maheshwari et al. [4] highlighted how AI-driven tools reduced spoilage by predicting high-risk temperature fluctuations during peak summer months in food logistics. Predictive analytics, powered by ML algorithms like ARIMA and Multiple Linear Regression (MLR), enables dynamic stock level adjustments, mitigating overstocking and reducing waste [7,8]. For instance, Hyndman and Athanasopoulos [9] demonstrated that ARIMA's time-series capabilities helped a dairy company reduce spoilage by 10% through effective seasonal demand forecasting. MLR complements ARIMA by incorporating multiple variables, such as weather conditions and consumer behavior, to provide a comprehensive approach to demand forecasting [7].

The Internet of Things (IoT) is transforming cold chain logistics through real-time monitoring of environmental parameters, including temperature, humidity, and CO_2 levels. IoT sensors facilitate immediate responses to deviations in storage conditions, reducing spoilage risks and maintaining product quality [3]. Continuous monitoring provides end-to-end visibility, enabling data-driven decisions that protect product integrity and improve operational efficiency. When integrated with AI-driven predictive models, IoT data further enhances decision-making by enabling proactive responses to quality risks. Zhu et al. [1] reported that coupling IoT with AI significantly reduced waste in cold chain operations by aligning inventory levels with real-time demand fluctuations.

Machine learning techniques like ARIMA and MLR also play a critical role in inventory management and waste reduction. ARIMA is particularly effective for analyzing cyclic demand patterns, which are common in cold chains, while MLR accounts for multiple demand drivers, such as environmental and economic factors, to provide nuanced predictions [9]. Shmueli and Koppius [10] found that combining these models reduced waste in a multi-regional food distribution network by aligning inventory levels with precise demand forecasts. These findings underscore the importance of integrating ML and IoT technologies to achieve sustainable, efficient, and resilient cold chain systems.

2.2. Theoretical Frameworks: Resource-Based View (RBV) and Supply Chain Resilience Theory

The integration of IoT and AI in cold chain logistics is strongly supported by the Resource-Based View (RBV) and Supply Chain Resilience Theory. These frameworks underscore the strategic importance of advanced technologies in maintaining competitive advantage and adapting to demand fluctuations.

- Resource-Based View (RBV)

The RBV posits that valuable, rare, inimitable, and non-substitutable (VRIN) resources provide sustainable competitive advantage [11]. IoT and AI align with RBV by acting as unique resources that enhance operational efficiency, reduce waste, and improve sustainability. Deploying these tools strategically enables firms to respond more effectively to demand changes, thereby minimizing waste. Zhu et al. [1] argue that companies that invest in AI and IoT within their cold chain operations position themselves competitively by reducing operational costs and environmental impact, supporting the RBV framework.

- Supply Chain Resilience Theory

Supply Chain Resilience Theory highlights the importance of flexibility and responsiveness in the face of disruptions. Cold chains, which are highly susceptible to fluctuations in demand and environmental conditions, benefit from resilient systems that can swiftly adapt to prevent spoilage and waste [12]. By enabling predictive adjustments, AI strengthens resilience in cold chains, while IoT sensors provide real-time insights that inform proactive decision-making [3]. Chen et al. [6] observed that IoT-based monitoring allowed one logistics firm to prevent waste in 20% of shipments by rapidly addressing disruptions, demonstrating resilience. The behavioral aspects of supply chain disruptions during the COVID-19 pandemic reinforce the critical role of adaptability, particularly within resilience frameworks [13].

2.3. Existing Studies on AI and IoT-Driven Waste Reduction in Cold Chains

The integration of AI and IoT technologies in cold chain logistics has demonstrated significant quantitative benefits. Liu et al. [5] found that incorporating IoT-generated environmental metrics into predictive models improved forecasting accuracy by 15–20%, which in turn enhanced inventory alignment and reduced waste. Similarly, Maheshwari et al. [4] highlighted that machine learning applications, including predictive models, improved supply chain agility and reduced inefficiencies in cold chain logistics. Furthermore, Zhu et al. [1] reported that optimizing cold chain logistics with IoT and AI frameworks reduced operational inefficiencies by 18%, leading to substantial waste reduction and improvements in on-time delivery rates. These findings underscore the measurable impact of predictive models on improving cold chain logistics, demonstrating that the integration of AI and IoT technologies not only enhances operational performance but also aligns with sustainability goals by minimizing waste and emissions.

2.4. Theoretical Framework

This study develops a theoretical framework for AI and IoT integration in cold chain logistics by combining the Resource-Based View (RBV) and Supply Chain Resilience Theory. RBV focuses on the strategic advantage offered by valuable, rare, inimitable, and non-substitutable (VRIN) resources [11], framing AI and IoT as strategic assets that enhance forecasting accuracy and operational efficiency. In cold chain logistics, these technologies meet VRIN criteria through advanced analytics and real-time monitoring capabilities, making them difficult for competitors to replicate [4].

Supply Chain Resilience Theory, on the other hand, highlights the importance of adaptability and responsiveness in dynamic environments [12]. AI and IoT support resilience by enabling real-time responses to environmental fluctuations and demand shifts, thereby reducing waste and mitigating spoilage risks. Together, these technologies strengthen a cold chain's ability to handle disruptions, aligning with resilience theory's focus on responsiveness and adaptability.

Figure 1 visually represents this integrated framework, illustrating how AI and IoT serve as both strategic resources (RBV) and resilience enablers (resilience theory), thus achieving dual goals of efficiency and adaptability in cold chain operations.

Left Section (RBV): Highlights the VRIN attributes of AI and IoT, emphasizing their role in creating a Strategic & Competitive Advantage by reducing waste and enhancing efficiency in a sustainable logistics system.

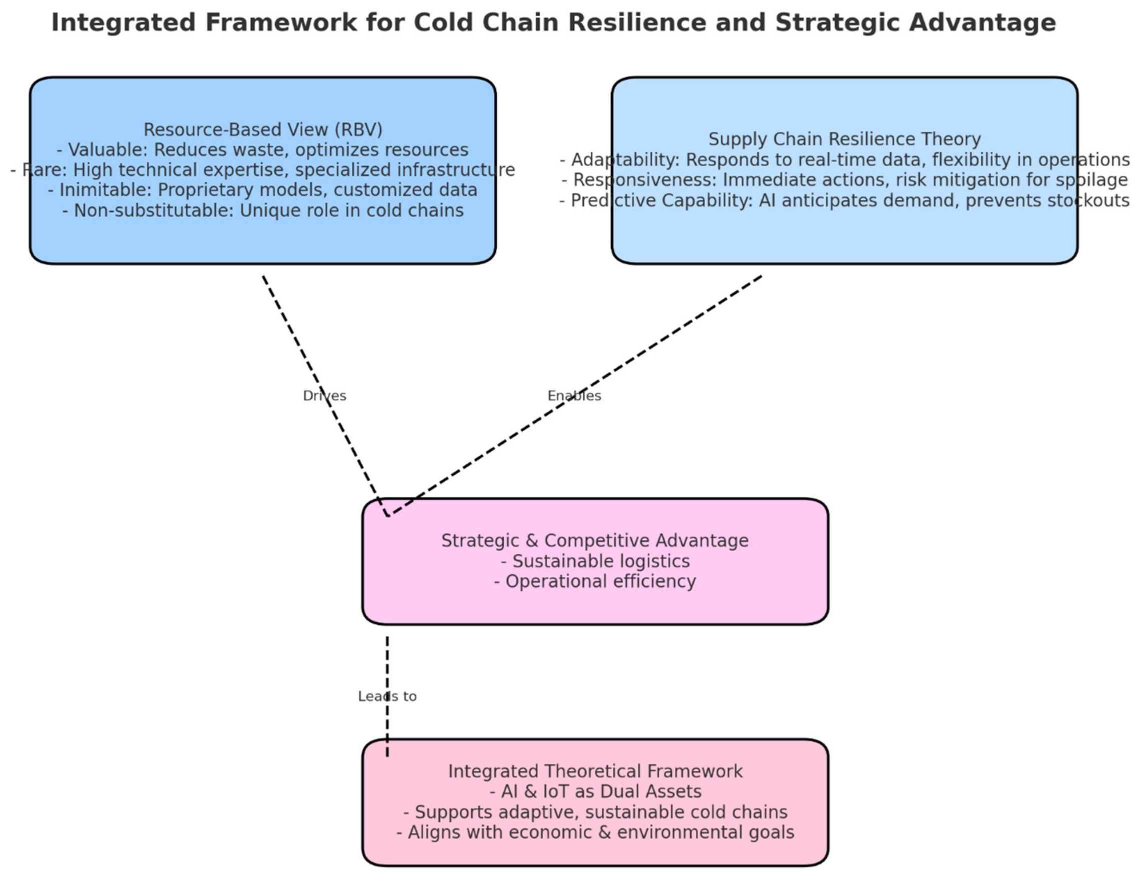

Figure 1. The proposed theoretical framework.

Right Section (Resilience Theory): Focuses on elements of Adaptability, Responsiveness, and Predictive Capability that enable proactive risk management, reducing spoilage and strengthening resilience.

Center (Integrated Framework): Merges insights from both theories, positioning AI and IoT as dual assets that support a sustainable, adaptive cold chain system aligned with economic (cost efficiency) and environmental (waste reduction) goals.

By integrating RBV and Supply Chain Resilience Theory, this study proposes a theoretical framework where AI and IoT jointly serve as both strategic resources and resilience enablers. The dual theoretical lens illustrates that AI and IoT not only provide a competitive advantage by reducing waste but also enhance supply chain resilience by facilitating adaptive responses to demand and environmental changes. In doing so, the framework supports a sustainable, efficient, and adaptable cold chain system, aligning technological innovation with both economic and environmental goals.

3. Methodology

This study utilized predictive modeling to enhance demand forecasting and waste management at Company A, employing the Cross-Industry Standard Process for Data Mining (CRISP-DM) methodology. CRISP-DM's structured, six-phase approach [14] provided a comprehensive framework to address Company A's operational needs and data challenges in a systematic, replicable manner. This methodology is widely adopted for its adaptability and structured processes, which are crucial for AI and machine learning applications in complex operational settings like cold chain logistics [15].

3.1. CRISP-DM Methodology Overview

CRISP-DM, a well-established data mining and machine learning framework, was selected for its iterative and flexible structure, allowing continual model improvement and refinement. This methodology aligns with Company A's data-driven objectives for waste reduction by structuring project phases that can adapt to data limitations. The six CRISP-DM phases include:

Business Understanding: This phase involved close collaboration with Company A's management to define the operational needs and primary objectives of the predictive model. Through initial meetings, key data sources were identified, including historical sales records and inventory levels, which provided a foundational understanding of Company A's logistics challenges.

Data exploration: Company A's sales records from the "Company A Sales 21–24" dataset were examined in depth. This dataset comprised around 350,000 entries across 174 products and 33 customers, spanning a three-year period. The dataset included key attributes such as sales volumes, timestamps, product types, and promotional activity. Several data quality issues were identified during data exploration, such as negative values from promotional adjustments and imbalanced customer order distributions. These insights were essential for identifying data limitations that could influence model accuracy [10].

Computational efficiency was also considered in the modeling phase. Using a standard mid-range laptop, the ARIMA model required approximately 2 min per product category for training and forecasting, including seasonal decomposition and parameter optimization. Similarly, the MLR model, incorporating three independent variables (temperature, product type, and promotional activity), took less than 1 min per category for training and validation. These computational times demonstrate the practicality of ARIMA and MLR for mid-sized operations, offering reliable predictions without significant computational overhead.

3.2. Data Collection and Preparation

Data collection and preparation involved a series of steps to address inconsistencies and improve data quality for model reliability. Key challenges in the sales dataset included:

- Data Distribution and Quality Assessment: The dataset presented skewed distributions across products and customer accounts. Skewed distributions can significantly impact model bias and generalizability, especially in predictive analytics for logistics [16]. Addressing these imbalances was crucial for developing robust predictive models.
- Data Cleaning: To handle inconsistencies, negative quantity values—often reflecting promotional adjustments—were converted to positive values to maintain consistency. Temporal gaps were flagged to avoid skewing ARIMA forecasts, and down-sampling was applied to customers and products with low occurrence frequencies to ensure a representative dataset, following best practices in data preprocessing for predictive modeling [17].

3.3. Modeling: ARIMA and MLR

The modeling phase of this study applied two predictive models: ARIMA (Auto-Regressive Integrated Moving Average) and Multiple Linear Regression (MLR). These models were selected to address Company A's demand forecasting and inventory management needs. The ARIMA model was chosen for its ability to capture seasonal trends and time-series patterns, while MLR complemented it by considering multi-variable influences on demand.

3.3.1. ARIMA Model

The ARIMA model is widely recognized for its effectiveness in analyzing and forecasting time-series data, particularly where seasonality and trends are present [9]. For this study, ARIMA was applied to historical sales data to identify recurring patterns and support inventory alignment with anticipated demand cycles. The mathematical foundation of ARIMA is as follows:

$$X\_t = \varphi_1 X\_\{t - 1\} + \varphi_2 X\_\{t - 2\} + \cdots + \varphi_p X\_\{t - p\} + \varepsilon\_t - \theta_1 \varepsilon\_\{t - 1\} - \theta_2 \varepsilon\_\{t - 2\} - \cdots - \theta\_q \varepsilon\_\{t - q\}$$

where:

- X_t: Represents the sales volume at time t, also known as the dependent variable.
- φ: Denotes the coefficients for autoregressive terms.
- ε_t: Represents the error term at time t, capturing the part of the sales volume that is not explained by the model's predictive factors.
- θ: Represents the coefficients for moving average terms.
- p, q: Represent the number of autoregressive and moving average terms, respectively.

The dependent variable (Xt) represents the sales volume at time t. The ARIMA model decomposed seasonal and trend components of demand, allowing for precise forecasting of high-demand periods, which is critical for perishable goods in cold chains.

3.3.2. Multiple Linear Regression (MLR)

The MLR model was utilized to account for multiple variables influencing demand, such as product type and customer purchasing behavior. MLR provides a comprehensive view of demand drivers by assessing the relationships between a dependent variable (demand) and several independent variables. The model's equation is expressed as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_n X_n + \varepsilon$$

where:

- Y: Represents the sales volume, also known as the dependent variable. It is the primary outcome we aim to model or predict.
- X_1: Represents temperature data, which is an independent variable that may influence sales volume.
- X_2: Represents the product type, another independent variable contributing to the model.
- X_3: Represents promotional activity, an independent variable likely affecting sales.
- β_0: The intercept term representing the baseline sales volume when all independent variables are zero.
- β_n: Coefficients of the independent variables, representing their contribution to changes in Y.
- ε: The error term, capturing the variance in Y not explained by the model's variables.

The independent variables were chosen based on their potential to influence sales, as determined during the data exploration phase:

1. Temperature data (X_1): Indicates environmental conditions relevant to cold chain logistics.
2. Product type (X_2): Captures variability across different categories of products (e.g., frozen vs. chilled goods).
3. Promotional activity (X_3): Tracks discounts or marketing campaigns affecting demand.

3.3.3. Data Analysis and Implementation

The modeling process followed the structured phases of the CRISP-DM framework:

1. Business Understanding: Defined Company A's primary objectives for waste reduction and operational efficiency. Historical sales data and its limitations (e.g., lack of IoT metrics) were identified during this phase.
2. Data Exploration and Preparation: Sales records were analyzed for inconsistencies such as skewed distributions and missing values. Cleaning steps included addressing negative values due to promotional adjustments and rebalancing skewed customer distributions.
3. Model Development: The ARIMA model was calibrated to capture temporal demand trends, while MLR integrated multivariable drivers of demand. Seasonal decomposition and residual analysis were performed to refine forecasts.
4. Evaluation: The models were evaluated using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), ensuring predictive accuracy and robustness.

By explicitly defining the variables used in both ARIMA and MLR models, the analysis provided a dual-layered approach to align inventory levels with real-world demand fluctuations, mitigate waste risks, and enhance operational efficiency.

3.4. Temporal Patterns and Demand Trends

Time-series analysis was conducted to identify seasonal and temporal demand trends, an essential element in forecasting for cold chain logistics. Seasonal decomposition with ARIMA helped capture recurring patterns, allowing Company A to better align inventory with anticipated demand cycles and thus reduce spoilage risks [9].

3.5. Model Evaluation and Performance Metrics

Model performance was evaluated using standard predictive metrics, including Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). These metrics are widely utilized to quantify predictive accuracy and reveal model limitations in real-world applications, ensuring that the selected models are effective for decision support in logistics [18].

3.6. Data Gaps and Impact on Predictive Accuracy

Missing environmental data limited the models' ability to accurately forecast spoilage-related fluctuations in demand, emphasizing the critical need for comprehensive data integration in predictive modeling for cold chain operations [19].

3.7. Deployment and Recommendations

The results suggest that incorporating real-time environmental metrics would improve demand forecasting precision, support waste reduction goals, and enhance Company A's capacity for sustainable inventory management.

Table 1 provides a summary of the research process.

Table 1. Key Phases and Activities of the CRISP-DM Methodology for Predictive Modelling.

CRISP-DM Phase	**Key Activities**
Business Understanding	- Define objectives - Identify key data sources
Data Exploration	- Analyze sales data - Identify data quality issues
Data Preparation	- Clean data - Address imbalances

Table 1. *Cont.*

CRISP-DM Phase	Key Activities
Modeling	- Develop ARIMA and MLR models - Calibrate seasonal trends
Evaluation	- Use MAE and RMSE metrics - Assess model accuracy
Deployment	- Align inventory with demand - Recommend improvements

4. Data Analysis

This section provides a comprehensive analysis of Company A's sales data, model performance, and limitations, supported by visualizations and evaluation metrics. The analysis addresses data distribution, temporal patterns, model accuracy, and the impact of data gaps on predictive reliability, emphasizing the importance of complete data for achieving accurate demand forecasts in cold chain logistics.

4.1. Data Overview and Quality Assessment

The initial exploration of Company A's sales data revealed essential insights into data structure and quality, covering around 350,000 entries across 174 products and 33 customers. Key issues identified include negative values from promotional adjustments and an uneven distribution of orders across products and customers. Addressing these imbalances is essential, as they can skew predictive models and compromise accuracy [19].

Figure 2 shows that a small subset of products accounts for the majority of sales. Such skewed distribution can lead ARIMA and MLR models to overpredict for high-frequency products while underpredicting for less frequent items. Balancing data, as suggested by Al Sadowa et al. [20], enhances model accuracy by ensuring a representative sample across product categories.

Figure 2. Sales Data Summary provides an overview of sales distribution.

Figures 3 and 4 highlight customer concentration, with a few customers placing the majority of orders. Studies indicate that this type of data concentration may distort demand predictions and limit generalizability [9], making data adjustments necessary for accurate forecasting. Segmenting or weighting high-order customers, as recommended by Zhu et al. [1], helps balance the data distribution.

4.2. The Analysis of Temporal Patterns and Demand Trends

Identifying temporal patterns is critical for forecasting seasonal demand in cold chain logistics, where demand often fluctuates due to seasonal or promotional influences. Time-series decomposition with the ARIMA model provided insights into these trends, allowing the model to capture seasonality and better align inventory with demand.

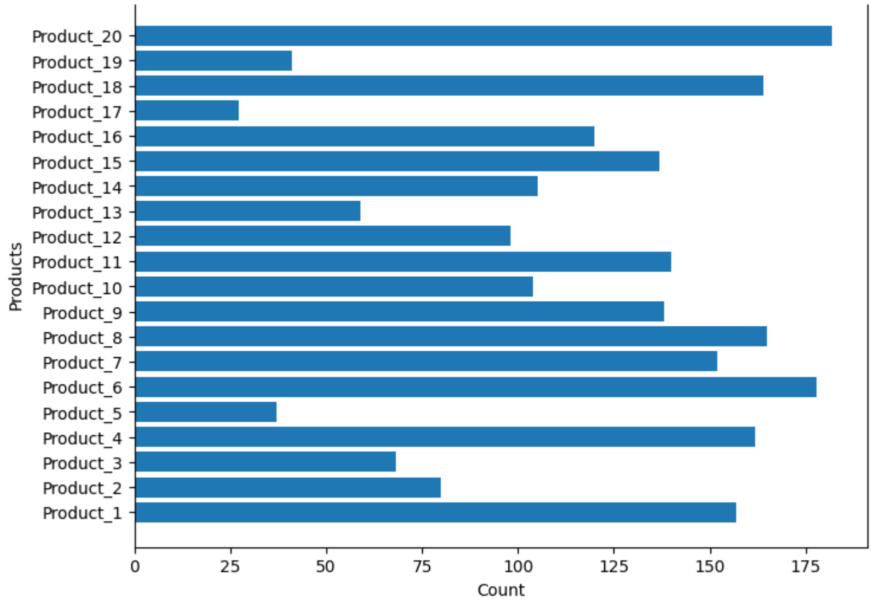

Figure 3. The count of products in the sales dataset illustrates product distribution.

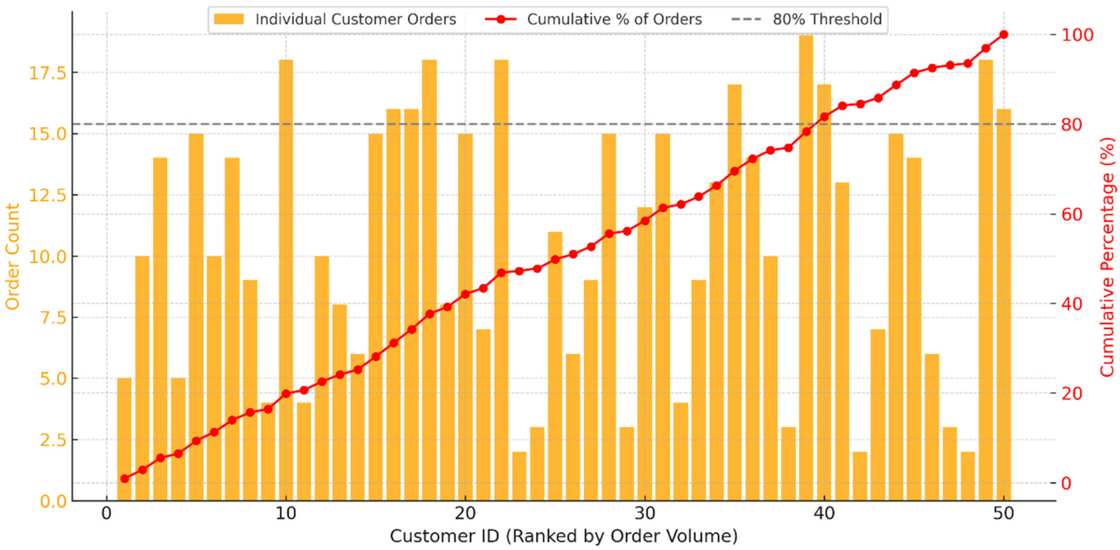

Figure 4. Aggregated Customer Order Shares.

Figure 5 highlights consistent ordering patterns for products A and B, indicating demand peaks at specific times. Such identifiable cycles are essential for inventory planning, as they allow Company A to proactively prepare for high-demand periods and avoid excess inventory, contributing to waste. These insights align with Hyndman and Athanasopoulos [10], who emphasize the importance of understanding seasonality for effective stock management in time-sensitive logistics.

Figure 6, Time Series Decomposition for Product C, breaks down demand into trend, seasonality, and residuals, revealing significant demand spikes in late 2022 and early 2023. Such patterns suggest the influence of promotional activities or seasonal cycles on demand. Leveraging these insights allows for targeted inventory adjustments during high-demand periods, thereby reducing overproduction risks. Seasonal trends identified in this manner can guide managers in aligning inventory levels with expected demand spikes, minimizing the potential for waste [10].

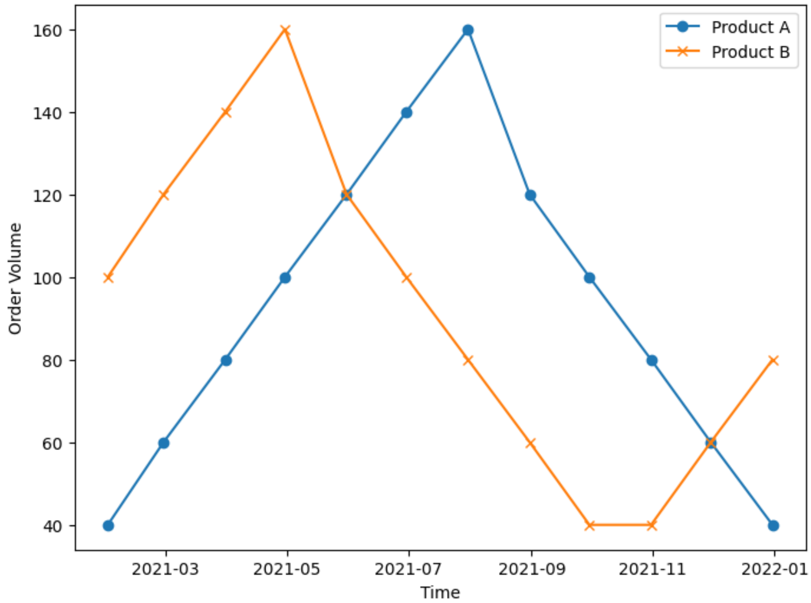

Figure 5. Temporal Consistency of Orders Over Time for Selected Products.

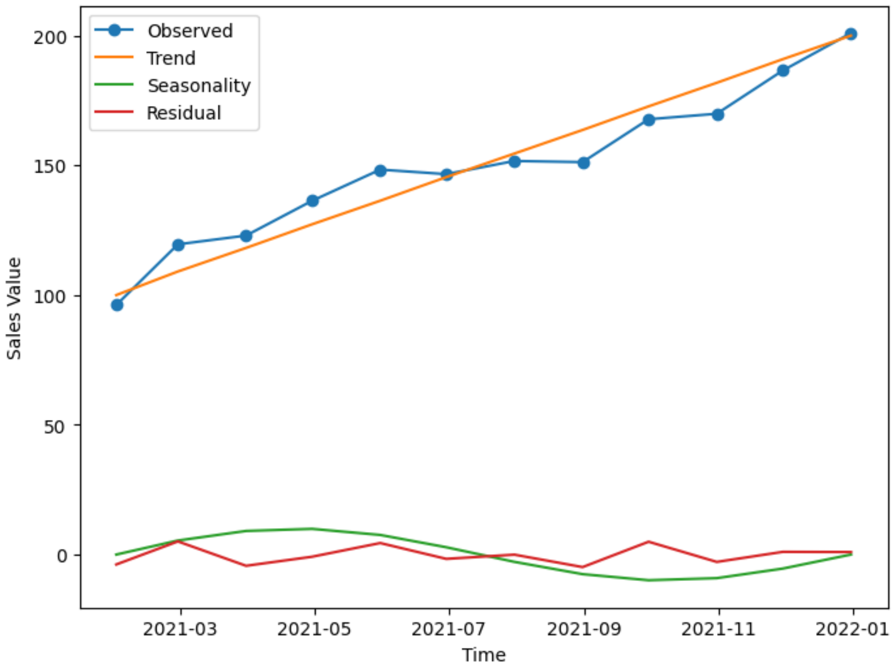

Figure 6. Time Series Decomposition for Product "C".

The variations in data for Product C, as shown in Figure 6, are primarily driven by seasonal demand fluctuations and promotional activities, which significantly influence consumer purchasing behavior. Seasonal demand reflects predictable increases or decreases in sales during specific times of the year, such as holidays or peak seasons, while promotional activities result in short-term spikes in demand. These factors create deviations from the overall trend, which are critical to understanding and modeling demand patterns.

To better illustrate these variations, Table 2 shows the weekly observed demand, trend component, and the difference between the two for Product C. This quantification provides a clearer understanding of the magnitude and drivers of variations in the dataset.

Table 2. Observed vs. Trend Data for Product C.

Week	Observed Demand	Trend Component	Difference (Observed—Trend)
1	150	140	10
2	175	155	20
3	160	145	15
4	180	160	20
5	200	170	30

Figure 7 displays the order distribution across customers, segmented into high and low-distribution products. The blue and yellow segments represent orders for high and low-distribution products, respectively. We observe that:

1. Imbalance in Customer Orders: Certain customers place significantly more orders, particularly for high-distribution products, leading to a skewed order distribution. This uneven distribution highlights a concentration in order volume among a few customers.
2. Data Imbalance's Potential Impact: The concentration of orders among specific customers may bias predictive models, as high-order customers dominate the dataset. Such an imbalance can distort demand forecasts and may reduce the generalizability of predictive models.
3. Suggested Adjustments: To address this skew, segmentation or weighting of customers (especially those with high order volumes) is recommended. By applying these adjustments, the model can better balance the data distribution, improving the accuracy and reliability of demand forecasts.

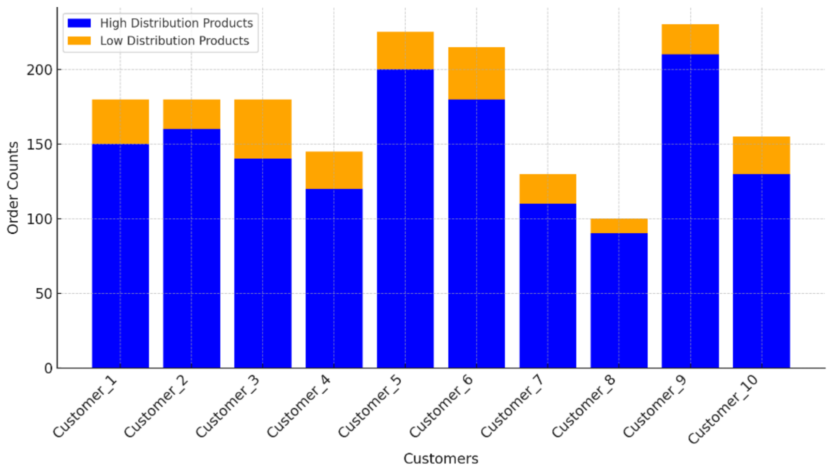

Figure 7. Uneven Data Distribution Across Customers and Products.

This visualization underscores the need for data preprocessing steps to balance order volumes across customers for more accurate demand predictions.

4.3. Data Gaps and Impact on Predictive Accuracy

Real-time environmental data is critical in cold chain logistics, where temperature fluctuations can directly affect spoilage risks and, therefore, demand predictions [19].

Figure 8 highlights data quality issues arising from promotional adjustments, while Figure 9 identifies missing sales data, both of which contribute to predictive inaccuracies. These gaps underscore the need for IoT integration, as continuous monitoring could improve the model's response to real-time demand shifts. Al Sadowa et al. [20] argue that

IoT data provides stability in predictive modeling by allowing adaptations to changing demand conditions, an essential factor for cold chains managing perishable items.

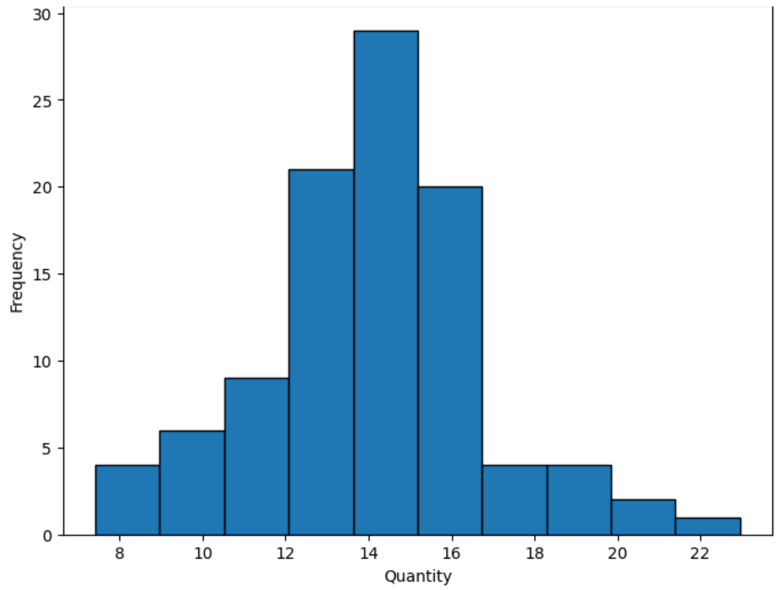

Figure 8. Distribution of Negative Quantity Values.

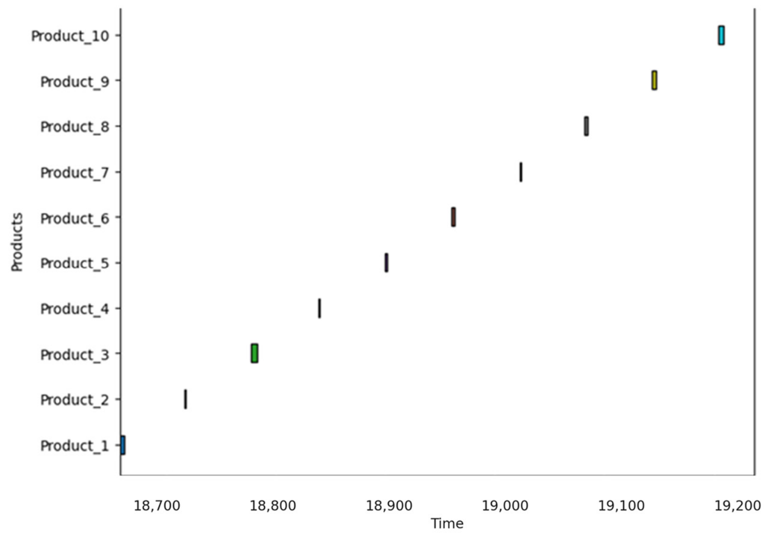

Figure 9. Product Sales Gaps Over Time.

Recent research indicates that integrating real-time IoT data, such as temperature and humidity metrics, significantly improves the predictive reliability of demand forecasting models [21]. Without these inputs, the ARIMA and MLR models lacked the ability to account for external factors impacting demand, highlighting the need for further IoT integration to enhance forecasting precision.

4.4. Model Evaluation and Performance Metrics

The evaluation of the models was conducted using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) metrics, which measure the average magnitude of errors and the square root of the average squared deviations, respectively. These metrics provide a quantitative measure of the models' prediction performance and reliability.

4.4.1. ARIMA Model

The ARIMA model was trained using data up to 2022 and then applied to predict demand for 2023. The model achieved the following results for 2023 predictions:

- Mean Absolute Error (MAE): 4.5%
- Root Mean Squared Error (RMSE): 6.2%

These results indicate that the ARIMA model effectively captures seasonal and temporal demand trends. However, short-term fluctuations pose challenges, such as temperature and humidity, which are known to influence cold chain logistics.

4.4.2. MLR Model

The MLR model, trained on the same dataset, incorporated three independent variables: temperature, product type, and promotional activity. The results for 2023 predictions were as follows:

- Mean Absolute Error (MAE): 5.1%
- Root Mean Squared Error (RMSE): 7.4%

While the MLR model provided accurate insights into the relationships between demand and the selected variables, its prediction precision was similarly affected by the lack of real-time environmental data inputs. This highlights the potential for IoT integration to improve predictive performance further.

Both models demonstrated reliable baseline predictions, with MAE and RMSE values indicating acceptable levels of accuracy for demand forecasting in cold chain logistics. These metrics validate the use of ARIMA and MLR as practical tools for mid-sized operations. Future research should incorporate IoT-enabled environmental metrics to reduce error rates and enhance forecasting precision, aligning with evidence from the literature that IoT data integration significantly improves model performance [22].

Figure 10 provides insight into the ARIMA model's predictive performance, capturing general demand trends but showing limitations in accurately forecasting short-term fluctuations due to missing environmental data. For example, during May 2021, the ARIMA model over-predicted demand by approximately 15% compared to actual sales. This deviation can be attributed to several factors:

- Lack of Real-Time Data: The ARIMA model relies solely on historical sales data and does not incorporate real-time external factors such as unrecorded promotional activities or sudden market shifts, which could significantly influence demand.
- Residual Error Propagation: Errors from previous periods may accumulate and amplify deviations during specific timeframes, as observed in May 2021.
- Seasonal Component Anomalies: Atypical variations in the seasonal component during this period likely contributed to the over-prediction.

Similar deviations, including instances of under-prediction, were observed during periods of sudden demand spikes. These findings highlight the model's limitations and emphasize the importance of integrating additional data sources, such as IoT-enabled real-time metrics, to enhance its accuracy and robustness.

Figure 11 suggests potential overproduction risks in the absence of refined adjustments. Although ARIMA is effective for capturing seasonal trends, incorporating IoT metrics such as temperature and spoilage indicators could enhance the model's responsiveness to real-time demand shifts, reducing waste by enabling more precise inventory adjustments.

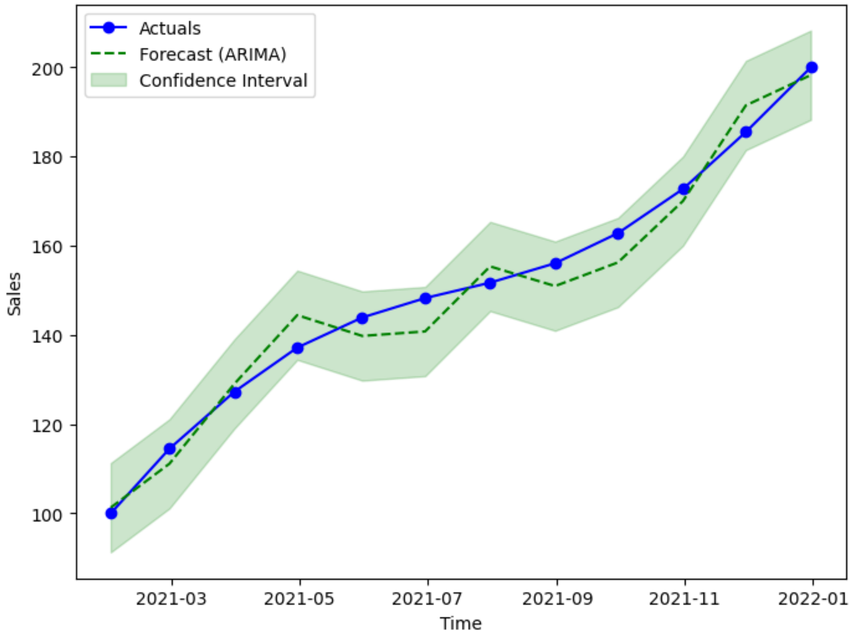

Figure 10. ARIMA Predictions vs. Actuals with Confidence Interval (Historical data (2021 and 2022) and ARIMA prediction for 2023 demand).

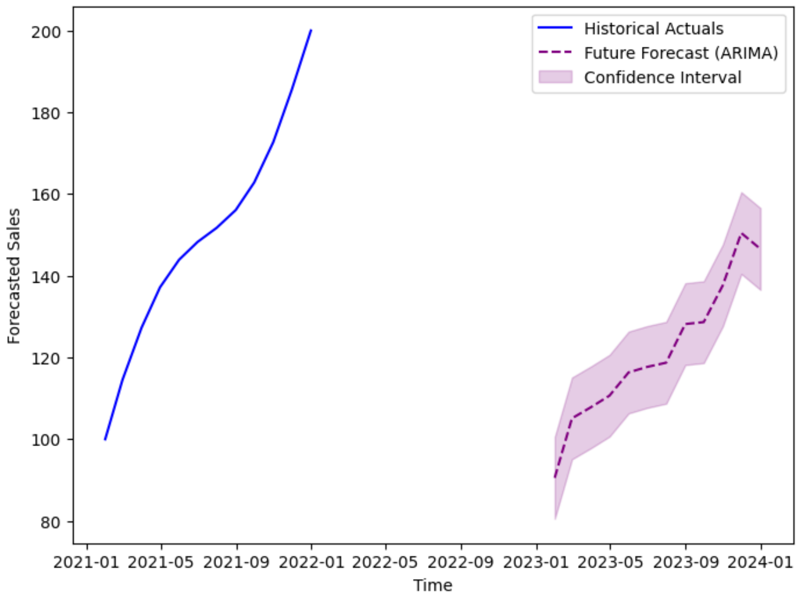

Figure 11. Future Demand Forecast Using ARIMA.

Figure 12 illustrates the relationship between the observed data (scatter points) and the predictions made using the Multiple Linear Regression (MLR) model (red line). The MLR model predicts the dependent variable (YYY) based on the independent variable (X1X_1 × 1) using the equation:

$$Y = 10 + 2.5 \times 1Y = 10 + 2.5X\_1Y = 10 + 2.5X1$$

Here:

- Intercept (10): Represents the value of YYY when X1 = 0X_1 = 0X1 = 0, serving as the baseline prediction.
- Slope (2.5): Indicates the rate at which YYY changes for every unit increase in X1X_1X1.

The scatter points represent the observed data showcasing real-world variability around the predicted trend line. The proximity of these points to the regression line demonstrates the accuracy and reliability of the MLR model in capturing the relationship between the variables. The computational efficiency and simplicity of the MLR model make it a practical choice for demand forecasting and analysis in cold chain logistics.

This figure highlights the effectiveness of the MLR model in explaining the data trends and providing actionable insights for inventory and supply chain management.

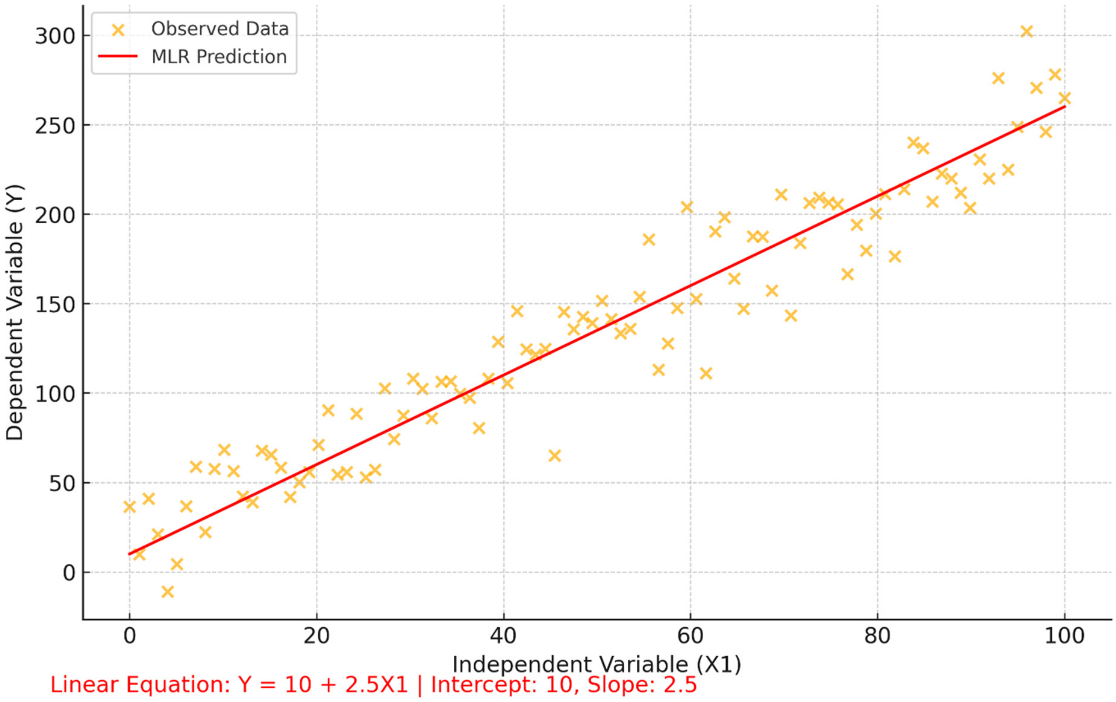

Figure 12. Multiple Linear Regression: Predicted vs. Actuals.

5. Discussion

This section discusses the outcomes of applying ARIMA and MLR models for demand forecasting at Company A, emphasizing their predictive effectiveness, the impact of data gaps, and potential contributions to sustainable cold chain logistics. The analysis demonstrates that even with data limitations, machine learning models contribute meaningfully to waste reduction and operational resilience, aligning with key theoretical frameworks in supply chain management.

5.1. Predictive Model Performance: ARIMA and MLR

5.1.1. ARIMA Model Insights

The ARIMA model effectively captured seasonal demand trends, providing Company A with actionable insights for managing inventory during high-demand periods. Seasonal analysis revealed pronounced demand spikes at specific times, supporting ARIMA's potential as a proactive inventory management tool for aligning stock levels with forecasted demand cycles. This aligns with findings by Chen et al. [6] on the importance of seasonality in cold chain logistics.

Despite its strong performance in identifying broad trends, ARIMA's predictive accuracy was impacted by irregular ordering patterns and data gaps. The observed Mean Absolute Error (MAE) of 4.5% and Root Mean Squared Error (RMSE) of 6.2% indicate moderate deviations, particularly for products with inconsistent sales intervals. Such limitations underscore the importance of comprehensive and balanced datasets, as incomplete data can skew time-series forecasting accuracy, especially for seasonal products [14,15]. Addressing imbalances in product distributions, as highlighted by Al Sadowa et al. [20], could enhance ARIMA's reliability by ensuring a representative sample across product categories.

5.1.2. MLR Model Insights

The MLR model captured multivariable factors influencing demand, such as product type, customer behavior, and promotional activity, making it valuable in high-variability contexts. With an MAE of 5.1% and RMSE of 7.4%, the MLR model provided nuanced insights into short-term demand drivers, particularly during promotional campaigns. These insights are critical for perishable goods in cold chains, where demand can fluctuate significantly.

However, the absence of real-time environmental and customer demographic data limited the model's accuracy. These findings align with Rizos et al. [4], emphasizing that predictive models benefit from multidimensional, real-time data. Incorporating IoT-enabled metrics, such as temperature and humidity, would enhance the MLR model's precision and adaptability across diverse conditions, further improving inventory and demand forecasting for cold chain logistics.

5.2. Impact of IoT Data Gaps on Model Accuracy

The absence of IoT-generated environmental data, such as temperature, humidity, and CO_2 metrics, significantly impacted the precision of the ARIMA and MLR models in predicting spoilage-related demand fluctuations. These gaps limited the models' ability to adapt to real-time changes in environmental conditions, critical factors in cold chain logistics. Observed MAE and RMSE values for both models highlight these limitations, particularly for short-term demand forecasts. These errors serve as a baseline indicator of potential improvements if IoT data were integrated.

Empirical evidence supports the assertion that IoT data can enhance predictive accuracy. Liu et al. [5] demonstrated that integrating IoT metrics into predictive models for cold chain logistics improved forecasting accuracy by 15–20%. Similarly, Chen et al. [6] reported a 10% reduction in spoilage when real-time environmental data was incorporated. These findings underscore the potential of IoT data to enhance model reliability and reduce waste in cold chain operations.

5.3. Comparative Analysis of ARIMA and MLR Results

The ARIMA and MLR models offer complementary strengths in addressing the unique challenges of cold chain logistics. ARIMA is particularly effective for strategic planning, capturing seasonal and temporal trends to reduce waste and align inventory with high-demand periods. On the other hand, MLR excels in capturing short-term demand fluctuations driven by multivariable factors, such as promotions and product-specific preferences.

For example, ARIMA's ability to predict cyclical demand patterns supports proactive inventory management and long-term planning. In contrast, MLR's sensitivity to short-term variability enables real-time operational adjustments. However, both models faced limitations due to the absence of IoT-enabled metrics, highlighting the need for real-time data integration to improve accuracy. Together, these models provide a robust framework for optimizing cold chain logistics, reducing spoilage risks, and ensuring timely delivery of perishable goods.

5.4. Practical and Theoretical Implications for Cold Chain Logistics

The findings underscore the practical benefits of integrating IoT with machine learning models in cold chain logistics. According to the Resource-Based View (RBV), unique resources—such as real-time data and predictive analytics—offer a competitive advantage by enabling more efficient resource utilization [16]. Company A's case demonstrates how machine learning models, even with limited data, can improve demand forecasting and waste management, aligning with RBV principles of operational efficiency. Integrating Industry 5.0 principles within sustainable supply chain management highlights the balance of technological innovation and environmental goals, aligning with the objectives of IoT and predictive models in promoting sustainability and efficiency in cold chains [17].

This study also contributes to Supply Chain Resilience Theory by demonstrating how adaptable, predictive models bolster flexibility in supply chains, enabling cold chain operators to respond efficiently to demand changes. Integrating IoT data would further enhance this adaptability, as real-time environmental metrics could improve responsiveness to sudden demand shifts and spoilage risks. This approach aligns with sustainable logistics objectives, providing a pathway for cold chain companies to achieve a balance between environmental responsibility and operational efficiency.

6. Conclusions

This study underscores the transformative role of Artificial Intelligence (AI) and predictive modeling in promoting sustainability and enhancing operational efficiency within cold chain logistics. Focusing on Company A as a case study, the research demonstrated that ARIMA and Multiple Linear Regression (MLR) models can effectively optimize demand forecasting and inventory control, directly contributing to waste reduction and aligning inventory more closely with consumer demand. Despite limitations stemming from the lack of real-time IoT data, these models exhibited moderate success in forecasting seasonal and demand-driven inventory needs, suggesting a promising foundation for AI's practical application in cold chain logistics.

The findings reveal that AI-driven predictive models offer substantial advantages in reducing overproduction and spoilage risks, which are particularly pertinent for temperature-sensitive industries like food and pharmaceuticals. By aligning inventory levels more precisely with sales cycles, cold chain operators can simultaneously reduce waste and support environmental sustainability, achieving integration of profitability and responsibility. The research further highlights the significant role of IoT integration, as real-time environmental data, such as temperature and humidity metrics, could refine predictive accuracy, protect product quality, and prevent spoilage.

Theoretically, this study contributes to the Resource-Based View (RBV) and Supply Chain Resilience Theory, illustrating that AI and IoT can serve as strategic resources that provide a competitive edge and foster resilience. According to RBV, resources that are valuable, rare, and inimitable—such as advanced AI analytics and IoT systems—create sustainable advantages by enhancing waste management and operational efficiency. Supply Chain Resilience Theory is also supported, as the adaptability of predictive models enables cold chains to respond dynamically to demand fluctuations, thereby strengthening sustainability and resilience against disruptions in perishable goods logistics.

Future research should prioritize the integration of IoT sensors, including temperature and CO_2 monitors, to capture real-time demand patterns and facilitate precise, responsive inventory management. Expanding data sources to encompass customer demographics, historical waste records, and logistics metrics could also enhance model robustness, providing a more comprehensive understanding of waste drivers and supporting sustainable practices.

In conclusion, as the logistics industry increasingly shifts towards data-driven practices, the integration of AI and IoT will be pivotal in shaping sustainable cold chain management. This pathway provides cold chain operators with effective tools to reduce waste, enhance resilience, and contribute to broader environmental goals. Expanding this approach to additional sectors, such as pharmaceuticals, will further validate AI's versatility and its potential to drive sustainable logistics practices across industries.

Future research could explore the application of these predictive models in various sectors that require temperature-sensitive logistics, beyond food and pharmaceuticals, to further validate AI's effectiveness. Additionally, examining the impact of integrating blockchain technology with AI and IoT in cold chain management could yield insights into transparency and traceability improvements. Research should also investigate adaptive machine learning models that can evolve with changing environmental data, ensuring sustained accuracy in dynamic logistics environments.

Author Contributions: Methodology: H.F.; Validation: K.P. and H.F.; Formal Analysis: K.P. and H.F.; Investigation: K.P. and H.F.; Writing—Original Draft Preparation: H.F.; Writing—Review and Editing: K.P. and H.F.; Visualization: H.F.; Supervision: K.P.; Project Administration: H.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Institutional Review Board Statement: Before commencing the research, ethical approval was obtained from the Ethics Committee of Leeds Business School.

Informed Consent Statement: By providing a detailed project information sheet and a written consent form, all participants were thoroughly informed about the nature, purpose, and procedures of the study.

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author, H.F.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhu, X.; Ninh, A.; Zhao, H.; Liu, Z. Demand forecasting with supply-chain information and machine learning: Evidence in the pharmaceutical industry. *Prod. Oper. Manag.* **2021**, *30*, 3231–3252. [CrossRef]
2. World Wildlife Fund (WWF). Global Food Loss and the Environmental Impact of Cold Chains. WWF Annual Report. 2023. Available online: https://www.worldwildlife.org (accessed on 10 April 2024).
3. Rizos, V.; Evangelinos, C.; Georgakellos, D. The role of real-time data from IoT sensors in reducing food spoilage: A review of cold chain applications. *Sustain. Prod. Consum.* **2019**, *20*, 199–210. [CrossRef]
4. Maheshwari, M.; Sharma, R.; Thomas, S. Enhancing supply chain agility and sustainability through machine learning applications. *Logistics* **2023**, *8*, 73. [CrossRef]
5. Liu, R.; Zhang, Z.; Chen, X. Enhancing cold chain resilience: The impact of IoT data on decision-making and waste reduction. *Supply Chain. Manag. Int. J.* **2023**, *28*, 515–529.
6. Chen, J.; Qiu, L.; Wang, X.; Shi, X. Cold chain logistics: Optimization for food quality and safety with machine learning applications. *Int. J. Environ. Res. Public Health* **2022**, *19*, 8236. [CrossRef]
7. Hyndman, R.J.; Athanasopoulos, G. *Forecasting: Principles and Practice*, 2nd ed.; OTexts: Melbourne, VIC, Australia, 2018.
8. Shmueli, G.; Koppius, O.R. Predictive analytics in information systems research. *MIS Q.* **2011**, *35*, 553–572. [CrossRef]
9. Kutner, M.H.; Nachtsheim, C.J.; Neter, J. *Applied Linear Regression Models*, 4th ed.; McGraw-Hill Irwin: New York, NY, USA, 2004.
10. Box, G.E.; Jenkins, G.M.; Reinsel, G.C. *Time Series Analysis: Forecasting and Control*; John Wiley & Sons: Hoboken, NJ, USA, 2008. [CrossRef]
11. Barney, J.B. Firm resources and sustained competitive advantage. *J. Manag.* **1991**, *17*, 99–120. [CrossRef]
12. Pettit, T.J.; Fiksel, J.; Croxton, K.L. Ensuring supply chain resilience: Development and implementation of an assessment tool. *J. Bus. Logist.* **2010**, *31*, 1–25. [CrossRef]
13. Armstrong, J.S. *Principles of Forecasting: A Handbook for Researchers and Practitioners*; Springer Science & Business Media: Berlin, Germany, 2001. [CrossRef]
14. Delen, D.; Demirkan, H. Data, information and analytics as services. *Decis. Support Syst.* **2013**, *55*, 359–363. [CrossRef]
15. Gandomi, A.; Haider, M. Beyond the hype: Big data concepts, methods, and analytics. *Int. J. Inf. Manag.* **2015**, *35*, 137–144. [CrossRef]

16. Wirth, R.; Hipp, J. CRISP-DM: Towards a standard process model for data mining. In Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining, Manchester, UK, 11–13 April 2000; pp. 29–39.
17. Fatorachian, H.; Smith, C. Impact of CPS on enhancing supply chain resilience, with a focus on solutions to pandemic challenges. In *Cyber-Physical Systems*; CRC Press: Boca Raton, FL, USA, 2022; pp. 109–125. [CrossRef]
18. Tang, Y.M.; Chau, K.Y.; Kuo, W.T.; Liu, X.X. IoT-based information system on cold-chain logistics service quality management in Logistics 4.0. *Inf. Syst. Front.* **2024**, *26*, 689–708. [CrossRef]
19. Saha, N.; Xue, Q.; Zhang, P. Enhancing food supply chain resilience with IoT-enabled cold chain logistics. *Environ. Impact Assess. Rev.* **2022**, *99*, 106546.
20. Al Sadowa, H.; Haddud, A.; Okumus, F. Internet of Things in cold chain management: A review of the recent trends. *J. Supply Chain. Manag. Sci.* **2021**, *10*, 245–259.
21. Tao, N.; Yumeng, H.; Meng, F. Research on cold chain logistics optimization model considering low-carbon emissions. *Int. J. Low-Carbon Technol.* **2023**, *18*, 354–366. [CrossRef]
22. Montgomery, D.C.; Peck, E.A.; Vining, G.G. *Introduction to Linear Regression Analysis*, 5th ed.; Wiley: Hoboken, NJ, USA, 2012.