ARTICLE TEMPLATE

Revolutionize Cold Chain: An AI/ML Driven Approach to Overcome Capacity Shortages

Ilya Jackson[a*], Jafar Namdar[a], Maria Jesús Saénz[a], Richard Augustus Elmquist III[a], Luis Rodrigo Dávila Novoa[a]

[a]Massachusetts Institute of Technology, Center for Transportation & Logistics, 1 Amherst Street, Cambridge, MA, USA;

ARTICLE HISTORY
Compiled January 28, 2024

ABSTRACT
This research investigates how Artificial Intelligence (AI) and Machine Learning (ML) forecasting methodologies can be leveraged for cold chain capacity planning, specifically utilizing Prophet and Seasonal Autoregressive Integrated Moving Average parametrized through grid search. In collaboration with Americold, the world's second-largest refrigerated logistic service provider, the study explores the challenges and opportunities in applying AI/ML techniques to complex operations covering 385 customers and a capacity of 73,296 pallet positions. We train and test several AI/ML and traditional statistical models using extensive data for every customer over 3.5 years. Based on the results, MAPE of 5% was achieved on the whole site level, and SARIMA outperformed ML models in most cases. Next, we show that developing and applying a Customer Segmentation Matrix has enabled more accurate forecasting and planning across various customer segments, addressing the issue of forecasting inaccuracies. This approach effectively improves forecasting inaccuracies, underscoring the significance of tailoring AI/ML models for demand forecasting within the cold-chain industry. Ultimately, this research presents an AI-driven approach that transcends mere forecasting, offering a practical pathway to manage capacity in light of the constraints.

KEYWORDS
Artificial intelligence; Machine Learning; Forecasting; Capacity Planning; Cold Chain; Supply Chain

1. Introduction

A "cold chain" is a logistical process that involves the transportation and storage of temperature-sensitive products within a specific temperature range to maintain their quality and safety. It includes a series of temperature-controlled processes from raw material acquisition to the delivery of products to end consumers (Khan and Ali 2021). Also, the ever-increasing demand for fresh agricultural products, pharmaceuticals, and biotechnological goods requires efficient and reliable cold chains (Yu and Xiao 2021). The deterioration of fresh food can lead to substantial economic losses and food safety issues (Wang and Zhao 2021). Cases of mishandling have resulted in significant product losses, like temperature-sensitive food items being discarded due to inadequate storage conditions (Tsang et al. 2018; Nerbovig 2017). The significance of the cold chain

*Corresponding author: Ilya Jackson. Email: ilyajack@mit.edu

becomes even more apparent when preservation during long-distance transportation results in approximately 40% value loss in fruit and vegetable (Negi and Anand 2015). Such losses not only impact the economic viability of businesses but also highlight the critical need for effective cold chains in preserving product quality and minimizing waste (Yu and Xiao 2021). Hence, to address these challenges, cold chains have emerged as a solution, employing temperature-controlled activities to slow down the deterioration of perishable items, attract safety-conscious consumers, and ensure food safety.

Capacity planning in cold chains is a complex task for several reasons. First, the products of cold chains typically possess shorter shelf lives and are highly susceptible to environmental factors such as temperature, humidity, and lighting intensity (Gormley, Brennan, and Butler 2000). Maintaining precise environmental conditions becomes paramount throughout the cold chain, requiring advanced refrigeration and dehumidification systems (Tsang et al. 2018). The real complexity arises from the broad range of ambient temperatures required, spanning from -25°C to +10°C, depending on the product type (Lana, Tijskens, and Van Kooten 2005; Soyer et al. 2010). This means different types of food products have varying temperature requirements. Perishable items like dairy, seafood, and fresh produce require different storage conditions, making it challenging to maintain consistent temperatures across the cold chain. Second, food safety regulations mandate strict temperature controls (Weng et al. 2015). Non-compliance can lead to legal consequences, recalls, and damage to a company's reputation. Third, most items within the cold chain typically consist of fresh agricultural or perishable products, which inherently come with extended supply lead times that are difficult to adjust due to environmental factors (Behzadi et al. 2018). Due to the vast distance between the production base and the target market in cold chains, the lead time is not only long but also usually quite unstable (Cai et al. 2013). Also, managing activities such as harvesting, post-harvest processes, packing, processing, storage, and transportation can be highly challenging due to sudden surges in supply. This indicates that implementing a cold chain involves a long lead time and significant capital investment in refrigeration and preservation equipment (Behzadi et al. 2018; Wang and Zhao 2021). This gets more challenging since seasonal demand fluctuations and changes in consumer preferences can lead to uneven demand patterns, complicating inventory management and cold chain logistics (Prentice and McLachlin 2008).

Above all challenges, the profit margins in this domain tend to be relatively narrow which highlights the necessity for effective capacity management practices and the utilization of contemporary decision-making technological solutions (Soto-Silva et al. 2016). As a result, many firms, like Driscoll's, outsource cold chain services to logistic service providers (LSPs) to reduce operational costs and allow agri-product suppliers and retailers to focus on their core business activities (Yu and Xiao 2021; Mariappan et al. 2023). This means the role of cold chain logistics service providers (LSPs) cannot be overlooked in ensuring the seamless functioning of cold chains. These specialized LSPs bring expertise and dedicated resources, playing a crucial role in the successful implementation and management of temperature-controlled activities, which are essential for maintaining the integrity and quality of perishable items throughout the cold chain.

Additionally, the significance of cold chains goes beyond cost savings, where the lack of cold chains results in substantial value and quality losses (Yu and Xiao 2021). Recall the case of Driscoll's and Americold; the investments made in refrigeration facilities are pivotal in maintaining the quality and freshness of their products throughout the entire supply chain (Wang and Zhao 2021). Americold, a prominent refrigerated

Logistic Service Provider (LSP), invests significantly in procuring and maintaining such facilities (Wang and Zhao 2021). These investments are crucial for maintaining the freshness and integrity of perishable goods during their journey through the supply chain.

Given the distinct characteristics of cold chains, the task of capacity planning becomes a daunting task. On the one hand, the time-sensitive nature and perishability of products within cold chains require prompt transportation and storage to mitigate spoilage and maintain product quality (Gormley, Brennan, and Butler 2000). This means the insufficient capacity of cold chains could cause significant economic losses and even lead to legal ramifications, product recalls, and reputational damage (Tsang et al. 2018; Nerbovig 2017). Furthermore, in contrast to the relatively rapid capacity adjustments in conventional supply chains, the process of modifying capacity within cold chains is arduous and time-intensive. This stems from the extensive investments required for advanced refrigeration and dehumidification systems required for storing and transporting temperature-sensitive goods (e.g., Wang and Zhao 2021; Dai et al. 2020). On the other hand, excessive under-utilization of capacity within cold chains is expensive, given the substantial investments necessary for these specialized systems.

In summary, recognizing the significant consequences of over-utilization and under-utilization in cold chains, accurate demand forecasting is vital in optimizing operations and capacity planning within these chains. Moreover, accurate demand forecasting ensures the timely transportation and storage of perishable and time-sensitive items in cold chains, thereby preventing wastage and upholding product quality. In other words, demand forecasting for highly constrained operations requires advanced techniques. This is because traditional forecasting methods, often relying on regression models adjusted with historical data, frequently deliver less accurate forecasts than newer approaches like AI and ML. AI/ML models excel in identifying complex non-linear patterns and exploring relationships with seemingly unrelated external data sources, resulting in improved decision-making across various sectors (Hastie et al. 2009). Adopting these advanced algorithms has successfully enhanced decision quality in diverse managerial contexts (Brynjolfsson and McAfee 2014; LeCun, Bengio, and Hinton 2015). Hence, state-of-the-art Artificial Intelligence (AI) and Machine Learning (ML) forecasting methods can effectively address the unique complexities of the cold chain. By applying AI/ML techniques, the primary objective is to enhance demand forecasting accuracy, optimize capacity planning, and improve overall cold chain operations.

Hence, we formulate our research question as follows:

- **Research Question**: *How can AI/ML forecasting methodologies be leveraged for cold chain capacity planning?*

To address the research question, we develop a novel forecasting framework meticulously designed for capacity planning in cold chains, which is necessary to ensure the quality of temperature-sensitive food products. The foundation of our study lies in a comprehensive review of existing literature on food cold chain management, demand forecasting methods, and qualitative interviews with key stakeholders within a real case study.

In our forecasting approach, we utilize two distinct modeling techniques: Seasonal Autoregressive Integrated Moving Average (SARIMA) and Facebook Prophet (FP). We selected SARIMA due to its well-established power and widespread application among both researchers and practitioners. SARIMA models offer an interpretability advantage and are based on established statistical principles. Their capability to handle time series data characterized by clear seasonal patterns makes them particularly suitable for the cold chain context. Conversely, AI/ML models, such as Facebook Prophet (FP), excel in detecting intricate non-linear patterns and uncovering relationships with seemingly unrelated external data sources. This proficiency results in enhanced decision-making across various sectors (Hastie et al. 2009). To measure the accuracy of our models, we utilize the Mean Absolute Percentage Error (MAPE) performance metric. The insights of our forecasting models are invaluable, enabling us to propose the creation of additional freezer capacity at the selected site of the case study. Additionally, we identify underutilized space in the cooler segments that could potentially be utilized to increase freezer capacity, thereby optimizing resource utilization.

As a real case study, we collaborate with Americold, the world's second-largest refrigerated logistics service provider (LSP), the 2021 IARW Global Top 25[1] list by the Global Cold Chain Alliance. Americold specializes in the ownership, operation, acquisition, and development of temperature-controlled warehouses, managing a vast network of 250 facilities across North America, Europe, Asia-Pacific, and South America, with an impressive overall capacity of approximately 1.5 billion cubic feet and a revenue of $2.7 billion, as reported in the Form 10-K of Americold Realty Trust in 2021[2]. By selecting a specific warehousing site as our minimum viable product, we extract extensive inventory data for every customer over a 3.5-year period. Leveraging this data, we propose segmentation criteria based on customer *inventory size* and *variability* (operationalized by the Coefficient of Variation for customer demand), enabling us to develop accurate forecasting models tailored to each segment.

Overall, our findings highlight the significance of AI/ML techniques in cold chains to manage temperature-sensitive food supply chains effectively. In line with the core idea that there is not a single universal algorithm or model that excels above all others (commonly referred to as the "no free lunch" principle), the results of this study also demonstrate that certain algorithms perform better than others based on the characteristics of the SKUs and customer segmentation. The managerial implications provided by the framework provide valuable guidance that can be beneficial for making decisions about selecting forecasting models and planning for capacity.

The subsequent sections of this paper are structured as follows. We review the related literature concerning cold chains and the integration of novel technologies like AI/ML models within this domain. Following this, we elaborate on the methodology employed and detail the procedure for data collection. Finally, we end this paper by proposing the framework and discussing the results.

2. Literature Review

In this section, we examine the relevant literature. First, we review the papers related to cold chains and emphasize the attributes that differentiate cold chains from other supply chains. Next, we discuss the adoption of technologies in cold chains and dynamic capacity planning. Following that, we analyze studies closely related papers to our research that employ AI/ML models in the context of cold chain or fresh food supply chains. The review of these sections will help us to pinpoint the research gap and position the current study.

[1]https://www.coldchainconference.org

[2]https://ir.americold.com/financials/sec-filings/default.aspx

2.1. *Cold Chains*

A cold chain refers to a systematically engineered process in which refrigerated or temperature-sensitive products are stored, transported, distributed, and sold in controlled low-temperature environments to ensure product quality and safety (Mohsin and Yellampalli 2017; Cang and Wang 2021; Meng et al. 2022). In the past decade, there has been significant attention to cold chains due to their expanding range of applications (Zheng et al. 2021). Cold chains have unique characteristics, one of which pertains to the types of items they handle. Typically, these supply chains deal with perishable and temperature-sensitive items like vegetables, fruits, fresh meat, fish, shrimp, and other raw foods (Meng et al. 2022). The quality of these products is a critical factor and a fundamental criterion for consumers' choices (Cang and Wang 2021). This means ensuring the preservation of freshness and product quality necessitates tighter delivery timeframes and controlled storage conditions, which lead to higher end-product quality and reduced losses due to spoilage (Soto-Silva et al. 2016; Dabbene, Gay, and Sacco 2008). Additionally, cold chains face the challenge of long and often unpredictable transportation times due to the considerable distance between the production source and the target market (Behzadi et al. 2018; Cai et al. 2013). This extended lead time makes fresh products susceptible to decay and deterioration. Moreover, end consumers are highly sensitive to both the retail price and the freshness of the products, leading to uncertain market demand (Cang and Wang 2021; Cai et al. 2013). Hence, these challenges underscore the critical role that cold chains play in preserving product quality and pricing, highlighting their complexity compared to other supply chains (Cai et al. 2013).

Numerous logistical models related to cold chains have been extensively examined in existing literature (e.g., Cai et al. 2013; Khodaee, Kayvanfar, and Haji 2022; Zheng et al. 2021; Jedermann et al. 2014). For instance, Jedermann et al. (2014) conducted an assessment of technical remedies and implementations to track the shelf-life of perishable commodities in cold-chain food logistics. Their study also focused on devising subsequent chain operations to curtail product wastage throughout the food distribution process.

Cold chains exhibit unique attributes that can lead to elevated CO2 emissions. For instance, these chains involve the preservation and movement of products at low temperatures, typically at or below freezing levels (Saif and Elhedhli 2016). This prerequisite mandates the utilization of refrigerated warehouses and vehicles, which consume significant energy for refrigeration, potentially amplifying the carbon footprint. Within this context, Saif and Elhedhli (2016) introduce a novel mathematical model for devising cold supply chains with environmental concerns in mind. This model accounts not only for CO2 emissions arising from energy consumption but also includes the effects of refrigerant gas leakage. In the context of E-commerce, Rodríguez Garcia et al. (2023) presents a framework for organizing operational costs in retail and warehouse e-fulfillment strategies, enabling brick-and-mortar grocery retailers to evaluate their advantages and disadvantages. Applying the Time-Driven Activity-Based Costing (TDABC) methodology, the framework is informed by insights from two prominent European supermarkets employing these strategies.

Studies have highlighted that product perishability in cold chains plays a critical role in shaping logistics operations and the design of distribution networks Hasani, Zegordi, and Nikbakhsh (2012); Van Kampen and Van Donk (2014); Dolgui et al. (2018); Viet, Behdani, and Bloemhof (2020). This means the inherent complexity of managing the cold chain for fresh fruits surpasses that of other supply chains (Soto-Silva et al. 2016).

These complexities emphasize the importance of efficient management practices and the integration of contemporary decision-making tools (Ahumada and Villalobos 2009; Akkerman, Farahani, and Grunow 2010). In the context of food and grocery logistics, Rodriguez Garcia et al. (2022) propose the building blocks for the value proposition and logistics strategy of grocery pure players. The study demonstrates that including fresh and frozen products in a retailer's portfolio significantly impacts the fulfillment process. Specifically, the authors highlight the effects on dispatch time slots, vehicle types, delivery time windows, and the appropriate packing methods.

In a comprehensive literature, Soto-Silva et al. (2016) show that a majority of the relevant research related to the cold-chain management of fresh fruits primarily focused on the areas such as truck transportation, vehicle routing, production allocation, and the planning and distribution network. Hence, there is a lack of studies focusing on applying AI/ML forecasting techniques for capacity planning in cold chains (Behzadi et al. 2018). This is the first study that not only proposes a new framework but also tests the validity of the framework in a real case study. The proposed framework can help to enhance the efficiency of cold chain operations, enabling accurate capacity allocation and ensuring on-time deliveries, which is particularly vital in industries like food and pharmaceuticals, where product quality and safety are important.

2.2. *Technologies in Cold Chains*

Some studies have investigated the adoption of new and disruptive technologies in cold chains with a particular focus on fresh and agri-food items. For instance, Wu, Fan, and Cao (2023) examines the approaches to integrating blockchain technology into the cold chain associated with the supply of fresh products. They demonstrate that embracing blockchain technology might not always be the most advantageous choice. This is because the outcome is tied to factors such as consumer acceptance of products without blockchain technology, the rate of fresh product deterioration, and the distribution of traceability costs among supply chain participants upon adopting blockchain technology. Li, Lee, and Gharehgozli (2023) examine and evaluate the primary blockchain platforms utilized within cold chains that are related to food supply networks. Through an exhaustive synthesis analysis, they explore the benefits and challenges linked with blockchain technology. The study underscores that blockchain improves visibility at every phase of the cold chain, increases transaction transparency, food safety, and quality, and simultaneously alleviates concerns regarding food fraud and wastage. Additionally, blockchain serves as a digital solution to reducing operational costs and enhancing efficiency within food supply chains (Li, Lee, and Gharehgozli 2023).

Defraeye et al. (2021) examine how Digital Twins (DT) can be useful in managing cold chains. They emphasize that DT provides valuable information to exporters, retailers, and consumers. This information includes specifics on the time left before a shipment's products expire, which can guide decisions about logistics and marketing. These technologies also help in spotting and predicting possible problems in supply chains that might lead to reduced food quality and waste.

Anticipatory shipping (AS) tools have been used in the context of cold chains (Viet, Behdani, and Bloemhof 2020). AS uses historical data to predict future orders and deliver products to nearby distribution centers before customers order, balancing customer demands and cost efficiency (Lee 2017). Viet, Behdani, and Bloemhof (2020) apply AS for cold chains related to agro-food items, addressing product perishability and imbalance issues. By integrating product quality data into analysis and optimizing production and transportation processes, the research shows anticipatory shipping can improve delivery service by up to 35.3% and cut costs by up to 9.3%, as validated through simulations using a Dutch floriculture supplier's data.

Internet of Things (IoT) devices offer valuable assistance in data collection by continuously and accurately gathering real-time information from various sources (Ko et al. 2015). Considering the unique characteristics of cold chains, IoT plays a significant role in simplifying processes such as gathering data, supervising product quality, managing logistics, and enhancing payment efficiency (Ruan and Shi 2016). The IoT devices enable real-time monitoring of temperature, humidity, and other critical parameters, ensuring the preservation of perishable goods and enhancing overall supply chain visibility and efficiency. In this context, Tsang et al. (2021) present an IoT system for planning deliveries with multiple temperature requirements (IoT-MTDPS). This system incorporates a two-stage multi-objective genetic algorithm optimizer (2PM-GAO). By utilizing IoT-MTDPS, the capacity to manage e-commerce orders is improved while ensuring customer contentment at a specified standard. Hence, established and continually expanding technologies such as the Internet of Things (IoT), cloud computing, and Wireless Sensor and Actuator Networks (WSANs) have offered unique opportunities to remotely control and monitor various regulated parameters in cold chain with minimal human intervention (Mohsin and Yellampalli 2017). Feng et al. (2019) employed a WSN network to dynamically monitor various quality parameters such as temperature, humidity, O2, and CO2 to enhance the safety of frozen shellfish in the cold chain supply chain. Lastly, Torres-Sánchez et al. (2020) introduced a real-time monitoring system that employed multiple non-linear regression techniques to predict the shelf life of fruits and vegetables based on factors such as temperature, relative humidity, and gas concentration.

Kinetic analysis, using past data and mathematical models or machine learning tools, is a helpful tool in numerous fields and industries, helping make future plans and strategies and generally forecasting. In the context of cold chains, Roduit et al. (2019) employed an advanced kinetic analysis to forecast the shelf life of cold chain items through the monitoring of time-temperature data using a data-logger. Wang et al. (2015) introduced a monitoring and decision system that leveraged a wireless sensor network (WSN) and an ontology-based knowledge representation approach to evaluate the quality of cold chain products. The proposed approach interprets data like temperature, humidity, and location directly obtained from sensors within a refrigerated truck to diagnose product status and trigger appropriate alerts.

2.3. *Dynamic Capacity Planning*

As noted in the literature, capacity planning plays an important role in improving business performance in industries with high capital investment costs (Uzsoy, Fowler, and Mönch 2018; Geng and Jiang 2009; Al-Shobaki and Mohsen 2008). Similarly, the huge investments required for advanced refrigeration and dehumidification systems in the cold chains industry (Dai et al. 2020; Behzadi et al. 2018), make capacity planning very relevant. Precise demand prediction is crucial for making decisions like capacity planning, labor scheduling, and production planning (Uzsoy, Fowler, and Mönch 2018).

Although studies have highlighted that forecasting is a crucial component of any planning process, a comprehensive capacity planning process requires more than just forecasting (Thomé et al. 2012; Noroozi and Wikner 2017; Grimson and Pyke 2007). This is due to several reasons. First, forecasts are not always entirely accurate, particularly given the frequent occurrence of black swan events (Taleb 2005) that are drastically affecting supply chains currently (Ivanov, Dolgui, and Sokolov 2019; Dolgui, Ivanov, and Sokolov 2018; Dolgui and Ivanov 2021). Second, even without these events, forecasts inherently lack the data points necessary for precise projections. For example, traditional forecasting methods often struggle to capture and handle the random fluctuations in demand effectively Huang, Chang, and Chou (2008). The inaccurate forecasts, in turn, impact the accuracy and efficiency of subsequent production scheduling and capacity planning (Huang, Chang, and Chou 2008). Third, even if a company possesses every data point at the present moment, it would be unrealistic to assume that each data point remains unchanged as inputs into a future forecast. In a similar line, the famous phrase "All models are wrong!" by Box (1976) is a common sentiment and has become even more prevalent with these unexpected events.

Therefore, the primary goal of this paper is to develop a dynamic and comprehensive capacity planning framework designed for cold chains. This process extends the demand forecasting process and includes the capacity to adjust and optimize resources, such as storage, customer segmentation, and labor planning. The provided framework considers various factors such as site-level forecasting, room-level forecasting, and customer segmentation. This comprehensive approach is aimed at achieving capacity planning that is both effective and flexible within the context of cold chain logistics.

2.4. *AI/ML Forecasting in Cold Chains*

To the best of our understanding, a gap exists in utilizing AI/ML models for the capacity planning of cold chains. Previous research has predominantly focused on using AI/ML tools in predicting time and monitoring temperature (Awad, Ndiaye, and Osman 2020; Aung and Chang 2014; Göransson, Jevinger, and Nilsson 2018). Mariappan et al. (2023) use AI/ML techniques to predict shipment times in cold chains. They analyze over 3 million real-world shipments from an e-pharmacy, creating a diverse ensemble of AI/ML models for improved accuracy. Their approach outperforms existing methods and proves effective for forecasting shipment times within the cold chain related to the e-pharmacy supply chain. He and Yin (2021) focus on cold chain logistics in the Chinese market and develop demand prediction models using AI/ML models. They use neural networks and grey prediction to study cold chain logistics demand. The authors show the neural network algorithm has slightly better accuracy, making it the preferred choice for demand forecasting in cold chain logistics management. Cannas et al. (2023) explore the utilization of artificial intelligence (AI) in Operations and Supply Chain Management (OSCM) by analyzing empirical data from 17 AI applications within six Italian companies. The study assesses how AI applications enhance various processes within OSCM and examines the advantages and challenges companies encounter during their adoption.

In the food supply chain, Priyadarshi et al. (2019) show that AI/ML tools, precisely the use of long short-term memory (LSTM) and support vector regression (SVR), for demand forecasting of certain vegetables, leads to improved outcomes. These enhancements are observed in terms of inventory turnover and days of stock coverage, which aid in averting stock shortages and reducing supply chain volatility. Although the results of the study cannot be generalized, they can still be used for forecasting various agricultural products at the retail stage, taking into account the distinct demand characteristics of each item. Haselbeck et al. (2022) compared nine modern AI/ML techniques with three traditional forecasting methods to predict horticultural sales. The results consistently favored AI/ML models, especially XGBoost, which came out on top in 14 out of 15 comparisons. Also, the authors demonstrate that including additional external factors, such as weather and holiday information, as well as meta-features, will improve the model's accuracy. In addition, they examined whether the algorithms could capture the sudden increase in demand of horticultural products during the SARS-CoV-2 pandemic in 2020, and XGBoost outperformed all other models.

To the best our knowledge, there is an existing gap in leveraging AI/ML models for capacity planning in cold chain logistics. This research introduces a novel AI/ML framework and validates its efficacy through a practical case study. The framework is based on a multi-level forecasting approach, which includes site-level, room-level, and customer segmentation. The results highlight that capacity planning in the cold chain is challenging; however, the framework can enable businesses to determine the right inventory levels, optimize resource allocation, ensure timely deliveries, and enhance cost efficiency. This is critical in time and temperature sensitive industries like food and pharmaceuticals, where product quality and safety are essential.

3. Methodology

This section presents the methodology behind the proposed cold chain capacity planning framework. We start by describing the dataset, the implementation of AI/ML for forecasting, the approach behind customer segmentation, and the eventual development of the capacity planning framework. As previously discussed, there is a noticeable gap of research that applies AI/ML techniques to the capacity planning of cold chains. This study aims to bridge this gap by collaborating closely with Americold. Our approach involves an initial step of segmenting customers into four distinct quadrants, which we achieve by utilizing the coefficient of variation (CV) and analyzing the inventory levels occupied by each customer. This segmentation strategy allows us to account for the diverse demand patterns exhibited by different customer segments. Subsequently, we apply a range of AI/ML models to each customer segment, tailoring our predictions to the unique requirements of each segment. Given the significant variations in demand patterns across different customer types, adopting a one-size-fits-all prediction model is not feasible, and thus, our approach ensures that we deploy the best-suited prediction model for each customer segment.

3.1. *Data*

To address the research question and test the proposed methodology, we collaborate with Americold, the world's second-largest refrigerated LSP. The company specializes in the ownership, operation, acquisition, and development of temperature-controlled warehouses, managing a vast network of 250 facilities across North America, Europe, Asia-Pacific, and South America, with an impressive overall capacity of approximately 1.5 billion cubic feet and a revenue of $2.7 billion. The dataset under consideration covers only one site with a total capacity of 73,296 pallet positions. The site contains both cooling chambers, with a total capacity of 13,632 pallet positions, and freezing chambers, with a total capacity of 59,664 pallet positions. The dataset covers the period from 14 June 2019 to 19 February 2023 and includes 385 individual customers.

The comprehensive dataset from Americold, encompassing diverse customer profiles and detailed transactional data over 3.5 years, offers a robust framework for assessing the efficacy of AI/ML models in the context of complex, real-world cold chain logistics, making it highly suitable for the scope of our study.

3.2. *Forecasting Algorithms*

Forecasting is central to many activities within an organization. For instance, organizations across all sectors of industry must engage in capacity planning to efficiently allocate scarce resources and establish a mechanism to measure and track performance in relation to a predefined baseline Taylor and Letham (2018). The algorithms under consideration include SARIMA and Prophet. In the subsequent experiments, we aim to juxtapose the performance of a state-of-the-art AI/ML forecasting algorithm, Prophet, with that of SARIMA, a traditional and well-established approach. This comparison elucidates the relative strengths and practical implications of advanced versus conventional methodologies within the cold chain capacity planning context.

3.2.1. Seasonal Autoregressive Integrated Moving Average

ARIMA(p, d, q) model is described by Equation 1. In this formulation, p corresponds to the order of the autoregressive component, d is the degree of differencing, and q is an order of a moving average component (Box et al. 2015).

$$y'(t) = c + \phi_1 y'(t-1) + ... + \phi_p y'(t-p) + \theta_1 \epsilon(t-1) + ...\theta_q \epsilon(t-q) + \epsilon(t) \quad (1)$$

where $y'(t) = y(t) - y(t-1)$ is the differenced series, the autoregressive component of order p (AR(p)) is described by Equation 2, and the moving average component of order q (MA(q)) is described by Equation 3.

$$y(t) = c + \phi_1 y(t-1) + ... + \phi_p y(t-p) + \epsilon(t) \quad (2)$$

$$y(t) = c + \epsilon(t) + \theta_1 \epsilon(t-1) + ...\theta_q \epsilon(t-q) \quad (3)$$

where $\epsilon(t)$ is a white noise (Hyndman and Athanasopoulos 2018). The ARIMA model can be represented in a more compact way using the backshift operator $By(t) = y(t-1)$. Using the backshift operator, the first-order difference can be transformed into $y'(t) = y(t) - y(t-1) = y(t) - By(t) = (1-B)y(t)$, therefore Equation 1 can be rewritten as Equation 4.

$$(1 - \phi_1 B - ... - \phi_p B^p)(1-B)^d y(t) = c + (1 + \theta_1 B + ... + \theta_q B^q)\epsilon(t) \quad (4)$$

This formulation is critical to define the seasonal extension of ARIMA (SARIMA). SARIMA is formed by adding additional seasonal terms to the ARIMA models, which can be represented as follows: SARIMA $(p, d, q)(P, D, Q)_m$, where m corresponds to the number of observations per year, for example, $m = 52$ in our study because we deal with data aggregated by week (Hyndman and Athanasopoulos 2018). The lowercase notation is for the non-seasonal parts of the model, and the uppercase notation is used for the seasonal parts ones. For example, SARIMA $(1, 1, 1)(1, 1, 1)_{52}$ corresponds to Equation 5.

$$(1 - \phi_1 B)(1 - \Phi_1 B^{52})(1-B)(1-B^{52})y(t) = (1 + \theta_1 B)(1 + \Theta_1 B^{52})\epsilon(t) \quad (5)$$

We selected the SARIMA model using an exhaustive grid search (Jiménez, Lázaro, and Dorronsoro 2008).

3.2.2. *Facebook Prophet*

Prophet is a Bayesian AI/ML algorithm developed by Facebook's Core Data Science team. Prophet aims to decompose time-series data $y(t)$ into three main components: a trend component $g(t)$, a seasonality component $s(t)$, and a holiday effect component $h(t)$ (Taylor and Letham 2018). Our application doesn't use $h(t)$ due to the data aggregation and the fact that the variability in capacity utilization during major holidays can be captured within $s(t)$. Therefore, the Facebook Prophet model is reduced to a decomposable time-series (Harvey and Peters 1990) of the form described by the Equation 6.

$$y(t) = g(t) + s(t) + \epsilon_t \tag{6}$$

where the error term, represented by ϵ_t, stands for any idiosyncratic changes not covered by the model. It is assumed within the model that ϵ_t follows a normal distribution.

This particular configuration bears a resemblance to a generalized additive model (Hastie and Tibshirani 1987), which constitutes a group of regression models where potentially nonlinear smoothers are applied to the regressors. In this context, only time is used as a regressor, but potentially several linear and nonlinear functions of time serve as components. The methodology of modeling seasonality as an additive component echoes the approach employed by exponential smoothing, as detailed by Gardner Jr (1985).

The trend is modeled as a piece-wise constant rate of growth (Equation 7).

$$g(t) = (k + a(t)^\intercal \delta)t + (m + a(t)^\intercal \gamma) \tag{7}$$

where k is the growth rate, m is the offset parameter. Adjustments vector $a(t)$ with the corresponding rates of adjustments δ and γ are responsible for the trend changes and are defined below.

Trend changes are modeled by explicitly defining changepoints where the growth rate is allowed to change. Given S changepoints at times $s_j, j = 1, \ldots, S$. A vector of rate adjustments can be defined using $\delta \in \mathbb{R}^S$, where δ_j is the change in rate that occurs at the moments of time s_j. The rate at t is then considered as the base rate k, plus all of the adjustments up to that point, namely $k + \sum_{j:t>s_j} \delta_j$. This procedure can be represented by defining a vector $a(t) \in 0, 1^S$ according to Equation 8 (Taylor and Letham 2018).

$$a_j(t) = \begin{cases} 1, & \text{if } t \geq s_j, \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

Based on that, the rate at time t can be calculated as $k + a(t)^\intercal \delta$. It is important to note that When the rate k is adjusted, the offset parameter m must also be adjusted. The adjustment at changepoint j is made using the coefficient γ_j that can be calculated according to he Equation 9 (Taylor and Letham 2018).

$$\gamma_j = \left(s_j - m - \sum_{l<j} \gamma_l \right) \left(1 - \frac{k + \sum_{l<j} \delta_l}{k + \sum_{l \leq j} \delta_l} \right) \tag{9}$$

Facebook Prophet relies on the Fourier series (Harvey and Shepard 1993) to construct a seasonal component $s(t)$. Setting P as the regular period we expect the time series to have, the arbitrary smooth periodic effects can be approximated according to Equation 10:

$$s(t) = \sum_{n=1}^{N} \left(a_n \cos \frac{2\pi n t}{P} + b_n \sin \frac{2\pi n t}{P} \right) \tag{10}$$

Fitting $s(t)$ is conducted through estimating the vector of parameters $\beta = [a_1, b_1, ..., a_N, b_N]$ (Taylor and Letham 2018). $P = 52$ in our case because we deal with yearly data aggregated by week.

3.2.3. Validation

In planning capacity for the cold storage company, single-point forecasts are of little utility, and it is essential to make forecasts over a certain horizon H.

Each of H estimates of the future data points is associated with some error. Therefore, it is essential to specify empirical metrics that may serve to measure forecasting accuracy, compare methods and track the performance of the forecasting model overtime. Additionally, such a metric helps to diagnose how error-prone the forecasting procedure is, which in its turn, allows capacity planning team to determine whether the forecast is trustworthy at all and to which extent. Such a metric can be formally defined as the empirical accuracy of a forecast of $h \in (0, H]$ periods ahead of time T (Equation 11).

$$\phi(T, h) = d(\hat{y}(T + h|T), y(T + h)) \tag{11}$$

where $\hat{y}(T + h|T)$ represents a forecast for h steps ahead of T made using historical observations up to time T and $d(\hat{y}(T + h|T), y(T + h))$ is a distance metric used to measure the forecasting accuracy (Taylor and Letham 2018).

We set the forecast horizon to 26 weeks in the following analysis ($h = 26$). As highlighted by De Gooijer and Hyndman (2006), the choice of a distance metric must be problem-specific. De Gooijer and Hyndman (2006) and Hyndman and Koehler (2006) provide a detailed overview of error metrics suitable for forecasting. In our study, we decided to focus on mean absolute percentage error (MAPE) presented in Equation 12 and Root Mean Squared Error (RMSE) presented in Equation 13.

$$MAPE = \frac{100\%}{T} \sum_{t=1}^{T} \frac{y_t - \hat{y}_t}{y_t} \tag{12}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{T} (\hat{y}_t - y_t)^2}{T}} \tag{13}$$

In the field of AI/ML forecasting, mainly when applied to forecasting tasks within supply chain and operations management, MAPE and RMSE stand out as two of the most commonly adopted metrics (Chou et al. 2023; Chuang, Chou, and Oliva 2021; Zhu et al. 2021). MAPE, favored for its interpretability in AI/ML models, does come with certain caveats. Notably, when actual values (y_t) are zero or approaching zero, MAPE can become infinite or severely skewed, a concern highlighted by (Gardner 1990). Additionally, MAPE's inherent structure means it penalizes positive errors more than negative ones, introducing an asymmetry, as pointed out by Makridakis (1993). However, in AI/ML applications where all data points are positive and significantly distant from zero, these limitations become less consequential. As Hyndman and Koehler (2006) concluded, under such conditions, MAPE remains a reliable metric. On the other hand, RMSE is chosen in AI/ML forecasting models due to its alignment with the data scale, making it particularly relevant for inventory management applications (Hopp and Spearman 2011). Yet, it is crucial to recognize RMSE's sensitivity to outliers (Hyndman and Koehler 2006). This fact means that in AI/ML models, where data might have occasional extreme values, RMSE can be disproportionately influenced.

In summary, while both MAPE and RMSE have their respective advantages and challenges, their widespread adoption in AI/ML forecasting applications, especially within the context of supply chain and operations management, underscores their utility and relevance.

3.2.4. Prediction intervals from bootstrapped residuals and capacity breaks estimation

As highlighted by Box (1976) "All models are wrong!". That is why the forecasting errors and the corresponding prediction intervals are central for capacity planning under uncertainty. A prediction interval for forecast $\hat{y}_{T+h|T}$ can be written as $\hat{y}_{T+h|T} \pm c\hat{\sigma}_h$, where $\hat{\sigma}_h$ is an estimate of the standard deviation of the h-step forecast distribution and c is the multiplier to control the width of the interval (Hyndman and Athanasopoulos 2018).

In order to estimate $\hat{\sigma}_h$ without making additional assumptions on the models and the distribution of residuals, we build prediction intervals from bootstrapped residuals. If forecast error is defined as $e_t = y_t - \hat{y}_{t|t+1}$, we can solve it for y_t and rewrite as follows $y_t = \hat{y}_{t|t+1} + e_t$. In this setting, we can simulate the next observation of a time series as follows $y_{T+1} = \hat{y}_{T+1|T} + e_{T+1}$, where $\hat{y}_{T+1|T}$ is the one-step forecast and e_{T+1} is the future error (Thombs and Schucany 1990). Assuming future errors will be close to past errors, e_{T+1} can be sampled from the past residuals. Adding the new simulated observation to our data set, we can repeat the process to obtain $y_{T+h} = \hat{y}_{T+h|T} + e_{T+h}$ and subsequently $\hat{\sigma}_h$. As a result, if y_{T+h} is viewed as a random variable $y_{T+h} = X \sim N(\hat{y}_{T+h|T}, \hat{\sigma}_h)$, its cumulative distribution function (CDF) corresponds to the probability that the site's total capacity or target capacity x will be sufficient $F_X = P(X \leq x)$ (Rohatgi and Saleh 2015). Based on that $1 - F_X = P(X \geq x)$ corresponds to the probability that the site will break capacity. Figure 1 illustrates the intuition behind this estimation. In a practical setting, the normality of residuals has to be tested using the Anderson-Darling (Anderson and Darling 1954) Normality test (or alternative test) prior to futher probabilistic reasoning.

At this stage, it is essential to once again explain the difference between total capacity and target capacity. The total capacity is the theoretical maximum capacity of the site restricted by the total cubic volume available across all the chambers. However, if

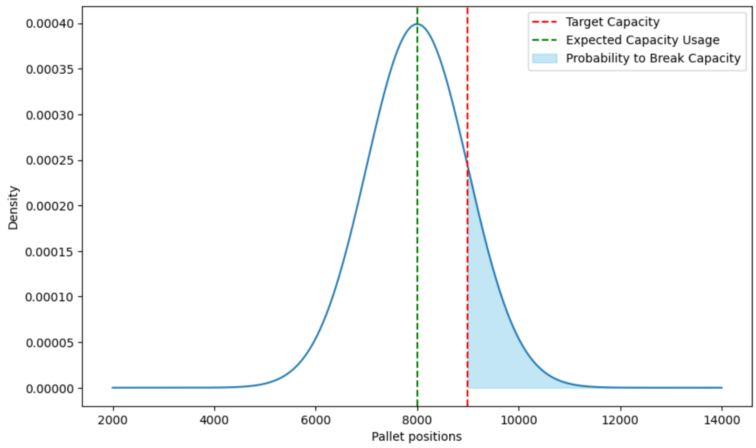

Figure 1.: Probabilistic view of the capacity. The shaded area corresponds to the probability that the site will break capacity x $(1 - F_X = P(X \geq x))$

Figure 1 Alt Text: Graphical representation of a probabilistic view of site capacity, with a shaded area indicating the likelihood of exceeding capacity.

the total capacity is reached, the picking and handling becomes substantially harder to perform, which slows the operation and reduces the overall efficiency. That is why in practical settings, the site uses the notion of target capacity (91% of the theoretical maximum for Americold), which is the optimal upper bound that allows the site to perform without jeopardizing its efficiency.

3.3. *Customer Segmentation*

We developed the Customer Segmentation Matrix inspired by the Growth Share Matrix (also known as BCG matrix). The Growth Share Matrix is a two-by-two matrix created by the Boston Consulting Group in the early 1970s as a tool to help companies with the task of portfolio management (Reeves and Moose 2023). It allows companies to manage different products in their portfolio based on two dimensions: market growth rate and market share (Hax and Majluf 1983; Seeger 1984).

In our model, we replaced these factors with inventory size (which is a primary revenue driver) and variability as our two main criteria for customer segmentation (Figure 2). To determine inventory size, we computed the average inventory during the analysis year. For variability, we utilized the Coefficient of Variation (CV) for outbound demand. The CV is calculated by dividing the standard deviation by the mean, indicating the dispersion around the mean. The lower the dispersion, the more stable and potentially predictable the volume is on any given day. We selected outbound demand for this metric as it is the most labor-intensive activity in a distribution center.

CV is a widely recognized metric in academia (Chou et al. 2023; Chuang, Chou, and Oliva 2021; Zhu et al. 2021) and industry (Hopp and Spearman 2011) for measuring volatility, especially in demand forecasting and inventory management. Its adoption stems from its ability to provide a normalized measure of dispersion, allowing for a more consistent comparison across different scales and units. This fact makes CV particularly valuable in assessing the relative variability of demand patterns, aiding businesses in making more informed inventory decisions.

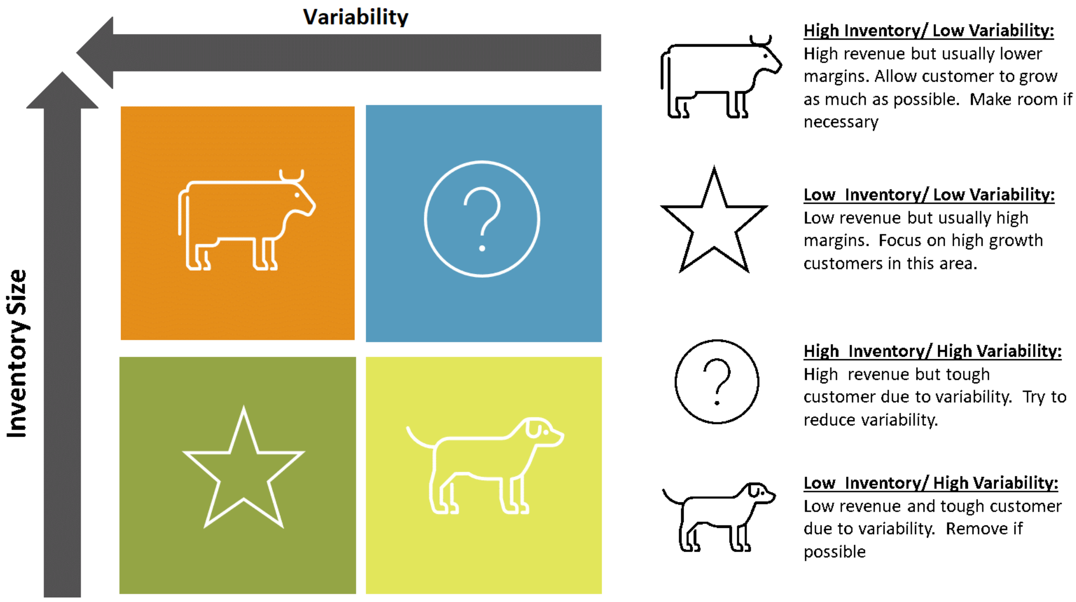

Figure 2.: Customer Segmentation Matrix inspired by the growth–share matrix.

Figure 2 Alt Text: Visual depiction of a Customer Segmentation Matrix, inspired by the growth–share matrix, showcasing different customer groups.

These two dimensions are then used to classify products or business units into four categories:

(1) High inventory / Low variability (Cash Cows). High revenue but usually lower margins. The proposed business strategy is to allow customers to grow as much as possible and reserve additional storage capacity if necessary.
(2) Low inventory / Low variability (Stars). Low revenue but usually high margins. The proposed business strategy is to focus on the customers with the potential to grow and become a “Cash Cow”.
(3) High inventory / High variability (Question Marks). These customers generate high revenue but are hard to deal with due to the high variability and, therefore, unpredictability. The proposed business strategy is to reduce variability through commitment agreements and relationship management.
(4) Low inventory / High variability (Dogs). These customers don’t contribute much to the revenue and are hard to deal with due to high variability. The proposed business strategy is to charge them above-average tariffs and potentially remove them to create space for more promising customers.

3.4. *Capacity Planning Framework*

The proposed capacity planning framework is designed in accordance with the principles of CRISP-DM (Cross-Industry Standard Process for Data Mining), a well-established, structured process model for planning and executing data mining projects (Shearer 2000). The elements of the capacity planning framework for Cold Chain can be connected with various stages of the CRISP-DM model as follows:

(1) Business Understanding: This phase delves deep into understanding the objec-

tives and requirements of the project from a business perspective. In the context of cold chain logistics, this is crucial due to the sensitivity of products and the high costs associated with temperature deviations. Within the capacity planning framework, this phase aligns with the literature review stage, which comprehends the business problem of the cold chain industry's intricacies. Feedback from industry experts further refines the understanding of business requirements and objectives.

(2) Cold chain logistics is data-intensive, with variables like temperature, humidity, and transit times playing pivotal roles. This phase involves the collection and initial exploration of such data. Within the capacity planning framework, it corresponds to the data analysis stage (segmentation and aggregation) and qualitative interviews. These processes ensure a comprehensive understanding of data, identifying data quality issues and exploring data properties and relationships.

(3) Data Preparation: Given the critical nature of cold chain logistics, ensuring data accuracy is paramount. This phase, in the CRISP-DM model, involves cleaning, formatting, and integrating data. Within the capacity planning framework, this aligns with the data validation phase, ensuring the data is accurate and primed for further processing.

(4) Modeling: This stage is vital for predicting potential disruptions or inefficiencies in the cold chain. It encompasses the selection and application of various models to the prepared data. Within the capacity planning framework, this phase is represented by developing and applying forecasting models tailored to cold chain nuances.

(5) Evaluation: Ensuring that models are robust and can handle the complexities of cold chain logistics is crucial. This stage in the CRISP-DM process involves assessing the models against the business objectives and requirements identified earlier. Within the capacity planning framework, this aligns with the steps involving exogenous factors and feedback from cold chain experts.

(6) Deployment: The final phase is where the rubber meets the road. In the context of cold chains, accurate capacity planning can mean the difference between product integrity and spoilage. Within the capacity planning framework, this phase is represented by refining forecasting models to produce actionable business insights. This step applies the model's findings to make informed decisions, ensuring optimal cold chain operations.

Connecting the elements of the capacity planning framework with the stages of the CRISP-DM model ensures a structured and effective approach to data mining projects.

4. Results

The following section illustrates the application of the proposed framework for capacity planning based on the real data provided by Americold. The section demonstrates how customer segmentation and AI/ML forecasts could be streamlined to facilitate critical decision-making regarding capacity planning.

4.1. *Forcasting-based Capacity Planning*

Since historical data is available up to February 19, 2023, and the forecast horizon is set to 26 weeks in the following analysis (h = 26), the forecast is made until August 20, 2023, for all the temperature controls and customer segments. This 26-week horizon has been chosen in line with the capacity planning context. In industries where inventory and storage play a pivotal role, having a medium-term outlook is crucial to ensure optimal utilization of resources and to anticipate any potential challenges or changes in demand (Gambaro et al. 2023). Furthermore, referring to the current managerial practices of Americold, a 26-week forecast aligns well with their biannual review cycles, allowing them to make informed decisions based on the most recent data and trends. We first identify the most accurate AI/ML model based on the time-series cross-validation. Table 1 summarizes the results for the whole site as well as for the freezer and cooler.

The site has a total capacity of 73296 pallet positions and a target capacity of 66846 pallet positions. At this stage, it is essential to once again explain the difference between total capacity and target capacity. The total capacity is the theoretical maximum capacity of the site restricted by the total cubic volume available across all the chambers. However, if the total capacity is reached, the picking and handling becomes substantially harder to perform, which slows the operation and reduces the overall efficiency. That is why in practical settings, the site uses the notion of target capacity (91% of the theoretical maximum), which is the optimal upper bound that allows the site to perform without jeopardizing its efficiency.

The algorithm forecasts that on August 20, 2023, there should be customer demand for 61598 pallet positions ($\hat{y}_{T+26|T}$). Given the $\hat{\sigma}_{26}$ of 7162 pallet positions, calculated using bootstrapped residuals, the probability of breaking the target capacity is 23.2%, and the probability of breaking the total capacity is 5.1%. Taking these numbers into consideration, the capacity planning team has to seriously consider expanding the site or building a new one (Figure 3). The practical limitation of reaching maximum capacity must be underscored, owing to the management of customers via negotiations and pricing strategies. Despite this constraint, the projected customer demand, coupled with a 23.2% probability of breaking the target capacity and a 5.1% chance of surpassing the total capacity, creates a critical situation. It could result in an inability to accommodate new customers or necessitate charging existing customers beyond standard rates to encourage their relocation to alternative sites. Such a scenario emphasizes the urgent need for the capacity planning team to contemplate either the expansion of the current facility or the development of a new one.

The freezer chambers, in total, occupy enough space to allocate 59664 pallet positions, with the target capacity at 53630 pallet positions. It is essential to point out that both the total and target capacity of a freezer increased on November 12, 2022, because the site management decided to convert one of the cooling chambers into a freezing one (Figure 4).

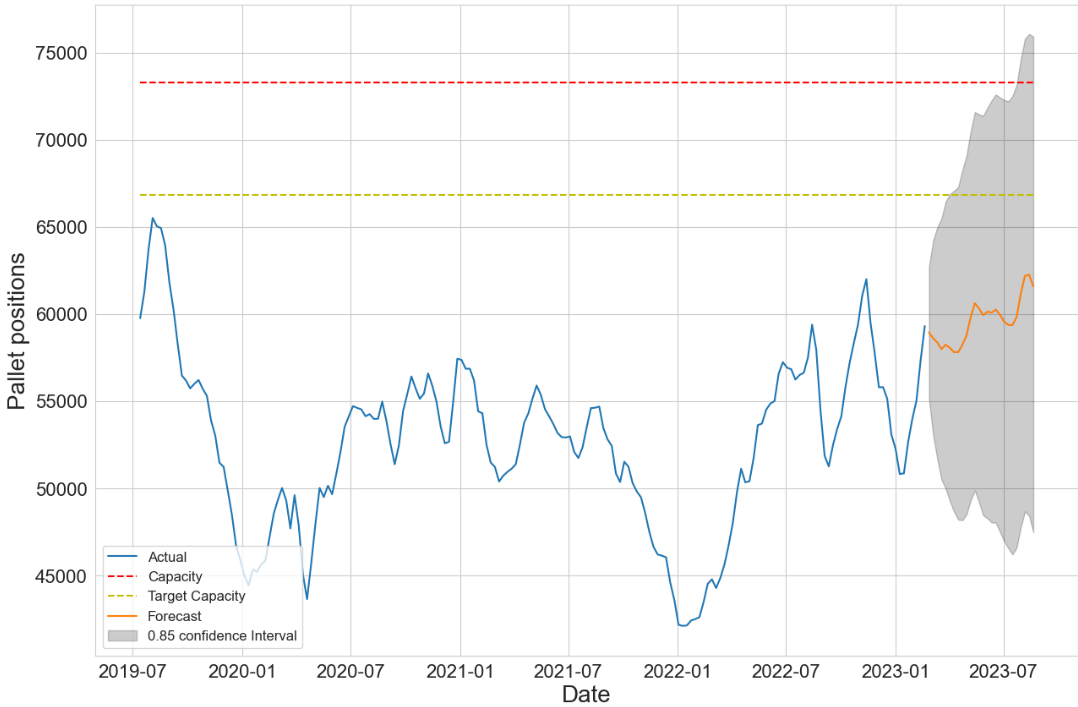

Figure 3.: Forecast at site level performed by SARIMA m(12) (1,0,1) (1,1,1). The figure contains the threshold for total and target capacities measured in pallet positions, historical time-series data, 26-week forecast, and credible interval.

Figure 3 Alt Text: Graphical illustration of a site-level forecast using the SARIMA m(12) (1,0,1)(1,1,1) model. The figure includes lines representing the threshold for total and target capacities in pallet positions, historical time-series data, a 26-week forecast projection, and a shaded credible interval for uncertainty estimation.

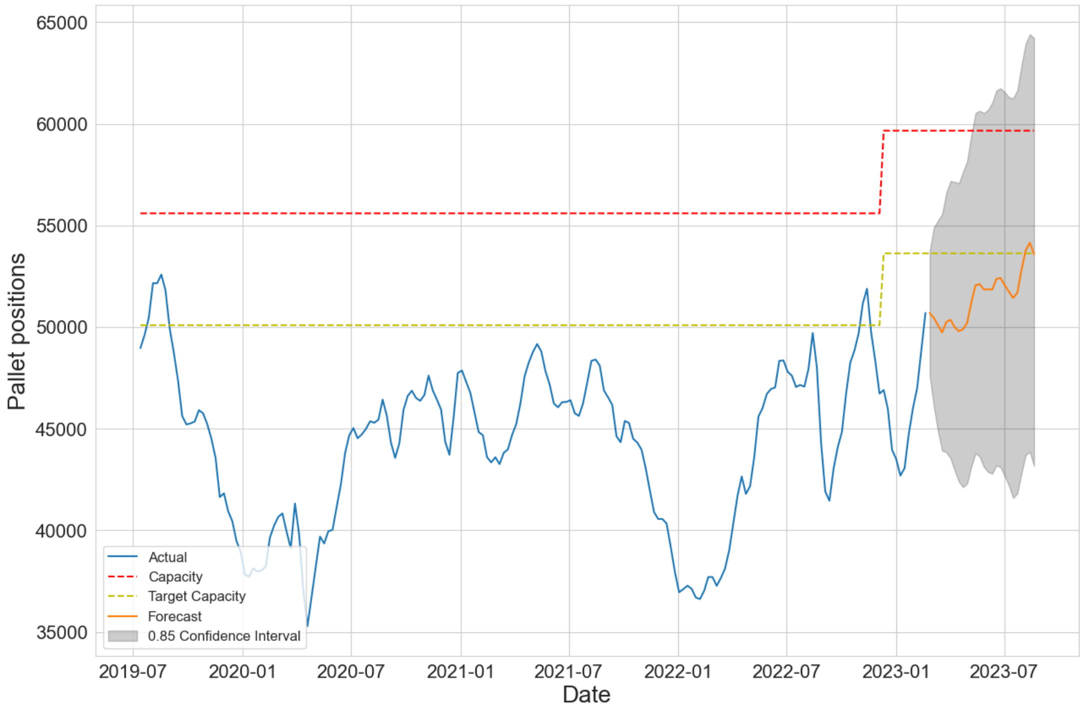

Figure 4.: Forecast for freezer performed by Prophet. The figure contains the threshold for total and target capacities measured in pallet positions, historical time-series data, 26-week forecast, and credible interval.

Figure 4 Alt Text: Chart illustrating a freezer capacity forecast conducted using the Prophet model. Displayed are thresholds for total and target capacities in pallet positions, past time-series data, predictions for the upcoming 26 weeks, and a shaded area indicating the credible interval for forecast uncertainty.

The algorithm forecasts that on August 20, 2023, there should be customer demand for freezing chambers equivalent to 53590 pallet positions. Given the sigma of 5320 pallet positions, calculated using bootstrapped residuals, the probability of breaking the target capacity is 49.7%, and the probability of breaking the total capacity is 12.7%. These projections underscore the significant risk of breaking the target capacity. In the short term, the capacity planning team must contemplate converting one of the cooling chambers to freezing as an immediate remedial measure. However, given the statistical trends, it becomes evident that a long-term solution will necessitate not merely a modification of the current infrastructure but an expansion of the existing site or the construction of an entirely new facility.

The cooling chambers, in total, occupy enough space to allocate 13632 pallet positions, with the target capacity at 13216 pallet positions. Since one of the chambers was converted to a freezer on November 12, 2022, both the total and target capacity of a cooler decreased (Figure 5. The algorithm forecasts that on August 20, 2023, there should be customer demand for cooling chambers equivalent to 8008 pallet positions. Given the sigma of 3684 pallet positions, calculated using bootstrapped residuals, the probability of breaking the target capacity is 7.9%, and the probability of breaking the total capacity is 6.3%. Given the relatively modest risk of breaching capacity thresholds for cooling chambers, the data reinforces the need to further adapt existing facilities. Specifically, the conversion of one of the cooling chambers to freezing appears justified and strategically aligned with projected demands.

Even though the analysis indicates a need to expand the site, it's essential to recognize that the planning and construction process might extend over a year. This timeframe emphasizes the urgency for immediate action. With the site hosting 385 individual customers, a one-size-fits-all approach is inappropriate since not all customers equally impact the total revenue. Considering the capacity constraints, it becomes crucial to further apply the capacity planning framework to enable the segmentation of customers based on various criteria and facilitate decisive action. Decisions must be made regarding which customers should be prioritized with reserved capacity and which may need to be relocated to different sites, charged above the prevailing tariff, or potentially even terminated. This strategic approach will optimize utilization and align customer management with both immediate needs and long-term projections.

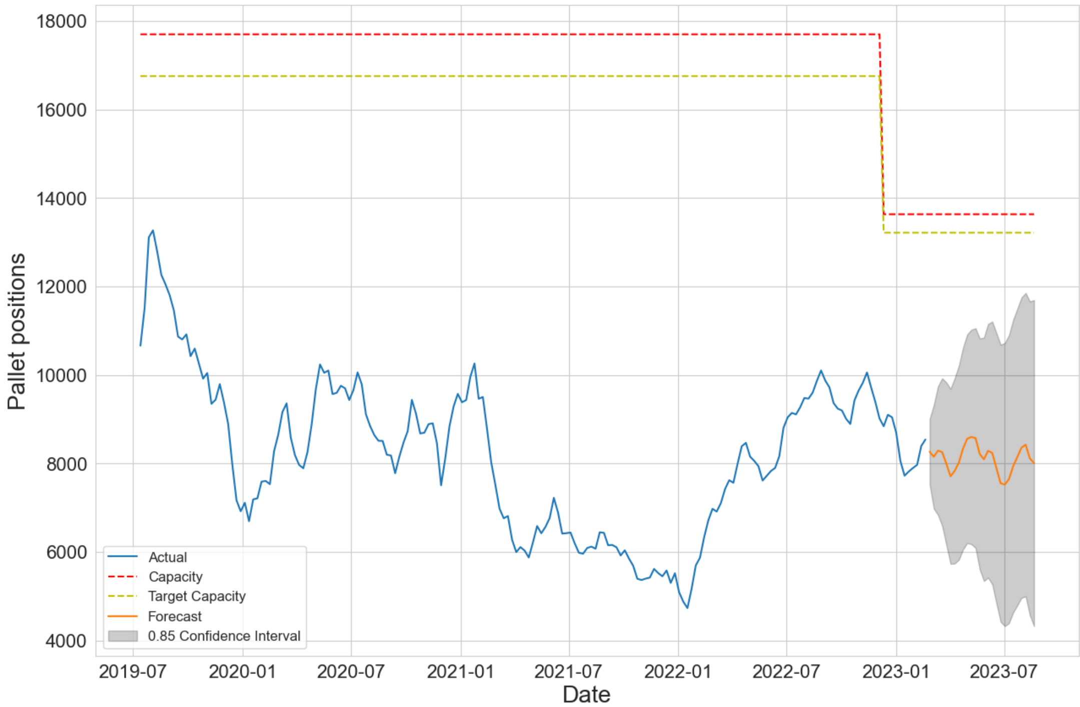

Figure 5.: Forecast for cooler performed by SARIMA m(52) (0,1,0)(1,0,0). The figure contains the threshold for totaland target capacities measured in pallet positions, historical time-series data, 26-week forecast, and credible interval.

Figure 5 Alt Text: Graph demonstrating a cooler capacity forecast using the SARIMA m(52) (0,1,0)(1,0,0) model. It features lines for the threshold of total and target capacities measured in pallet positions, previous time-series data, a 26-week forecast path, and a credible interval shaded to denote forecast reliability.

4.2. *Customer Segmentation*

After setting up the local database and conducting the data cleaning procedures, we were able to identify 385 individual customers who possessed stock at the site during the period covered by the dataset. This posed a significant challenge since generating 385 distinct customer forecasts would result in significant forecasting inaccuracies and render the outputs incomprehensible.

Segmenting was deemed crucial for generating useful forecasts for the organization due to the presence of new customers annually and the overall large number of customers. Using the Customer Segmentation Matrix described in Section 3.3 with the data provided by Americold, we generated the output shown in Figure 6, where each point represents all 385 customers in the data set. The Y-axis displays the average pallets for the analysis year, which was identified as 2022 because the current date was less than 6 months into 2023. In 2022, the largest customer had an average of 5,521 pallets on hand. The X-axis shows the CV for each customer in the dataset, calculated across the entire data set.

To finalize the segmentation for each customer, three additional features were required for the proper scaling of the model. Firstly, we identified the storage temperature for each customer item, which allowed us to categorize all items into Freezer, Cooler, or Dry. Secondly, we identified customers within the multi-vendor consolidation program (MVC).

Multivendor Consolidation (MVC) is a strategic approach in supply chain management where shipments from multiple vendors are consolidated into a single, larger shipment (Glock and Kim 2014). It is particularly beneficial when a large retailer, such as Walmart, sources products from various suppliers in the same region. By employing MVC, companies like Americold assist low-volume customers in achieving cost efficiencies and improved service levels. This consolidation not only reduces transportation costs by transitioning from multiple Less Than Truckload shipments to more cost-effective full truckloads but also ensures fewer delays, leading to consistent and reliable delivery times. Such a strategy is crucial for meeting stringent requirements, like Walmart's on-time in-full (OTIF) standards (Bower 2021). Through MVC, suppliers can also benefit from better transportation rates and services, fostering enhanced relationships and ensuring smoother operations in complex retail environments.

Regardless of their category, customers were given an MVC program flag along with the temperature stored. Lastly, we decided to treat key accounts individually due to their importance to the company (e.g. key customer 1). Americold believes it was important to track these customers individually due to their significance to the site. With these features, we were able to reduce the total number of unique customers from 385 to 12 actionable segmentation groups using this new customer segmentation model (Figure 7).

The conducted segmentation allows us to forecast capacity utilization for each customer individually. which is a game-changer in dealing with capacity shortages. Performing the forecast for each customer segment individually uncovered some compelling trends and patterns. Figures 8 and 9 provide us with a more detailed understanding of the cooler business, which can be used to develop effective business strategies. Upon analyzing Figure 8, it is further evidence that the cooler business has not shown growth over the last few years. The "Low_Inv-Low_CV_Cooler" segment, in particular, has driven the flat to a negative growth trend that persists in the forecast. This lack of growth highlights the potential need for strategic interventions to improve the performance of this customer segment or find new customers that could help fill

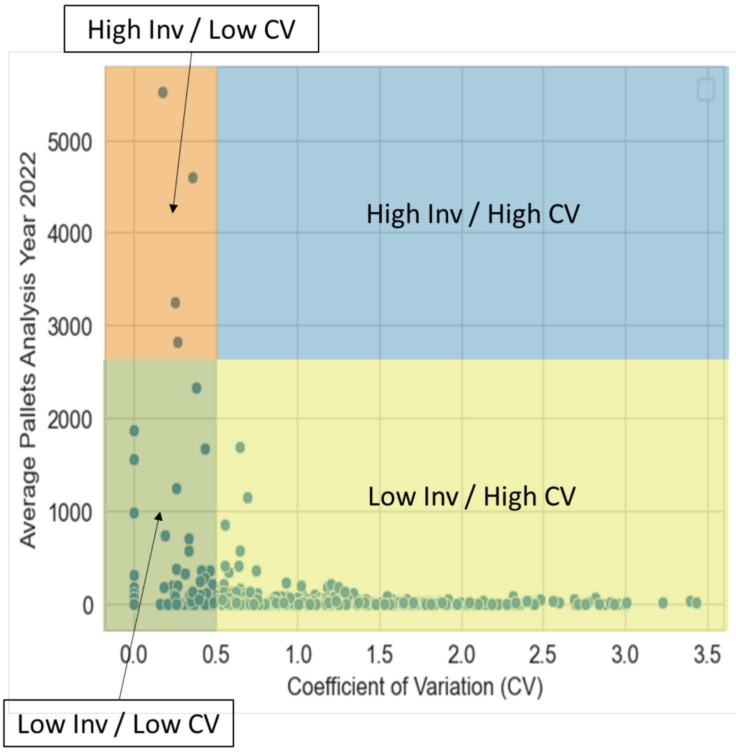

Figure 6.: Scatterplot showing average pallets in 2022 (Y-axis) vs. the Coefficient of Variation (X-axis) for each customer within the data set overlayed with the customer segments created (Low Inv/ Low CV, High Inv / Low CV, Low Inv / High CV, High Inv / High CV).

Figure 6 Alt Text: Scatterplot illustrating the relationship between average pallets in 2022 on the Y-axis and the Coefficient of Variation on the X-axis for each customer in the data set. The plot is segmented into four quadrants to show customer categories: Low Inventory/Low CV, High Inventory/Low CV, Low Inventory/High CV, and High Inventory/High CV, with distinct markers or colors for each segment.

the inventory positions available.

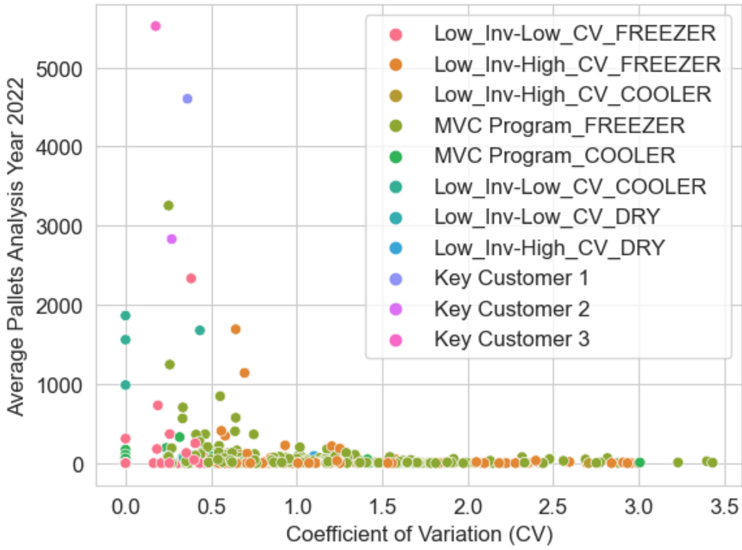

Figure 7.: Final Customer Segmentation. Scatterplot showing average pallets in 2022 (Y-axis) vs. the Coefficient of Variation (X-axis) for each customer within the data set, with each customer color-coded based on their customer segments.

Figure 7 Alt Text: Scatterplot depicting the final customer segmentation, with the average number of pallets in 2022 on the Y-axis and the Coefficient of Variation on the X-axis for each customer. Each data point is color-coded to represent different customer segments, visually distinguishing the groups within the dataset.

In Figure 9, a significant finding was that the "MVC Program_FREEZER" customer segment is projected to experience steady growth over the next 6 months. On the other hand, the remaining customer segments are mostly expected to remain flat or decline. These forecasted trends suggest that targeted interventions might be needed to stimulate growth in the freezer customer segments that are expected to remain flat. Overall, this analysis reveals important insights that can be leveraged to develop effective business strategies.

Dry (ambient) customers account for the smallest portion of pallet positions. Americold does not prioritize ambient customers strategically, and these pallet positions were made available primarily because existing customers required a small amount of ambient space. As of 03/01/2023, there are no more available Dry (ambient) pallet positions, and there are no plans to create more in the future. Hence, don't conduct any further analysis on this category.

Table 1 summarizes the accuracy of forecasting models for different customer segments. At first glance, the average difference of 1.7% in accuracy between the best and second-best models might seem negligible. However, such a difference can have profound implications in the context of supply chain management and inventory optimization.Given the current scenario of ongoing capacity shortages, even minor inaccuracies in forecasting can lead to significant operational challenges. A seemingly small forecasting error can result in stockouts or overstock situations, both of which can have detrimental effects on customer satisfaction and operational costs. In such a tight market, businesses cannot afford to have excess inventory occupying valuable storage space, nor can they risk running out of stock and losing sales opportunities. Furthermore, recent estimates from McKinsey underscore the financial implications of

Table 1.: Accuracy of best-performing forecasting algorithms at the level of the whole site, freezer, cooler, and different customer segments.

Level of aggregation	Best algorithm	Best MAPE *	2nd Best algorithm	2nd Best MAPE *	Delta **
Whole site	SARIMA m(12) (1,0,1) (1,1,1)	5.28% (2.09)	Prophet	5.87% (3.11)	0.59%
Cooler	SARIMA m(52) (0,1,0)(1,0,0)	11.73% (4.21)	SARIMA m(12) (1,0,0)(1,1,0)	11.92% (5.92)	0.19%
Freezer	Prophet	6.11% (1.87)	SARIMA m(52) (0,1,0)(1,0,0)	11.73% (4.21)	5.62%
Key Customer 1	SARIMA m(12) (2,1,1)(0,1,1)	13.17% (6.29)	SARIMA m(52) (0,1,1)(0,1,1)	13.91% (6.15)	0.74%
Key Customer 2	SARIMA m(12) (1,0,2)(0,1,1)	12.84 (5.20)	SARIMA m(52) (1,0,1)(0,1,0)	15.11% (3.26)	2.27%
Key Customer 3	SARIMA m(52) (1,0,1)(0,1,1)	14.55% (3.99)	SARIMA m(12) (2,0,0)(1,1,0)	14.58% (3.42)	0.03%
Low_Inv-High_CV_COOLER	SARIMA m(52) (0,1,0)(1,0,0)	20.62% (4.90)	SARIMA m(12) (2,0,2)(1,0,1)	21.69% (5.56)	1.07%
Low_Inv-High_CV_FREEZER	SARIMA m(52) (1,0,1)(1,1,1)	9.90% (2.29)	SARIMA m(12) (2,0,0)(0,1,1)	12.38% (3.65)	2.48%
Low_Inv-Low_CV_COOLER	SARIMA m(12) (1,0,1)(1,1,0)	17.79% (11.17)	SARIMA m(52) (0,1,0)(1,0,0)	17.80% (7.87)	0.01%
Low_Inv-Low_CV_FREEZER	SARIMA m(52) (1,0,1)(1,0,0)	16.86% (6.32)	SARIMA m(12) (1,0,0)(0,1,1)	19.66% (2.38)	2.8%
MVC Program_COOLER	SARIMA m(12) (1,0,1)(0,1,1)	11.86% (8.93)	Prophet	14.13% (10.57)	2.27%
MVC Program_FREEZER	Prophet	7.65% (2.00)	SARIMA m(52) (0,1,0)(1,0,0)	9.01% (5.15)	1.36%

*Parentheses contain standard deviation of MAPE

**Delta represents the absolute difference in MAPE between 1st and 2nd best forecasting algorithms

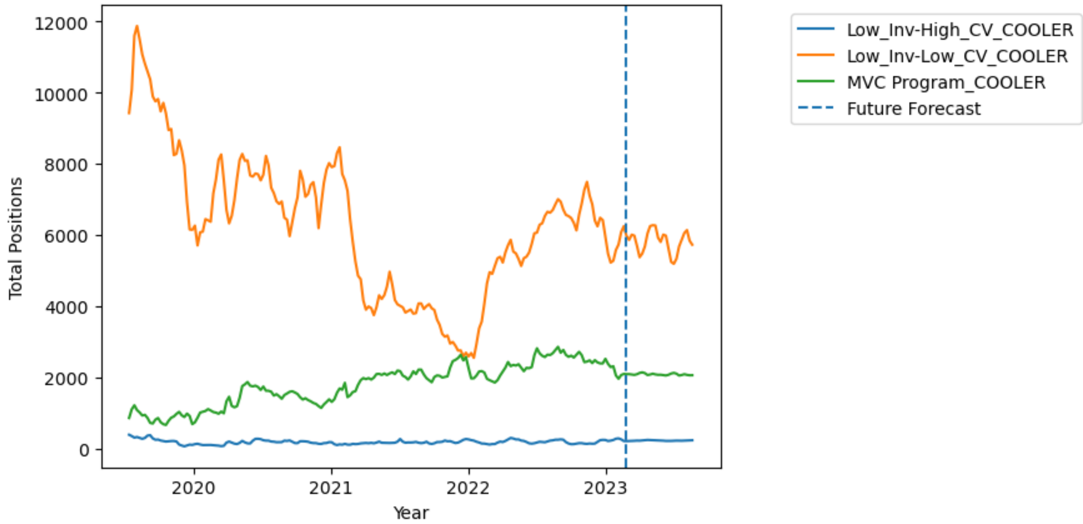

Figure 8.: Time series showing the actual and a 26-week forecast for each customer segment within the cooler. The blue dotted line separates the actual data from the forecasted future.

Figure 8 Alt Text: Time series graph displaying actual data and a 26-week forecast for each customer segment within the cooler. The graph uses a blue dotted line to distinctly separate historical data from the projected forecast, facilitating a clear comparison between actual and predicted values for each segment.

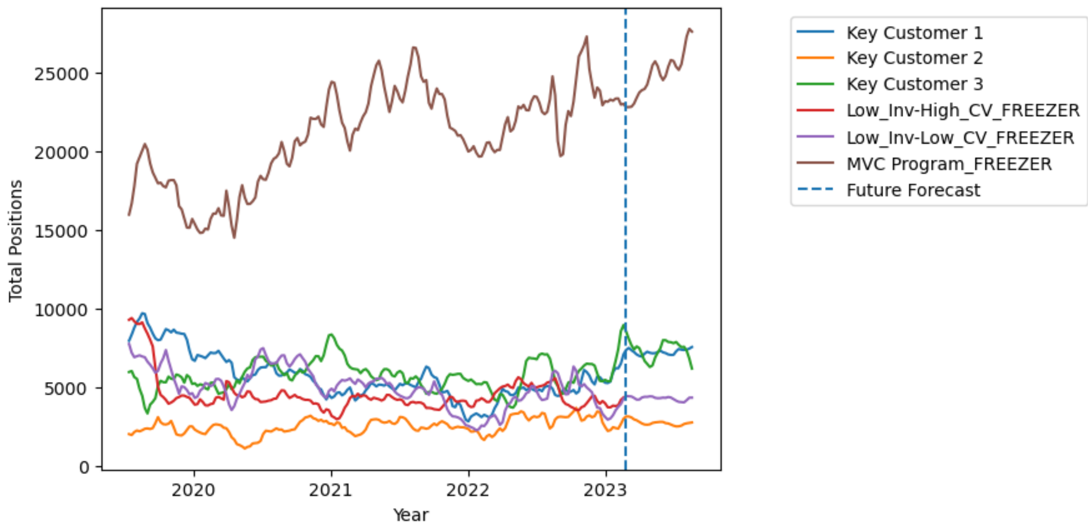

Figure 9.: Time series showing the actual and 26-week forecast for each customer segment within the freezer. The blue dotted line separates the actual data from the forecasted future.

Figure 9 Alt Text: Time series chart illustrating actual data and a 26-week forecast for each customer segment in the freezer. A blue dotted line is used to distinctly demarcate the transition from actual historical data to the forecasted future, providing a visual comparison of actual versus predicted trends for each customer segment.

these discrepancies in forecasting accuracy. For fast-moving items, a mere 5% dip in forecasting accuracy can lead to a 2.5% surge in inventory costs, which ties up capital and increases the costs associated with storage, handling, and potential obsolescence. Moreover, the same dip can result in a 1.5% drop in revenue due to missed sales opportunities and decreased customer trust (Chui, Manyika, and Miremadi 2018).

5. Discussion

5.1. *Framework for capacity planning*

As previously mentioned, capacity planning within the cold chain domain is a challenging task due to the potential economic consequences stemming from inadequate cold chain capacity. The lack of capacity could result in substantial financial losses, legal complications, as well as adverse impacts on reputation and necessitate product recalls (Tsang et al. 2018; Nerbovig 2017). Additionally, distinct from the swift capacity adaptations observed in conventional supply chains, the task of changing capacity within the cold chain context is intricate and time-consuming. Moreover, the financial commitments associated with refrigeration infrastructure expansion are notably substantial. Such expansions are vital within cold chains as refrigeration facilities play a crucial role in preserving the quality and freshness of their commodities across supply networks (Wang and Zhao 2021).

Considering the distinctive attributes inherent in the cold chain, managers are faced with the necessity of formulating both short-term actionable planes and long-term strategic plans—particularly when the impending break of capacity is evident (Figure 10). This study contributes a distinct vantage point primarily focused on short-term actionable plans. Firstly, the study underscores that the strategy of uniformly reducing overall customer orders when approaching maximum capacity is inadvisable. This is due to the inherent variability among customers in terms of revenue generation and product classifications (e.g., freezer, cold storage, etc.). Secondly, an observation emerges that customer segmentation yields multifaceted advantages encompassing revenue management and capacity allocation. Notably, in the context of capacity management, the research reveals the absence of a universally applicable forecasting technique spanning all customer categories.

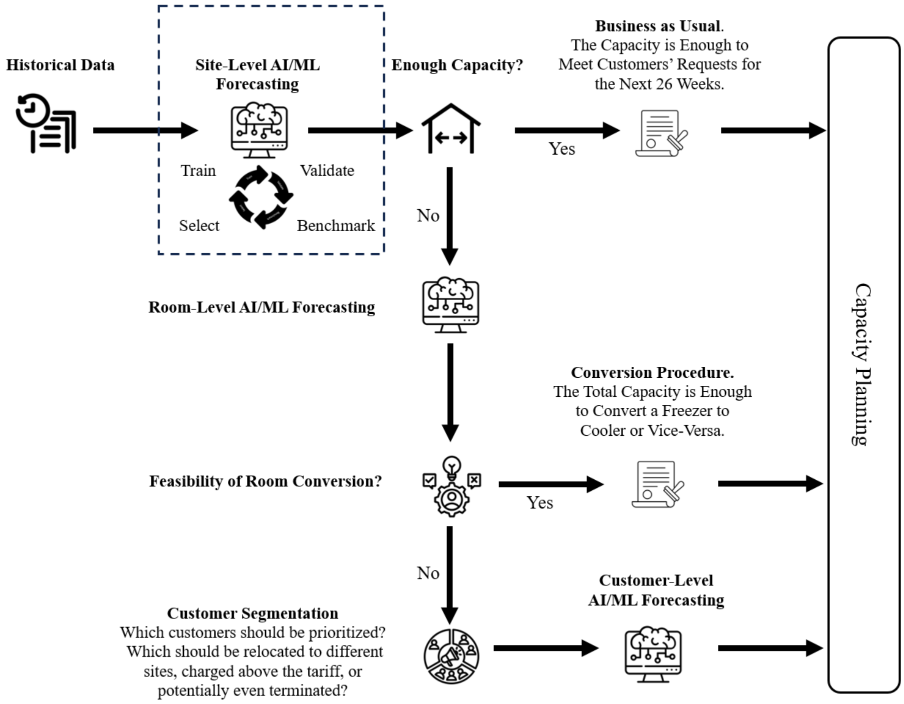

Figure 10.: The proposed framework for Cold Chain Capacity Planning.

Figure 10 Alt Text: Graphic representation of the proposed framework for Cold Chain Capacity Planning, showing the structured process and key components involved in effectively managing and planning capacity within a cold chain system.

This framework serves as a guiding compass for identifying and distinguishing between customers who contribute significantly to the bottom line (in terms of revenue) and those whose contributions are relatively less substantial. This classification holds intrinsic value, as it empowers managers with insights to make informed decisions when it comes to optimizing the allocation of precious capacity resources. In this manner, the capacity that might have been underutilized by less lucrative customers can be strategically repurposed to cater to the demands of the most lucrative customers, depicted in the top-left quadrant of Figure 6.

Furthermore, the framework's utility transcends mere capacity management, extending to a profound impact on pricing strategies. By identifying less lucrative customers, managers can implement pricing adjustments tailored to encourage these customers to migrate from a lower revenue contribution (i.e., the right quadrants depicted in Figure 6), to a more lucrative one, ultimately depicted in the top-left quadrant of Figure 6. This strategic pricing shift, aligned with the principles of the framework, not only addresses the immediate challenge of capacity shortages but also actively helps revenue optimization. Hence, the comprehensive framework serves as an invaluable tool. It not only aids in optimizing capacity allocation but also offers a blueprint to strategically realign pricing structures. By harmonizing these crucial aspects, the framework serves as a multifaceted solution that navigates the complexities unique to the cold chain industry while maximizing both operational efficiency and financial gains.

5.2. *Managerial implications*

Integrating AI/ML methodology in cold chain management has opened new horizons for improving efficiency and responsiveness, including:

- Implementing AI/ML methodologies in cold chain capacity planning offers the potential to revolutionize the industry, enhancing efficiency, responsiveness, and adaptability.
- Using the Customer Segmentation Matrix helps in accurately forecasting and planning different customer segments, addressing the complexity of the diverse customer base.
- By segmenting customers, companies can tailor their approaches to individual needs and trends, thereby reducing forecast inaccuracies.
- Americold's collaboration provides a tangible case study demonstrating how advanced technology can manage a vast network of facilities, signifying practical applications of AI/ML.

The landscape of cold chain management demands not only accurate forecasting but a more holistic, dynamic approach, which entails the following:

- The proposed framework extends beyond mere forecasting, providing a more holistic approach to capacity planning, aligning with organizational needs and market demands.
- A dynamic planning process enables businesses to be more agile and responsive to fluctuations in market trends and unexpected disruptions (e.g., black swan events (Taleb 2005)).
- The research underscores the inherent uncertainty and limitations of forecasting. Recognizing these limitations, managers should adopt a cautious and flexible approach to planning, constantly evaluating and updating models to adapt to

ever-changing market dynamics.
- Given the fast-paced evolution of AI and ML, continuous learning and improvement should be integral to a firm's strategy.

5.3. *Limitations and future research*

It is also essential to emphasize that the results obtained from Americold should be considered with caution when applied to other organizations or industries. Considering the specific dynamics, challenges, and customer base, tailored solutions and adaptations might be required.

Just like many research studies, our study has certain limitations and also suggests interesting directions for future research. Firstly, we've focused our study on just one location within Americold Cold Chain. It would be valuable to examine how well our proposed framework works in different locations with diverse geographical settings. Secondly, we've divided customers into four groups based on their inventory and coefficient of variation (CV), as shown in Figure 6. Expanding the way we categorize customers could be worth exploring, for instance, by also considering the types of products they deal with. This might provide a more comprehensive understanding of how different factors influence customer behavior and could offer more specific insights for both Americold Cold Chain and similar businesses. Lastly, in future research, it could be beneficial to consider the problems related to the planning of limited capacity in a broader theoretical landscape of the shortage economy (Ivanov and Dolgui 2022).

6. Conclusion

In conclusion, this study provides compelling insights into how AI/ML forecasting methodologies can be leveraged for cold chain capacity planning. Collaborating with Americold and utilizing the Customer Segmentation Matrix, we have demonstrated that effective segmentation is crucial for generating practical forecasts in light of the complexity related to handling a large number of customers. While forecasting is fundamental to capacity planning, the challenge lies in acknowledging the inherent inaccuracies and the potential impact of unforeseen black swan events. The capacity planning framework presented in this paper not only addresses the conventional challenges but also prepares cold storage companies to deal with unexpected events, reinforcing the understanding that while all models might be imperfect, they can still be immensely useful.

The contribution of this study extends beyond just the academic sphere and offers a practical, AI-driven approach that can be applied by cold storage companies to overcome capacity shortages. By connecting the dots between AI/ML modeling, customer segmentation, and real-world application, we have proposed a capacity planning framework that is both relevant and adaptable in today's ever-changing supply chain landscape.

Disclosure of interest

The authors report no conflict of interest.

Data availability statement

The authors confirm that the data supporting the findings of this study are available and the subsample of anonymized data can be provided upon reasonable request.

References

Ahumada, Omar, and J Rene Villalobos. 2009. "Application of planning models in the agri-food supply chain: A review." *European journal of Operational research* 196 (1): 1–20.

Akkerman, Renzo, Poorya Farahani, and Martin Grunow. 2010. "Quality, safety and sustainability in food distribution: a review of quantitative operations management approaches and challenges." *OR spectrum* 32: 863–904.

Al-Shobaki, Salman, and Mousa Mohsen. 2008. "Modeling and forecasting of electrical power demands for capacity planning." *Energy Conversion and Management* 49 (11): 3367–3375.

Anderson, Theodore W, and Donald A Darling. 1954. "A test of goodness of fit." *Journal of the American statistical association* 49 (268): 765–769.

Aung, Myo Min, and Yoon Seok Chang. 2014. "Temperature management for the quality assurance of a perishable food supply chain." *Food Control* 40: 198–207.

Awad, Mahmoud, Malick Ndiaye, and Ahmed Osman. 2020. "Vehicle routing in cold food supply chain logistics: a literature review." *The International Journal of Logistics Management* 32 (2): 592–617.

Behzadi, Golnar, Michael Justin O'Sullivan, Tava Lennon Olsen, and Abraham Zhang. 2018. "Agribusiness supply chain risk management: A review of quantitative decision models." *Omega* 79: 21–42.

Bower, Pat. 2021. "Improving OTIF Metrics with Supply Chain & S&OP Best Practices." *The Journal of Business Forecasting* 40 (1): 20–28.

Box, George EP. 1976. "Science and statistics." *Journal of the American Statistical Association* 71 (356): 791–799.

Box, George EP, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. 2015. *Time series analysis: forecasting and control.* John Wiley & Sons.

Brynjolfsson, Erik, and Andrew McAfee. 2014. *The second machine age: Work, progress, and prosperity in a time of brilliant technologies.* WW Norton & Company.

Cai, Xiaoqiang, Jian Chen, Yongbo Xiao, Xiaolin Xu, and Gang Yu. 2013. "Fresh-product supply chain management with logistics outsourcing." *Omega* 41 (4): 752–765.

Cang, Ying-mei, and Du-chun Wang. 2021. "A comparative study on the online shopping willingness of fresh agricultural products between experienced consumers and potential consumers." *Sustainable Computing: Informatics and Systems* 30: 100493.

Cannas, Violetta Giada, Maria Pia Ciano, Mattia Saltalamacchia, and Raffaele Secchi. 2023. "Artificial intelligence in supply chain and operations management: a multiple case study research." *International Journal of Production Research* 1–28.

Chou, Yen-Chun, Howard Hao-Chun Chuang, Ping Chou, and Rogelio Oliva. 2023. "Supervised machine learning for theory building and testing: Opportunities in operations management." *Journal of Operations Management* 69 (4): 643–675.

Chuang, Howard Hao-Chun, Yen-Chun Chou, and Rogelio Oliva. 2021. "Cross-item learning for volatile demand forecasting: An intervention with predictive analytics." *Journal of Operations Management* 67 (7): 828–852.

Chui, Michael, James Manyika, and Mehdi Miremadi. 2018. "McKinsey. Notes from the AI frontier: Applications and value of deep learning." `https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-applications-and-value-of-deep-learning`. Last accessed on 23-06-2023.

Dabbene, F, Paolo Gay, and N Sacco. 2008. "Optimisation of fresh-food supply chains in uncertain environments, Part I: Background and methodology." *Biosystems Engineering* 99 (3): 348–359.

Dai, Jing, Wen Che, Jia Jia Lim, and Yongyi Shou. 2020. "Service innovation of cold chain logistics service providers: A multiple-case study in China." *Industrial Marketing Management* 89: 143–156.

De Gooijer, Jan G, and Rob J Hyndman. 2006. "25 years of time series forecasting." *International journal of forecasting* 22 (3): 443–473.

Defraeye, Thijs, Chandrima Shrivastava, Tarl Berry, Pieter Verboven, Daniel Onwude, Seraina Schudel, Andreas Bühlmann, Paul Cronje, and René M Rossi. 2021. "Digital twins are coming: Will we need them in supply chains of fresh horticultural produce?" *Trends in Food Science & Technology* 109: 245–258.

Dolgui, Alexandre, and Dmitry Ivanov. 2021. "Ripple effect and supply chain disruption management: new trends and research directions." *International Journal of Production Research* 59 (1): 102–109.

Dolgui, Alexandre, Dmitry Ivanov, and Boris Sokolov. 2018. "Ripple effect in the supply chain: an analysis and recent literature." *International Journal of Production Research* 56 (1-2): 414–430.

Dolgui, Alexandre, Manoj Kumar Tiwari, Yerasani Sinjana, Sri Krishna Kumar, and Young-Jun Son. 2018. "Optimising integrated inventory policy for perishable items in a multi-stage supply chain." *International Journal of Production Research* 56 (1-2): 902–925.

Feng, Huanhuan, Jing Chen, Wei Zhou, Vilai Rungsardthong, and Xiaoshuan Zhang. 2019. "Modeling and evaluation on WSN-enabled and knowledge-based HACCP quality control for frozen shellfish cold chain." *Food Control* 98: 348–358.

Gambaro, Anna Maria, Gianluca Fusai, ManMohan S Sodhi, Caterina May, and Chiara Morelli. 2023. "ICU capacity expansion under uncertainty in the early stages of a pandemic." *Production and Operations Management* .

Gardner, Everette S. 1990. "Evaluating forecast performance in an inventory control system." *Management science* 36 (4): 490–499.

Gardner Jr, Everette S. 1985. "Exponential smoothing: The state of the art." *Journal of forecasting* 4 (1): 1–28.

Geng, Na, and Zhibin Jiang. 2009. "A review on strategic capacity planning for the semiconductor manufacturing industry." *International journal of production research* 47 (13): 3639–3655.

Glock, Christoph H, and Taebok Kim. 2014. "Shipment consolidation in a multiple-vendor–single-buyer integrated inventory model." *Computers & Industrial Engineering* 70: 31–42.

Göransson, Malin, Åse Jevinger, and Johan Nilsson. 2018. "Shelf-life variations in pallet unit loads during perishable food supply chain distribution." *Food Control* 84: 552–560.

Gormley, Ronan T, Martine H Brennan, and Francis Butler. 2000. *Upgrading the cold chain for consumer food products.* Teagasc.

Grimson, J Andrew, and David F Pyke. 2007. "Sales and operations planning: an exploratory study and framework." *The International journal of logistics management* .

Harvey, Andrew C, and Simon Peters. 1990. "Estimation procedures for structural time series models." *Journal of forecasting* 9 (2): 89–108.

Harvey, Andrew C, and N Shepard. 1993. "Structural Time Series Models. Handbook of Statistics." *Barking: Elsevier Science Publishers* .

Hasani, Aliakbar, Seyed Hessameddin Zegordi, and Ehsan Nikbakhsh. 2012. "Robust closed-loop supply chain network design for perishable goods in agile manufacturing under uncertainty." *International journal of production research* 50 (16): 4649–4669.

Haselbeck, Florian, Jennifer Killinger, Klaus Menrad, Thomas Hannus, and Dominik G Grimm. 2022. "Machine learning outperforms classical forecasting on horticultural sales predictions." *Machine Learning with Applications* 7: 100239.

Hastie, Trevor, and Robert Tibshirani. 1987. "Generalized additive models: some applications." *Journal of the American Statistical Association* 82 (398): 371–386.

Hastie, Trevor, Robert Tibshirani, Jerome H Friedman, and Jerome H Friedman. 2009. *The*

elements of statistical learning: data mining, inference, and prediction. Vol. 2. Springer.
Hax, Arnoldo C, and Nicolas S Majluf. 1983. "The use of the growth-share matrix in strategic planning." *Interfaces* 13 (1): 46–60.
He, Bo, and Lvjiang Yin. 2021. "Prediction modelling of cold chain logistics demand based on data mining algorithm." *Mathematical Problems in Engineering* 2021: 1–9.
Hopp, Wallace J, and Mark L Spearman. 2011. *Factory physics*. Waveland Press.
Huang, M-G, P-L Chang, and Y-C Chou. 2008. "Demand forecasting and smoothing capacity planning for products with high random demand volatility." *International Journal of Production Research* 46 (12): 3223–3239.
Hyndman, Rob J, and George Athanasopoulos. 2018. *Forecasting: principles and practice*. OTexts.
Hyndman, Rob J, and Anne B Koehler. 2006. "Another look at measures of forecast accuracy." *International journal of forecasting* 22 (4): 679–688.
Ivanov, Dmitry, and Alexandre Dolgui. 2022. "The shortage economy and its implications for supply chain and operations management." *International Journal of Production Research* 60 (24): 7141–7154.
Ivanov, Dmitry, Alexandre Dolgui, and Boris Sokolov. 2019. *Handbook of ripple effects in the supply chain*. Vol. 276. Springer.
Jedermann, Reiner, Mike Nicometo, Ismail Uysal, and Walter Lang. 2014. "Reducing food losses by intelligent food logistics." .
Jiménez, Álvaro Barbero, Jorge López Lázaro, and José R Dorronsoro. 2008. "Finding optimal model parameters by discrete grid search." In *Innovations in Hybrid Intelligent Systems*, 120–127. Springer.
Khan, Amin Ullah, and Yousaf Ali. 2021. "Sustainable supplier selection for the cold supply chain (CSC) in the context of a developing country." *Environment, development and sustainability* 1–30.
Khodaee, Vahid, Vahid Kayvanfar, and Alireza Haji. 2022. "A humanitarian cold supply chain distribution model with equity consideration: The case of COVID-19 vaccine distribution in the European Union." *Decision Analytics Journal* 4: 100126.
Ko, Dasik, Yunsik Kwak, Dojin Choi, and Seokil Song. 2015. "Design of cold chain application framework (CCAF) based on IOT and cloud." In *2015 8th International Conference on u-and e-Service, Science and Technology (UNESST)*, 11–13. IEEE.
Lana, M Moreira, LMM Tijskens, and O Van Kooten. 2005. "Effects of storage temperature and fruit ripening on firmness of fresh cut tomatoes." *Postharvest Biology and Technology* 35 (1): 87–95.
LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. 2015. "Deep learning." *nature* 521 (7553): 436–444.
Lee, CKH. 2017. "A GA-based optimisation model for big data analytics supporting anticipatory shipping in Retail 4.0." *International Journal of Production Research* 55 (2): 593–605.
Li, Kunpeng, Jun-Yeon Lee, and Amir Gharehgozli. 2023. "Blockchain in food supply chains: A literature review and synthesis analysis of platforms, benefits and challenges." *International Journal of Production Research* 61 (11): 3527–3546.
Makridakis, Spyros. 1993. "Accuracy measures: theoretical and practical concerns." *International journal of forecasting* 9 (4): 527–529.
Mariappan, Mahesh Babu, Kanniga Devi, Yegnanarayanan Venkataraman, Ming K Lim, and Panneerselvam Theivendren. 2023. "Using AI and ML to predict shipment times of therapeutics, diagnostics and vaccines in e-pharmacy supply chains during COVID-19 pandemic." *The International Journal of Logistics Management* 34 (2): 390–416.
Meng, Bingbing, Xuelai Zhang, Weisan Hua, Lu Liu, and Kunlin Ma. 2022. "Development and application of phase change material in fresh e-commerce cold chain logistics: A review." *Journal of energy storage* 55: 105373.
Mohsin, Afreen, and Siva S Yellampalli. 2017. "IoT based cold chain logistics monitoring." In *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*, 1971–1974. IEEE.

Negi, Saurav, and Neeraj Anand. 2015. "Cold chain: a weak link in the fruits and vegetables supply chain in India." *IUP Journal of Supply Chain Management* 12 (1): 48.

Nerbovig, A. 2017. "Cold weather causes food spoilage at Lucky's market; replenishment expected Saturday evening." *Billings Gazette* 1.

Noroozi, Sayeh, and Joakim Wikner. 2017. "Sales and operations planning in the process industry: a literature review." *International Journal of Production Economics* 188: 139–155.

Prentice, Barry E, and Ron McLachlin. 2008. "Industry Issue Paper: Refrigerated Food Transport from Canada to Mexico: Cold Chain Challenges." In *Journal of the Transportation Research Forum*, Vol. 47, 118–131.

Priyadarshi, Rahul, Akash Panigrahi, Srikanta Routroy, and Girish Kant Garg. 2019. "Demand forecasting at retail stage for selected vegetables: a performance analysis." *Journal of Modelling in Management* 14 (4): 1042–1063.

Reeves, Martin, and Sandy Moose. 2023. "What Is the Growth Share Matrix?" Available from: `https://www.bcg.com/about/overview/our-history/growth-share-matrix/` Retrieved 30.03.2023.

Rodriguez Garcia, Miguel, Iria Gonzalez Romero, Ángel Ortiz Bas, and J Carlos Prado-Prado. 2022. "E-grocery retailing: from value proposition to logistics strategy." *International Journal of Logistics Research and Applications* 25 (10): 1381–1400.

Rodríguez Garcia, Miguel, Angel Ortiz, José Carlos Prado-Prado, and Andrew Lyons. 2023. "Grocery e-fulfillment costs: comparing retail store and warehouse strategies." *International Journal of Production Management and Engineering* .

Roduit, Bertrand, Charles Albert Luyet, Marco Hartmann, Patrick Folly, Alexandre Sarbach, Alain Dejeaifve, Rowan Dobson, et al. 2019. "Continuous monitoring of shelf lives of materials by application of data loggers with implemented kinetic parameters." *Molecules* 24 (12): 2217.

Rohatgi, Vijay K, and AK Md Ehsanes Saleh. 2015. *An introduction to probability and statistics*. John Wiley & Sons.

Ruan, Junhu, and Yan Shi. 2016. "Monitoring and assessing fruit freshness in IOT-based e-commerce delivery using scenario analysis and interval number approaches." *Information Sciences* 373: 557–570.

Saif, Ahmed, and Samir Elhedhli. 2016. "Cold supply chain design with environmental considerations: A simulation-optimization approach." *European Journal of Operational Research* 251 (1): 274–287.

Seeger, John A. 1984. "Research note and communication. Reversing the images of BCG's growth/share matrix." *Strategic Management Journal* 5 (1): 93–97.

Shearer, Colin. 2000. "The CRISP-DM model: the new blueprint for data mining." *Journal of data warehousing* 5 (4): 13–22.

Soto-Silva, Wladimir E, Esteve Nadal-Roig, Marcela C González-Araya, and Lluis M Pla-Aragones. 2016. "Operational research models applied to the fresh fruit supply chain." *European Journal of Operational Research* 251 (2): 345–355.

Soyer, Ayla, Berna Özalp, Ülkü Dalmış, and Volkan Bilgin. 2010. "Effects of freezing temperature and duration of frozen storage on lipid and protein oxidation in chicken meat." *Food chemistry* 120 (4): 1025–1030.

Taleb, Nassim. 2005. "The black swan: Why don't we learn that we don't learn." *NY: Random House* 1145.

Taylor, Sean J, and Benjamin Letham. 2018. "Forecasting at scale." *The American Statistician* 72 (1): 37–45.

Thombs, Lori A, and William R Schucany. 1990. "Bootstrap prediction intervals for autoregression." *Journal of the American Statistical Association* 85 (410): 486–492.

Thomé, Antônio Márcio Tavares, Luiz Felipe Scavarda, Nicole Suclla Fernandez, and Annibal José Scavarda. 2012. "Sales and operations planning: A research synthesis." *International journal of production economics* 138 (1): 1–13.

Torres-Sánchez, Roque, María Teresa Martínez-Zafra, Noelia Castillejo, Antonio Guillamón-

Frutos, and Francisco Artés-Hernández. 2020. "Real-time monitoring system for shelf life estimation of fruit and vegetables." *Sensors* 20 (7): 1860.

Tsang, Yung Po, King Lun Choy, Chun-Ho Wu, George TS Ho, Cathy HY Lam, and PS Koo. 2018. "An Internet of Things (IoT)-based risk monitoring system for managing cold supply chain risks." *Industrial Management & Data Systems* 118 (7): 1432–1462.

Tsang, Yung Po, Chun-Ho Wu, Hoi Yan Lam, King Lun Choy, and George TS Ho. 2021. "Integrating Internet of Things and multi-temperature delivery planning for perishable food E-commerce logistics: A model and application." *International Journal of Production Research* 59 (5): 1534–1556.

Uzsoy, Reha, John W Fowler, and Lars Mönch. 2018. "A survey of semiconductor supply chain models Part II: demand planning, inventory management, and capacity planning." *International Journal of Production Research* 56 (13): 4546–4564.

Van Kampen, Tim, and Dirk Pieter Van Donk. 2014. "Coping with product variety in the food processing industry: the effect of form postponement." *International journal of production research* 52 (2): 353–367.

Viet, Nguyen Quoc, Behzad Behdani, and Jacqueline Bloemhof. 2020. "Data-driven process redesign: anticipatory shipping in agro-food supply chains." *International Journal of Production Research* 58 (5): 1302–1318.

Wang, Min, and Lindu Zhao. 2021. "Cold chain investment and pricing decisions in a fresh food supply chain." *International Transactions in Operational Research* 28 (2): 1074–1097.

Wang, Yujun, Jianjun Yi, Xiaomin Zhu, Jinlong Luo, Baiyang Ji, et al. 2015. "Developing an ontology-based cold chain logistics monitoring and decision system." *Journal of Sensors* 2015.

Weng, Xin'gang, Jiuyi An, Weifeng Chen, et al. 2015. "Research on Agricultural Products Cold Chain Logistics Safety Supervision System: Mechanism Improvement and Supervision Measures." *Journal of Service Science and Management* 8 (06): 903.

Wu, Xue-Yan, Zhi-Ping Fan, and Bing-Bing Cao. 2023. "An analysis of strategies for adopting blockchain technology in the fresh product supply chain." *International Journal of Production Research* 61 (11): 3717–3734.

Yu, Yunlong, and Tiaojun Xiao. 2021. "Analysis of cold-chain service outsourcing modes in a fresh agri-product supply chain." *Transportation Research Part E: Logistics and Transportation Review* 148: 102264.

Zheng, Feifeng, Yaxin Pang, Yinfeng Xu, and Ming Liu. 2021. "Heuristic algorithms for truck scheduling of cross-docking operations in cold-chain logistics." *International Journal of Production Research* 59 (21): 6579–6600.

Zhu, Xiaodan, Anh Ninh, Hui Zhao, and Zhenming Liu. 2021. "Demand forecasting with supply-chain information and machine learning: Evidence in the pharmaceutical industry." *Production and Operations Management* 30 (9): 3231–3252.