Supply Chain Modeling for Temperature-Sensitive Pharmaceutical Goods

by

Ricardo Guadarrama Arias
Bachelor of Science in Food Process Engineering, ITESO Universidad Jesuita de Guadalajara

and

Abdulrahman S S Gweder
Bachelor of Science in Electrical Engineering, University of Tripoli

Submitted to the Program in Supply Chain Management
in Partial Fulfillment of the Requirements for the Degree of

Master of Applied Science in Supply Chain Management

at the

Massachusetts Institute of Technology

May 2022

Signature of Author: ____

Department of Supply Chain Management
May 6, 2022

Signature of Author: ____

Department of Supply Chain Management
May 6, 2022

Certified by: ____

Dr. Özden Tozanlı Yilmaz
Postdoctoral Associate
Capstone Advisor

Accepted by: ____

Prof. Yossi Sheffi
Director, Center for Transportation and Logistics
Elisha Gray II Professor of Engineering Systems
Professor, Civil and Environmental Engineering

Supply Chain Modeling for Temperature-Sensitive Pharmaceutical Goods

by

Ricardo Guadarrama Arias

and

Abdulrahman S S Gweder

Submitted to the Program in Supply Chain Management

on May 6, 2022 in Partial Fulfillment of the

Requirements for the Degree of Master of Applied Science in Supply Chain Management

ABSTRACT

The pharmaceutical industry relies on cold chain supply infrastructure to preserve the integrity of temperature-sensitive products. Specifically, passive controlled temperature solutions present an increased challenge as the inherent uncertainty of ambient temperatures and process lead time variations in the delivery dramatically increase the risk of inactivating the product. Despite the sponsor company's efforts to test their package solutions in laboratory temperature-control chambers, a lack of visibility exists on the likelihood of success of the packaging solution in real-life conditions. Hence, developing predictive forecasting capabilities for their deliveries across the United States can provide significant financial and operational benefits. This research studies and compares two families of methods of predicting temperature ranges of the goods: statistical methods, Autoregression and ARIMA, and machine learning methods, K-Nearest Neighbor, Support Vector Machines, Random Forest, Quantile Regression, and Long Short-Term Memory Neural Networks. Additionally, one-step and multi-step ahead forecasting techniques were analyzed in all models to determine the best forecasting approach. In addition, the forecasting models were tested on two types of packaging solutions, one for the summer profiles and the second for the winter profile. The results confirm that one-step ahead models outperform multi-step ahead forecasting for long-term horizons when compared by RMSE and MAE. Both statistical and machine learning models accurately predicted training and test set values with relatively lower RMSE. Nonetheless, it was found that testing the models in new external temperature conditions presented contradictory results for predicting the internal temperature, mainly due to the limited data set utilized to train and validate the models. Quantile Regression, on the other hand, successfully predicted the internal temperature of the payload's given new ambient conditions. Therefore, we concluded that a forecasting model can be implemented as part of a predictive risk assessment analysis, considering the impact of variability in both temperature and process lead times for the sponsor company's passive-controlled temperature solutions. These models can be extended for future applications with different configurations of insulator materials, amount of gel packs, and package dimensions.

Capstone Advisor: Dr. Özden Tozanlı Yilmaz
Title: Postdoctoral Associate

ACKNOWLEDGMENTS

We would like to thank our advisor, Özden, for her hard work and mentorship guiding us through this experience and for being an all-round lovely person. We are grateful to Chuck and Chris for their insights, sense of humor, and constant collaboration. Finally, Toby for her through review during the writing process and her kindness.

- Ricardo & Abdulrahman

I would like to thank my capstone partner Abdulrahman for his energy, amazing spirit, and for the moments shared during this project. To my parents, Ernesto and María de Jesús, for your encouragement and support to make all challenges surmountable. To my friends, specifically to Mónica and Ignacio, without your support I would not have fulfilled my dream of coming to MIT. I also want to thank my family and friends for joining me during this journey with their good rapport and will.

Lastly, I want to thank MIT and the MIT Supply Chain Management program for giving me such a wonderful opportunity to work on this project and to continue learning throughout the journey of my master's program. Thank you.

- Ricardo

I would like to thank Richo, my research partner, for all his work making this happen and the friendship we developed. The SCM class in its entirety for their friendship and companionship. I would like to thank my mother, Mariem, for being a living example of being an incredible human, one who inspires me to be my best self. My father, Salah, for his unending support that got me to where I am. I would like to thank Rima and Gadk, my sisters, for always pushing me to be my best. Finally, I'd like to thank my best friend, Weam, his support especially when I doubted myself kept me on the right track.

- Abdulrahman

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The pharmaceutical industry makes up a considerable part of the global economy (Grand View Research, 2021) and has experienced significant growth during the past two decades. Pharma revenues worldwide totaled US $1.27 trillion in 2020 (Mikulic, 2021), contributing 1.5% of the global GDP of US $84.54 trillion (O'Neill, 2021).

The pharmaceutical industry relies on cold chain supply infrastructure; these networks are vital to preserve the integrity of temperature-sensitive products. In terms of value, cold chain-reliant logistics accounted for more than 27% of pharmaceutical logistics cost in 2020 (Grand View Research, 2021).

As many raw materials, active pharmaceutical ingredients, and finished products require the use of cold chain logistics, temperature control throughout the supply chain is crucial to maintain their efficacy. In addition to the pharmaceutical industry's specific supply chain needs, the industry has also experienced considerable growth within its temperature-sensitive portfolio. The share of temperature-sensitive goods overall is only likely to increase as new and more complex therapies are being developed, e.g., treatments derived from living cells that are very sensitive to changes in temperature. The growth is also driven by the increased need for COVID vaccines. Moreover, pharmaceutical logistics and their cold chain operations are some of the most heavily regulated supply chain specialties (Singh, 2005) as safety is a big concern for governments and oversight bodies.

The sponsor company for this project is a multinational pharmaceutical company with a large presence throughout the United States and more broadly, the world. The sponsor company has a large catalog of temperature-sensitive products which constitute a large part of their business, giving this research particular importance.

Due to the large operational area of the United States, and as the sponsor company delivers supplies through numerous lanes throughout the United States, temperature variations between geographical locations must be considered, in addition to weather conditions during transit. In addition, the shipping lanes have distinct stages in the supply chain, where changes in the environmental conditions directly impact the delivered package's inner temperature and its performance.

Currently, the company aims to find the best utilization of financial resources in its supply chains, as they rely on third-party shippers to transport their cold chain goods. The company is considering changing delivery schedules and shipping solutions. As of now, the cold chain delivery success rate, which is the percentage of shipments delivered within the temperature bounds, is above 98%. High success rates raise questions about diminishing returns on investment and whether the delivery operations are implemented in the most cost-effective way. The company is also currently moving towards the standardization of their packaging solutions for their temperature-sensitive goods to reduce complexity and add agility to their warehouse operations, for example, by using one pack-out instead of a summer configuration and winter configuration.

1.1 Problem Statement and Objectives

As the sponsor evolves into a biopharma company, product formulation and delivery have resulted in highly complex supply chains. Due to the overall company strategy, the number of products requiring temperature control throughout the supply chain is vastly increasing.

Furthermore, global regulatory agencies currently have more critical requirements pertaining to the temperature controls needed in products, including its room temperature, to guarantee proper functionality, thus, creating increased challenges to control passive container temperatures along the supply chain.

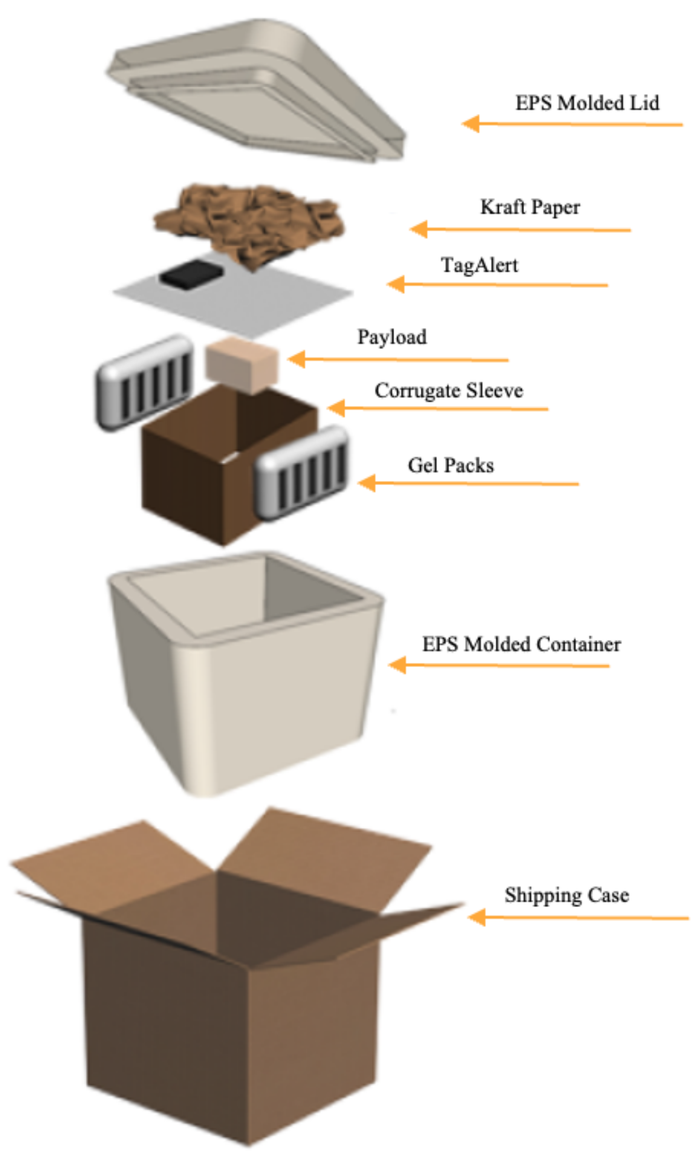
To handle temperature-sensitive products, two families of temperature control solutions are used by the pharmaceutical industry: active systems and passive systems (Zaharia, 2021). This research will focus on the cooling solutions: active cooling and passive cooling. Active cooling uses a heat-reducing mechanism through an electromechanical apparatus to preserve temperature, which is entirely dependent on energy consumption to function. Passive cooling on the other hand, uses a configuration of thermal insulators and cooling packs that protect the payload from excess heat entering from its surroundings. In addition, passive cooling requires no external power source after being packed and shipped. An essential component of passive cooling solutions is cooling gel packs, which are bags filled with a high thermal capacity gel. These gel packs are used as a heat sink to protect temperature-sensitive cargo.

Two main types of thermal conditioning gels are used: summer and winter. The "summer" type is used for temperatures where the outside weather conditions are higher than the target temperature range, and these gel packs are used frozen. In the summer configuration, space between the gels and the product is added to avoid freezing the product. "Winter" gels, on the other hand, are only refrigerated and are used when the ambient temperature conditions are lower than the lower bounds of the target temperature range.

This research will examine passive temperature control solutions utilizing the EPS 11 Box and its different configurations. Key elements comprise the boxes themselves, which act as insulators, the gel packs used to preserve temperature, the TagAlert temperature monitor, and the payload. Figure 1.1 depicts an example configuration for a summer lane.

Figure 1.1

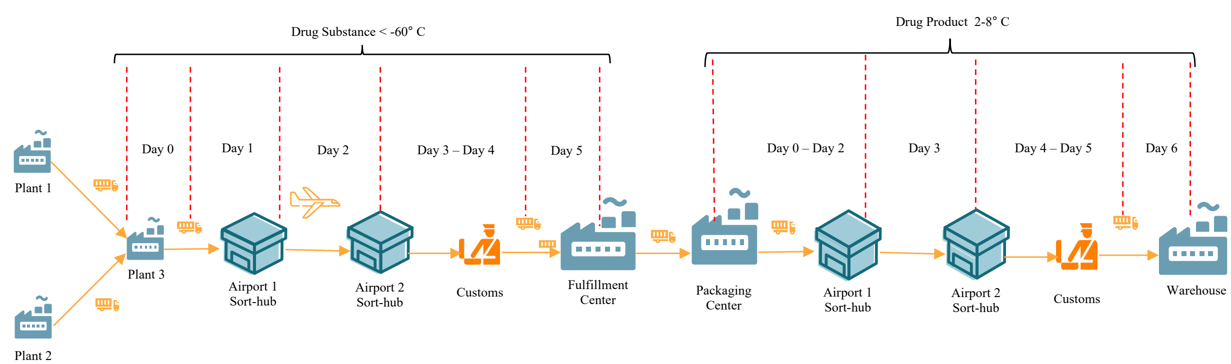
EPS 11 Box Summer Configuration Example



Each drug substance production involves a variety of production steps with many facilities across the globe. In addition, shipping of many raw materials is done below -60°C, by utilizing packaging solutions that use liquid nitrogen and dry ice among others as a means of temperature control. A representative overview of the company's supply chain is illustrated in Figure 1.2, where two primary components are presented, the manufacturing process and the in packaging and shipping process.

Figure 1.2

Typical Representation of Drug Manufacturing Process within the Supply Chain



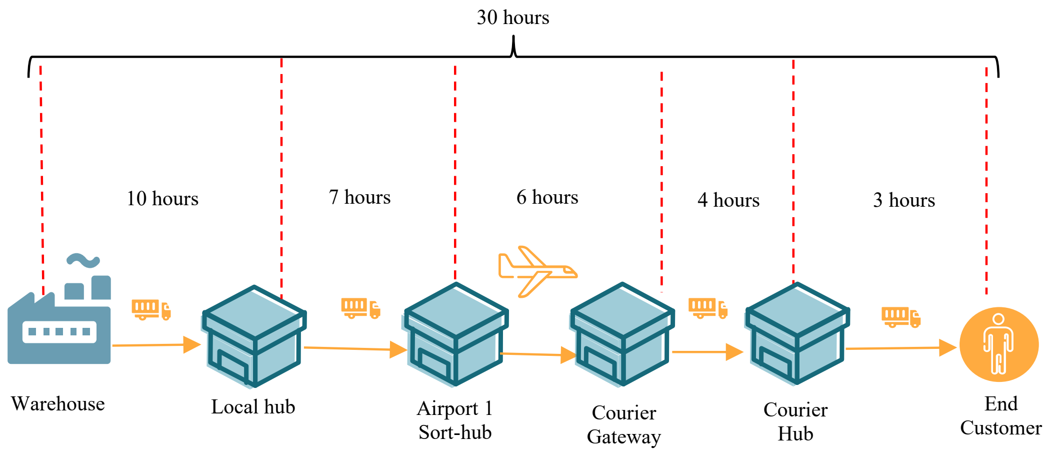
When raw materials are transformed and packed in the final vials or prefilled syringes as drug products, the final temperature requirements are within the range of 2° to 8° C. This is the moment when the delivery process initiates.

Upon reception of the finished and packaged goods to the sponsor company's warehouse, the shipping and the delivery process to the end customer is fulfilled from the sponsor company's warehouse.

The sponsor company's uses a representation of their existing supply chain of approximately 30 hours for delivery to the end customer. The numbers in this representation will be used in this research and are based on conservative estimates of the longest delivery windows in the US. Figure 1.3 depicts this process.

Figure 1.3

Typical Representation of Drug Shipping Process to the End Customer



The stages of the delivery process are as follows:

Step 1. Shipping starts from the product storage in the warehouse, where a transporter picks up the container from the distribution center and transports it to a local hub. (Approximately 10 hours)

Step 2. The product container is loaded in a second truck and delivered to the airport and handled at a sort-hub. (Approximately 7 hours)

Step 3. The product is placed on the tarmac and travels to the airport gateway at the destination. (Approximately 6 hours)

Step 4. The product is loaded in the courier's trucks to be delivered and stored at the courier's hub. (Approximately 4 hours)

Step 5. The drug products are delivered to the end customer. (Approximately 3 hours)

The research and development arm of the sponsor company currently performs internal testing and qualifications using the Parenteral Drug Association (PDA) Qualification Process for all its cold chain shipping solutions. Due to the complexity of the possible temperature control solution configurations, lab testing for new solutions is highly resource-intensive due to several reasons:

1) High-cost investments are required for each test.

2) The potential shipping solutions and configurations are numerousness.

3) Development and qualification testing for packaging solutions require time for approval, which is not aligned with expected lead times.

As such, testing is a time-consuming activity, and it is not practical to create a thermal profile for each case delivered within these containers. As of 2021, the annual volume of passive containers shipped by the company in the U.S. is over 160,000. This creates the need for predictive model that monitor and evaluate temperature profiles of their passive containers. Moreover, assessing the likelihood of success or failure of a proposed packaging solution might increase the percentage of successful deliveries to their destinations. Providing early information on potential failure payloads would provide insights to the operations teams to focus their resources, thus leading to cost savings in the long term.

The objective of this research project is to deliver a predictive model that enables the company to evaluate the performance of their passive containers for drug products along the supply chain and will predict whether the internal temperature stays within the required bounds during transit to the end customer. The geographical scope is based in the United States, using data from sample passive container laboratory tests that contain temperature time series profiles for their summer and winter seasons.

This study will also examine the packaging options used to ship temperature-sensitive goods for the sponsor company and study the current solutions to find a mathematical model that can better predict the success of a specific solution for a given scenario. The main deliverables for the model will include the following:

1) An initial test model for the sample case data to predict the inner container temperature along a typical representation of the company's supply chain.

2) A validated model that will predict the likelihood of solution success against different ambient temperature profiles.

3) A model which will predict the success of the container due to changes in its components and estimate its performance against various ambient temperatures.

4) Recommendations on improving the configuration of cases to certain stress conditions or scenarios.

5) Optimal configuration suggestions for the containers in summer and winter ambient conditions.

2. LITERATURE REVIEW

In this research, different traditional forecasting, machine learning-based forecasting, and simulation techniques will be explored to identify the most suitable approach to model the temperature changes throughout the sponsor company's temperature-controlled supply chain. The literature review section will discuss the current literature and give a summary of approaches other researchers have used to solve similar problems. Additionally, passive packaging solutions for the cold supply chain will be discussed to understand the main factors that impact the inner temperature profile performance.

The literature review will be organized as follows:

Section 2.1 discusses temperature prediction methodologies, machine learning-based approaches, and simulation. This section will also address possible approaches to implement an initial test model. Possible approaches to validate the model will be discussed. The solutions will be discussed and compared to identify existing gaps in the literature and define preliminary characteristics in the model.

Section 2.2 presents the configurations of passive packaging in cold supply chains to predict the performance of the packaging solution due to changes in components across various ambient temperatures.

Finally, Section 2.3, will identify and summarize gaps in previous research and propose a solution, and contributing to the literature.

2.1 Temperature Prediction Models

Estimating the inside temperature of passive packaging is a complicated task due to the number of variables that affect the inner controlled atmosphere. This task can be tackled with two different classes of models: physical or white-box models and data-driven or black-box models (Mustafaraj, Lowry, & Chen, 2011). Physical models are based on energy and mass balance integral–differential equations, where physical properties of the system are known. In contrast, data-driven models are extrapolatory and explanatory by relating to past or future information. Data-driven models investigate statistical information to understand correlation within historical or causal data, estimating patterns when relationships of factors in the system are unknown. Black box models depend entirely on experimental data and therefore can be used when measurements of the system are available. A third category of models is also used that combines both approaches, which are referred to as gray box models, (Berthou, Stabat, Salvazet, & Dominique, 2014). This project will focus on black box modeling approaches due to the lack of thermal property and physical characteristics data of the packaging. The approaches used will range from traditional statistical forecasting to machine learning models.

2.1.1 Forecast Prediction Models

Concerning temperature prediction models, temperature forecasting is usually employed to predict the likelihood of future conditions with many implications and aspects in society. In food supply chain management, predicting temperature along the supply chain is critical to guarantee food quality, as perishable items degrade depending on environmental conditions of storage and transportation facilities, thus impacting the supply chain's performance (Rong, Akkerman, & Grunow, 2011).

Nowadays, forecasting includes many methodologies and techniques that can be applied for temperature predictions. These methodologies can be broadly classified into quantitative and qualitative forecasting; the former is employed when historical data is available, given the aspects of past patterns will continue in the future. Some examples of quantitative forecasting include the naive approach, causal or econometric forecasting methods, time series analysis, and artificial intelligence, among others (Narvekar & Fargose, 2015). This project will focus on quantitative analysis of the relationships across historical and casual variables to predict the inner temperature of the packaging solution across the company's supply chain. This project will therefore study how different forecast methodologies may be applied to increase forecast accuracy for the temperature model of the passive temperature control shipping solution.

The most common quantitative forecasting approaches used in the industry are time series forecasting and regression methods. These models use historical data to predict what is likely to happen in the future. Most time series and regression methods rely on key parameters, such as level, trend, and seasonality, or a combination of them. Depending on the method selected, recent historical data might have a higher impact by adding different weights to the values used. The impact of real-time data is even more crucial for temperature forecasting as the shipping process has distinct stages in the supply chain, where changes in the environmental conditions directly impact the delivered package's inner temperature and its performance (Ben Taieb, Bontempi, Atiya, & Sorjamaa, 2011).

In this project, forecasting methods that do not include trends or use seasonality as the main factor of prediction were discarded, as the sponsor company delivers throughout the United States during the whole year. Some relevant approaches were explored including autoregression, double exponential smoothing and automatic model selection algorithms for

ARIMA, given their proven track of considerable accuracy. (Makridakis, Spiliotis, & Assimakopoulos, 2018)

2.1.2 Traditional vs. Machine Learning Techniques for Forecasting

Machine learning methods have been proposed in the past few decades as an alternative to statistical methods for time series forecasting. As a result, a promising array of their applications in many fields is available and is ever increasing. Machine learning is used in an enormous range of applications, from predicting financial series, to macroeconomic variables, to stock market analysis, to image recognition. The preference is especially pronounced where traditional forecasting methods have certain limitations (Makridakis, Spiliotis, & Assimakopoulos, 2018).

Traditional forecasting methods such as exponential smoothing and moving averages differ vastly from machine learning methods in their capabilities. For example, traditional time series take into consideration a single or a few factors, including trends, levels, and seasonality, and their limitations on the number of predicting factors for creating accurate forecasts. On the other hand, machine learning analytics may incorporate an unlimited source of predictor variables by using learning algorithms to identify underlying drivers and uncover insights from the provided data. Machine learning algorithms are also highly dependent on data availability. Table 2.1 summarizes the main differences between traditional and machine learning approaches.

Table 2.1

Comparison of Traditional and Machine Learning Forecasting for Time Series Models (adopted from Kharfan & Chan (2018))

	Traditional Forecasting	Machine Learning Forecasting
Number of features	Limited to single or a few	Extensive
Data source	Mainly historical data	Multiple
Algorithms	A few single-dimension algorithms	An array of integrated algorithms
Need for manual data manipulation and cleaning	High	Low
Data requirements	Low	High
Technology requirements	Low	High

Due to the advancement of computational processing and capabilities, an increasing amount of research comparing both traditional and machine learning performance and accuracy for time series analysis exists. However, according to the literature, there are mixed outcomes in the performance of both approaches. In their work, Makridakis et al. (2018) studied the eight most common statistical methods for time series analysis, from Naïve to ARIMA, against ten different machine learning models, most prominently using K-Nearest neighbors, CART regression Trees, and Support Vector Regression. It was shown that simple statistical models might outperform ML forecasts when data is limited, failing to learn from each forecasting horizon. In contrast, Smadi & Mjalli (2007) implemented a forward neural network (FFNN) and autoregression time series models (AR) for forecasting the annual air temperature data, where the former gave better forecasts and aid in identifying the dynamics of temperature time series. Another study by Mateo et al. (2013) focused on inner temperature forecasts in buildings, compared Extreme Learning Machines and Multilayer Perceptron (MLP) Artificial Neural Networks against simple machine learning methods based on linear approximations, such as multiple linear regression, robust linear regression, and autoregressive exogenous models. It was found that both MLP models and simple autoregression methods have comparable outcomes. The above suggests the importance of comparing traditional and complex methodologies to benchmark the performance of modeling approaches.

Additional improvements on traditional time series forecasting combined with machine learning methods incorporate the prediction of several future observations of a given sequence of historical observations, referred to as multi-step ahead time series forecasting (Ben Taieb & Hyndman, Boosting multi-step autoregressive forecasts, 2014). In this project, three main strategies of interest related to temperature multi-step ahead forecasting will be applied: Recursive (Rec), Direct (Dir), and Direct Recursive (DirRec) combination, (Ben Taieb, Bontempi, Atiya, & Sorjamaa, 2011). The recursive strategy focuses on creating one model for predicting one time-step and incorporating the predictive value as an input to predict subsequent n time steps to the horizon's end. On the other hand, the Direct method creates distinct forecast models per step ahead, as no approximations of values are used. Lastly, the DirRec method aims to combine Direct and Recursive methods to decrease the accumulated errors as the direct method while analyzing the intricate, complex relationship of historical values from the recursive methods.

Mentink (2018) applied a DirRec methodology to predict the internal package temperature based on the ambient temperature profile during transportation for a pharmaceutical company, applying an optimized quantile linear regression approach on internal and external temperature coefficients. In another study by Suradhaniwar et. al. (2021), one-step vs. multi-step ahead recursive forecasting was compared using agrometeorological time series of temperature and humidity. In their research, SARIMA, Linear Regression, and Support Vector Regression machine learning methods were compared against deep learning approaches, such as Multilayer Perceptron and Recurrent Neural Networks, the latter with better performance on both forecasting strategies. As such, the literature reviewed shows the modeling of the temperature the with promising applications to our package forecast problem.

2.1.3 Deep Learning Methods and ML Applications

Deep learning methods are a category of machine learning approaches that use algorithms that mimic the human brain's learning process. This concept is applied Artificial Neural Networks (ANN), which learn from large amounts of data. ANN differs from other machine learning methods from a modeling perspective as they are trained with more than two non-output layers, i.e., hidden layers, that extract and infer patterns from the data through a black-box modeling approach.

Many applications of deep learning for temperature forecasting and cold supply chain management exist. Xu et al. (2014) used neural networks to categorize the risk of environmental fluctuations in the supply chain management, including temperature control, humidity monitoring, and temperature interruption time, assessing their impact on different points in the cold chain logistics, such as temporary storage, loading and transport. The output of the system was represented by a binary variable (0 meaning risk is low and 1 meaning risk is high). Indicators were then developed at different points in the supply chain to examine the effects of temperature fluctuation. In another study, Tan et. al. (2020) used backpropagation neural network (BP-NN) and Long Short-Term Memory Neural Network (LSTM-NN) as reliable methods to predict glazed frozen squid storage time at different temperatures (Tan et. al., 2020).

According to Makridakis et. al. (2018), the most used neural networks time series forecasting based on the current research are Multi-Layer Perceptron (MLP), Bayesian Neural Network (BNN), Radial Basis Functions (RBF), Generalized Regression Neural Networks (GRNN), Recurrent Neural Network (RNN), and Long Short-Term Memory Neural Network (LSTM- NN). All deep learning techniques aim to perform computations through multi-layer neural networks where data is processed for building a data-driven model. Considered the foundation of deep learning architecture, MLP is a type of feedforward ANN supervised learning approach built by three essential layers: input, output, and hidden sections. BNN extends the MLP method by optimizing the network parameters using Bayesian estimations with Gauss-Newton algorithms. RBF follows the same MLP network, switching the sigmoid activation function for a linear combination of functions radially to the center. GRNN is a variant of RBF with an additional summation layer, with improved learning capabilities and lower processing times. RNN Models uses MLP with feedback connections to consider previous states with current inputs. At the same time, LSTM extends this concept with “gates”, helping to regulate the flow of information through each unit.

This study will implement a LSTM-NN deep learning method given its ability to understand sequential data, interpret time series given time lags of unknown duration, and remember information for long periods of time.

2.1.4 Simulation Model Approach

Simulation is the process of creating a software model of a physical system to predict its performance in the real world, and as the goal of any cold chain setup is the preservation of temperature integrity of goods throughout the supply chain, possible insights can be found through simulation-based approaches. Zwierzycki et al. (2011) at the Institute of Machines and

Motor Vehicles of Poznan University of Technology approached this problem by creating original computer software for heat exchange simulation to forecast temperature changes during transport (Zwierzycki, et al., 2011). The experiment involved cooling the load and monitoring its temperature. In the tests, such parameters as thermal insulation of the body, the original temperature of the load, cooling efficiency of the unit, and ambient temperature were examined to provide insights.

Zwierzycki et al. (2011) also analyzed two cases through their simulation. The decisive factor resulting from the numerical forecast was confirmed by the results of an insurance company's investigations for both cases. This research will use a similar approach to Zwierzycki et al. (2011) to simulate possible scenarios to test the resulting model.

2.2 Internal Configurations on Passive Temperature Control Solutions

Temperature monitoring and control have been rapidly advancing in recent years due to the nature and complexity of sensitive and perishable product logistics. Every aspect of the supply chain, such as production, storage, and transportation, is involved to ensure quality and performance of goods produced and distributed to end customers. (Xu, Zhang, Gong, & Guan, 2014).

Perishable and sensitive products are also a fundamental source of revenue for the cold chain logistics enterprises, typically including biological substances, vaccines, pharmaceutical products and drugs, dairy products, fresh food, and horticultural products, among others. To safe gaurd these revenue stream and mitigate risk, the role of packaging systems in logistics is crucial in protecting product quality and shelf life. Compliance with standards and general rules, as well as concerns about environmental sustainability are also at the forefront when considering passive temperature control solutions (Accorsi, Manzini, & Ferrari, 2013).

Sensitive products in passive temperature control solutions are usually shipped in dry containers, equipped with a wide variety of insulating materials. The choice of the materials in the package depends largely on the product's water content and its resulting interaction with the environmental humidity and temperature (Accorsi, Manzini, & Ferrari, 2013).

Previous research has already defined the primary decision factors to consider when designing the configuration of packaging options (Pillai, 2014). Focusing on the pharmaceutical industry, Figure 2.1 demonstrates the rationale for cold chain packaging choices, guiding drug makers and carriers on the best way these products might be delivered.

Figure 2.1

Decision-making for Packaging Options in the Cold Supply Chain. (Mentink, 2018)

Level 1

- Cost of Packaging

Level 2

- Shipper's concern about the capacity of the packaging system to withstand unexpected ambient temperature

Level 3

- Actual travel distance
- Selection of distribution circuit
- Shipper's concern about the supervision and maintenance during transportation
- Shippers concern on the packaging system's reliability during transportation

Level 4

- Shipper's concern about the temperature profile of the origin and destination location

Level 5

- Size and weight
- Temperature range requirements of the product
- Temperature excursion tolerance level

Level 6

- Product type
- Time to Market

The importance of the decision-making levels differ by industry. Level 5 is key for the pharmaceutical industry, especially since many products are high value and of high temperature sensitivity.

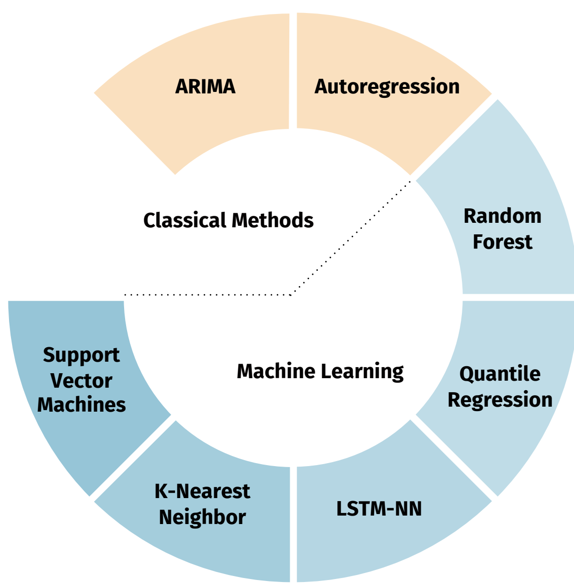
Many factors affect the decision making process of the final configuration of a packaging solution. Most of the crucial factors depend on the product type, temperature range requirements, ambient conditions of the origin and destination, seasonality, and length of the supply chain delivery process. The main elements for passive packaging components that drug makers can adjust are the number and initial temperature of gel packs, the addition of thermal blankets for improved insulation quality, and phase change materials[1] as a buffer for absorbing latent heat that provides stability of the inner temperature conditions. As per the problem statement, further exploration of packaging configuration options is needed, and insights may ultimately lead to savings in overall landed costs. This research will explore and incorporate these configurations in the modeling approaches.

2.3 Conclusion

As there are limited studies in the literature on how passive cold chain solutions can be modeled with pre-qualification lab testing results, this research will apply and compare the following of the previously mentioned approaches (see Figure 2.2).

Figure 2.2

Modeling Methods Used in this Research



[1] A phase change material is a substance which releases or absorbs energy at phase transition to for heating or cooling applications.

Utilizing the given data, machine learning and classical forecasting approaches will be used to create the model, while different scenarios will be applied to test the proposed solution for the supply chain operations of the sponsor company throughout the United States, across different temperature conditions and package configurations.

The methods used in this research were selected based on their capability to incorporate the external temperature as a predictor of the internal passive package solution with minimal effort. Additionally, from a business perspective, models are only useful if they can be interpreted and implemented by the end-user. This careful consideration led this research to implement such models as mentioned earlier.

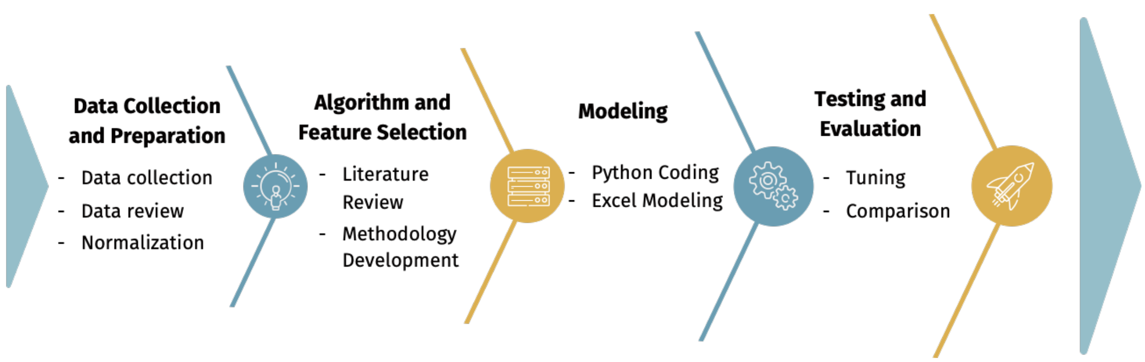
The expected outcome is to provide the sponsor company with simple but powerful methods that assess the impact the external temperature conditions may have on their deliveries, evaluating the conditions where the payload will reach their destinations within their temperature bounds.

3. DATA AND METHODOLOGY

This section will cover the methodology used to model the cold chain operations for the sponsor company. The key outputs to deliver will be predictions of temperature ranges of the goods. This research will start by collecting the data from the sponsor company on their network setup and the packaging solutions used. After the data is collected, the data available will be reviewed, and the preliminary analysis will be carried out. Upon the results of the analysis, the data will be transformed into a shared format between all datasets while also removing any redundancies and correlated data features that could skew the model by means of deletion and averaging. The candidate algorithms and relevant data features will then be selected and modeled. Finally, this project will conclude by comparing the models created through testing and evaluation. Figure 3.1 shows the research workflow.

Figure 3.1

Research Workflow



3.1 Data Collection

Time series temperature data from prequalification testing was collected from the sponsor company for three of their packaging solutions which utilize the EPS 11 box. This data results from lab testing in a temperature-controlled environment of the temperature changes at various stages throughout the supply chain. The sponsor company does the prequalification testing by setting up three test boxes in a temperature-controlled chamber. The temperature in the chamber is then adjusted to simulate the changes in ambient temperature during shipping and delivery as per the standard UPS delivery scenario environments for summer and winter for:

1) UPS facilities,
2) modes of transportation, and
3) any other stops on the way to the final destination.

Table 3.1 shows the temperature steps simulated in the prequalification testing and the corresponding stage in the UPS supply chain network.

Table 3.1

The UPS Environment Temperature Profile Conditions for Winter and Summer

Step	*Mode of Transport*	*Total Time (hours)*	*Winter (°C)*	*Summer (°C)*
Stage in DC	Storage	9.5	22	22
Pick up at DC to Local Hub	Truck	0.5	10	35
At Local Hub	Storage	2.5	5	30
From Local Hub to Sort Hub	Truck	5.0	5	28
At Sort Hub	Storage	1.5	5	28
Stage on Tarmac	Storage	0.25	0	28
In transit to Airport Gateway	Aircraft	2.75	5	28
At Airport Gateway	Storage	1.25	0	28
In transit to local UPS facility	Truck	0.75	-5	28
At destination	Storage	1.25	5	30
With Courier	Truck	2.25	0	30
With Courier	Truck	3	0	35

The EPS 11 box has two configurations; the summer configuration with gels conditioned between -3°C and -10°C, and the winter configuration with gels conditioned at 5°C. The decision between using the summer profile and the winter profile for a package is made according to the temperature at the destination, i.e., if the temperature at the destination is above 42 °F ≈ 5.5 °C, the summer configuration is used. Otherwise, it is packed as per the winter configuration.

Table 3.2 shows an excerpt of the raw data provided. The original datasets reviewed are split into three feature families:

1) Time,
2) The actual thermal data from the thermocouples[2] TC1-TC11, and
3) The step temperature, which mimics the temperatures while the shipment is being delivered.

Table 3.2

Excerpt of the Prequalification Data for -3°C Test Configuration

Step	Time	Product Thermocouples									Chamber Thermocouples	
		Box 1			Box 2			Box 3				
		TC1 (°C)	TC2 (°C)	TC3 (°C)	TC4 (°C)	TC5 (°C)	TC6 (°C)	TC7 (°C)	TC8 (°C)	TC9 (°C)	TC10 (°C)	TC11 (°C)
22	09:25:00	3.5	3.4	3.5	4.1	4.0	3.8	4.6	4.6	4.5	15.6	4.3
22	09:35:00	4.4	4.6	5.4	4.4	2.6	3.8	3.3	1.2	2.4	18.9	19.4
22	09:45:00	3.5	3.9	3.9	3.5	1.6	2.8	2.8	1.1	2.3	19.6	20.5
22	09:55:00	3.7	4.6	4.5	3.8	2.1	3.3	3.3	1.7	3.1	19.8	20.9
22	10:05:00	4.2	5.1	5.0	4.3	2.7	3.8	3.9	2.3	3.7	19.8	21.0

[2] Thermocouples are an electrical device consisting of two dissimilar electrical conductors used to measure temperatures by the flow of electricity

The 11 thermocouple time series are split as follows:

1- Three thermocouples on each of the three test packages. TC1- TC9. The thermocouples are positioned on the vials, the exterior cartons, and the TagAlert systems for each test package, for a total of nine product temperature monitors.
2- Two testing chamber thermocouples, TC10 and TC11, positioned on the top left front edge and the bottom right rear edge of the chamber.

3.2 Data Review

This section will discuss and review the available data. The data review started with transformation of the data and a preliminary analysis. Since the prequalification testing tests three versions of the same solution at the same time in the same environment, the data is of high quality. Upon preliminary analysis, we see that the nine temperature readings from the thermocouples on the boxes are highly correlated, as can be seen from Table 3.3.

Table 3.3

Correlation between TC1-TC11 for the -3°C Configuration

	TC1	*TC2*	*TC3*	*TC4*	*TC5*	*TC6*	*TC7*	*TC8*	*TC9*
TC1	1	-	-	-	-	-	-	-	-
TC2	0.998	1	-	-	-	-	-	-	-
TC3	0.9957	0.9994	1	-	-	-	-	-	-
TC4	0.9981	0.9925	0.9885	1	-	-	-	-	-
TC5	0.9882	0.9773	0.9708	0.9955	1	-	-	-	-
TC6	0.9997	0.9972	0.9947	0.9986	0.9899	1	-	-	-
TC7	0.9987	0.9938	0.9899	0.9997	0.9938	0.9989	1	-	-
TC8	0.9958	0.9888	0.9837	0.9987	0.9959	0.9964	0.9989	1	-
TC9	0.9908	0.9969	0.9987	0.9815	0.9611	0.9899	0.9833	0.9762	1

The chamber thermocouples, TC10 and TC11, are also very close to each other, which is expected as the temperature changes are not instantaneous in the chamber. The average difference between TC10 and TC11 is 1.2°C for each step, with the maximum difference being 1.6°C, and the minimum being 0.1°C

3.3 Data Preparation

The datasets for the three configurations were transformed and normalized. Initially, the time and date data were normalized into a time series in hours, starting at 0 with a step of approximately 0.167 hours, representing 10-minute increments.

The reading from the thermocouples placed on the vails containing the product will be the input considered, i.e., TC1, TC4, and TC7. These thermocouples were chosen as they are the closest to the real temperature of the product. These readings are all highly correlated, as can be seen from Table 3.4.

Table 3.4

Correlation between TC1, TC4, and TC7 for the -3°C Configuration

	TC 1 (°C) On Vial	**TC 4 (°C) On Vial**	**TC 7 (°C) On Vial**
TC 1 (°C) On Vial	1	-	-
TC 4 (°C) On Vial	0.9974	1	-
TC 7 (°C) On Vial	0.9994	0.9978	1

Since both thermocouple TC10 and TC11 readings in the dataset move in tandem, the temperature of the testing chamber will be taken as an average between them, i.e., from the top left front edge of the chamber and the bottom right rear edge.

3.4 Statistical Methods

This section describes the two types of traditional statistical algorithms selected to compare the outcome of machine learning methods: Linear Regression and Autoregressive Integrated Moving Average (ARIMA).

3.4.1 Linear Regression – Autoregressive models

Linear regression is a method that aims to model a linear relationship between two types of variables: the input variable x is known as the explanatory variable, and the output variable y is referred to as the dependent variable. If there are two or more explanatory variables, then the method is referred to as multiple linear regression. The form of the linear relationship is as follows:

$$y = b_0 + b_1 x_{i1} + b_2 x_{i2} + \cdots + b_n x_{in} + \epsilon \tag{3.1}$$

where

y: Predicted or dependent variable

x: Feature or explanatory variables

b: Coefficient for each explanatory variable

ϵ: Modeling error, noise, and disturbances for forecasting model

Frequently in time series data, observations at the current and previous time steps impact the values in the steps ahead. This method, called autoregression or regression of self, assumes that input variables are taken as observations at previous time steps, called lag variables represented as y_t, y_{t-1}, to y_{t-n}, each lag contains its own autoregressive coefficient b. As an example, predicting the next time step using the observations at the current and previous steps, the regression model takes the following form:

$$y_{t+1} = b_0 + b_1 y_t + b_2 y_{t-1} + b_{\mathrm{n}} y_{t-\mathrm{n}} + \epsilon \tag{3.2}$$

Given that the regression model utilizes data from the same input variable at previous time steps, there is a statistical analysis to measure the correlation between the predicted value and the number of lagged observations of the dependent variable. Correlation is helpful to identify the number of lag variables to use, and the weighting of the previous values impacting the predicted values.

The classical AR model does not take into consideration external factors, only lags from the input variable. Given these limitations, the AR method cannot evaluate the impact of the ambient temperature on the next predicted inner package temperature. Thus, in this capstone project, AR will be used only as a benchmark for complex models that incorporate external factors.

3.4.1.1 Autoregressive linear models with Exogenous factors (ARX).

A special extension of the simple AR method is incorporating external variables, i.e., variables that are not affected by other variables in the system. An everyday example would be how weather, pesticides, farmer skill, and availability of seeds are independent variables in the process of crop production. In our specific model problem, ambient temperature is a variable that is external to our package delivery but has a direct impact on the inner product temperature. This type of variable is called an exogenous factor, also known as a causal factor.

By using ARX, this capstone project will aim to estimate the relationship of the inner package temperature lags, while assessing the impact of the ambient temperature profiles as an exogenous variable. Preliminary correlation analysis done in Section 4.2 shows that ambient temperature highly correlates with all inner temperature measures from the thermocouples used on the summer and winter data (TC1 to TC9). A model that attempts to link historical past values and ambient temperature is shown below, as applied by Mentik (2018), following a simple regression model format. For forecasting the inside temperature y_t is replaced by $y_{t,l,p}$, given the forecast model depends on the lane (l) and season, and the packaging option used (p) for each step for the lag feature (h) in the horizon (H):

$$y_{t+h,l,p} = y_{t+h-n,l,p} + \beta_{h,p}\left(X_{t+h,l} - y_{t+h-n,l,p}\right) \tag{3.3}$$

Equation (3.3) predicts the inside temperature based on the previous inside temperature steps $y_{t+h-n,l,p}$ plus the impact of the ambient temperature compared to the inner temperature. In this formula, the bigger the difference between the ambient temperature and the inside temperature, the higher the impact. Given the previous inside temperature $y_{t+h-n,l,p}$, a modifier factor β estimates how vulnerable packaging option p is to the ambient temperature conditions $X_{t+h,l}$.

In this capstone project, the approximated ARX model will be used as a forecast model and as a comparison for other classical and machine learning models. Section 3.6.4 will discuss how this model is adapted to include multi-step ahead methods for improved accuracy.

3.4.2 Autoregressive Integrated Moving Average (ARIMA)

ARIMA models are time series methods that incorporate a linear regression of lags from autoregression (p), an integrated part from the differences of raw observations (i), and moving average used to measure the dependency between an observation and residual errors from the moving average model applied to lagged observations (q). The model incorporates a dependent relationship across historical values and lagged observations. The model is then transformed into a stationary time series by removing trend and seasonal factors. A regression is developed considering the error values between actual data and fitted values. The benefits of an ARIMA model are that it can assign appropriate weights for older lags, in contrast to more traditional methods like exponential smoothing which cannot, as they always assign a higher weight to recent values.

Classical ARIMA models do not take into consideration external factors. Given these limitations, ARIMA methods cannot evaluate the impact of the ambient temperature into the next predicted inner package temperature value. In this capstone project, ARIMA will be used only as a benchmark for complex models that incorporate exogenous factors.

3.5 Machine Learning Methods

This section describes the five types of machine learning algorithms selected to compare the outcome of traditional statistical forecasting methods: Random Forests, Support Vector Machines, Neural Networks, and K-Nearest Neighbors.

3.5.1 Regression Trees and Random Forests

Regression trees are a subgroup of a family of nonlinear predictive models, namely prediction trees. Regression Tree (RT) models are unsupervised machine learning algorithms that classify the array of outcome records into smaller regions by creating splits on predictors, i.e., partitioning, leading to more manageable data interactions for analysis. These splits, called prediction trees, represent the recursive partition. Each terminal node represents a partition cell and has attached a simple model that applies in that cell only. The prediction trees create logical rules from the training set that are interpretable. In the case of regression trees, a further regression can be applied to each subset of data to improve forecast accuracy.

Random Forests extends the regression trees by applying many individual decisions trees as an ensemble. The ensemble method integrates the predictions from multiple decision tree algorithms by generating random samples from testing data and using a random subset of the predictors at each run. The outcome aims to improve the performance and accuracy of the forecast compared to the individual decision trees.

3.5.2 Supported Vector Machines (SVM)

Supported Vector Machines (SVM) are supervised learning algorithms used for both classification and regression analysis, (Shmilovici, 2009). SVM aims to partition the time series data in different classes by a line or a hyperplane. By setting this separator and identifying the supporting vectors, which are the closest points from the partition, the algorithm aims to maximize the distance across different classes, known as margin, and to evaluate the fit's quality by minimizing the expected error of the loss function. An optimized hyperplane is created from the training set to separate the data into classes in all SVM applications. SVM can also incorporate kernel functions and hyperparameters to classify nonlinearly distributed data, thus applying a regression method to predict future values.

3.5.3 k-Nearest Neighbors (k-NN)

k-Nearest Neighbors are a non-parametric approximation learning algorithm that is based on the idea that new values can be classified or predicted given the similarities to the training data. The closeness of two samples is defined by a distance function, where the algorithm finds k training examples closest to the feature vector and returns the majority label, in case of classification, or the average label, for the case of regression. Typically, Euclidean distances are used to estimate closeness across neighbors.

3.5.4 Artificial Neural Networks (ANN)

An Artificial Neural Network (ANN) is a set of algorithms that aims to learn and identify patterns resembling the workings of the human brain. Neural Networks aim to understand a relationship across input and output data, capturing complex non-linear relationships among the variables in a flexible format. The typical ANN architecture consists of three main elements: input layer, hidden layer, and output layer. The Input layer receives and incorporates the information into the model, the hidden layer calculates relationships and patterns across the input data, and lastly, the output layer consolidates and returns the results from the network. Each layer is composed of interconnected units or nodes that, depending on their assigned weight, can impact the outcome of the nodes of the next layers. Activations functions can be defined to determine whether a given node can be activated or not.

There are many configurations of neural networks used for time series regressions. In our capstone project, a Long Short-Term Memory Neural Network (LSTM-NN) will be implemented. Deep learning methods differ from simple neural networks given the models are trained with more than two non-output layers. The parameters are updated accordingly using a backpropagation algorithm, estimated by the correct weights for each connection between the nodes of adjacent layers in our training set. Once the weights are learned correctly, the trained neural network can be used to make predictions for future values.

3.6 Multi-step Ahead Forecasting Strategies

Time series forecasting approaches are classified into two main categories: one-step forecasting and multi-step ahead forecasting. The former method predicts a single value of a historical time series. Multi-step forecasting approaches, on the other hand, consist of predicting the next series of steps given a historical time series composed of N observations. Since predicting the inner package temperature requires the actual ambient temperature conditions, the forecast methods cannot rely on methods that works with actual data from previous time horizons. This project applied an n-step ahead forecasting method, testing three Single-Output strategies: Recursive, Direct, and DirRec, adopted from Ben Taieb & Hyndman (2014).

f: Functional dependency between past and future observations for the package temperature.

y_t : The forecast of the variable predicted on time *t.*

d: Number of previous observations used to predict future values.

h: Point in time horizon = [*t, …, H*]

H: End time of the horizon referencing the package temperature forecasts for test set

N: Number of observations available in the training set

w: Fixed value composed of modeling error, noise, and disturbances for forecasting model

3.6.1 Recursive Method

In the recursive method, which is also called Iterated or Multi-Stage strategy, a single model *f* is trained to perform a one-step-ahead forecast, feeding the prediction into the model as an input to predict the subsequent time steps.

$$y_{t+1} = f(y_t, \dots, y_{t-d+1}) + w,\ t\ \in [d, \dots, N] \tag{3.4}$$

The recursive method aims to minimize the one-step-ahead prediction error variance, repeating the forecast process until the number of steps to the forecast are reached for the entire horizon. With the trained one-step ahead model being $\hat{f}$ the forecasts are given by:

$$\hat{y}_{N+h} = \begin{cases} \hat{f}(y_N, \dots, y_{N-d+1}) & if\ h = 1 \\ \hat{f}(\hat{y}_{N+h-1}, \dots, \hat{y}_{N+1}, \dots, y_{N-d+h}) & if\ h\ \in \{2, \dots, d\} \\ \hat{f}(\hat{y}_{N+h-1}, \dots, \hat{y}_{N+h-d}) & if\ h\ \in \{d+1, \dots, H\} \end{cases} \tag{3.5}$$

An example of a training and test data split is shown in Table 3.5, where predicted values are also included in the training set with other historical features to forecast the time steps ahead.

Table 3.5

Recursive Strategy, Train, and Test Set Split

Training Set (N)				Test Set (H)			
t-4	t-3	t-2	t-1	t			
t-5	t-4	t-3	t-2	t-1	t		
t-6	t-5	t-4	t-3	t-2	t-1	t	
t-7	t-6	t-5	t-4	t-3	t-2	t-1	t

A special case of this method exists when t equals d, where the model only gets 1 lag feature and transform into one-step ahead model.

3.6.2 Direct Method

In comparison with the Recursive strategy, the direct method relies only on historical data to forecast the predicted new values. Additionally, the direct method utilizes different forecasting models to increase the forecasting accuracy, where each model is estimated independently. In other words, it does not employ approximated values to compute the forecasts, thus eliminating the accumulation of errors. For each time series:

$$y_{t+h} = f(y_t, \dots, y_{t-d+1}) + w \tag{3.6}$$

where $t \in \{d, \dots, N-H\}$ and $h \in \{1, \dots, H\}$

the forecasts are obtained by using the H learned models $\hat{f}_h$ as follows:

$$\hat{y}_{N+h} = \hat{f}(\hat{y}_N, \dots, \hat{y}_{N-d+1}) \tag{3.7}$$

The main drawbacks of direct strategy are that it requires considerable computational time given the many models generated to predict each value in the horizon. This strategy also prevents us from considering complex dependencies across forecast values between models. In Table 3.6, an example of the various individual models is shown, each with a defined time lag.

Table 3.6

Direct Strategy, Train, and Test Set Split

Training Set (N)				Test Set (H)			
t-4	t-3	t-2	t-1	t			
t-5	t-4	t-3	t-2	-	t		
t-6	t-5	t-4	t-3	-	-	t	
t-7	t-6	t-5	t-4	-	-	-	t

3.6.3 Direct Recursive Strategy (DirRec)

The direct/recursive method (DirRec) applies a combination of the principles from the direct and recursive strategies previously discussed. The DirRec uses the approximations forecasts of the previous step to predict new values as the Recursive method, simultaneously using different models for every step as the Direct Strategy.

$$y_{t+h} = f(y_{t+h-1}, \dots, y_{t-d+1}) + w \tag{3.8}$$

where $t \in \{d, \dots, N-H\}$ and $h \in \{1, \dots, H\}$.

To obtain the forecasts, the end of the horizon H learned models are used as follows:

$$\hat{y}_{N+h} = \begin{cases} \hat{f}(y_N, \dots, y_{N-d+1}) & if\ h = 1 \\ \hat{f}(\hat{y}_{N+h-1}, \dots, \hat{y}_{N+1}, y_N, \dots, y_{N-d+1}) & if\ h \in \{2, \dots, H\} \end{cases} \tag{3.9}$$

Combining the strengths of Direct and Recursive methodologies allows us to capture the complex dependencies of inner package multiple temperatures time series by calculating approximated values for each step and minimizing the accumulation of errors generated by the recursive method alone.

3.6.4 Temperature Forecast Model Based on Ambient Temperature Profile

In Section 3.4.1.1 a model that predicts the inside temperature based on the previous inside temperature steps $y_{t+h-n,l,p}$ plus the impact of the ambient temperature is discussed and represented in Equation (3.3). Assuming the ambient temperature has a more significant impact predicting future inner temperature values, a naïve simplification can be done to the model to consider only one previous forecast ($n = 1$) instead of considering multi-step previous forecasts. Also, we can assume the factor β is not different for every time horizon (h) but only specific cumulative periods *i.* In this case, rearranging and applying these concepts in Equation (3.10) can be written as follows:

$$y_{t+h,l,p} = (1 - \beta_{i,p}) * \ y_{t+h-1,l,p} + \beta_{i,p} * X_{t+h,l} \tag{3.10}$$

Equation (3.11) resembles DirRec strategy discussed on Section 4.4.3, where:

$$y_{t+h,l,p} = f_{h,p} * y_{t+h-1,l,p} + w_{h,l,p} \tag{3.11}$$

Where $f_{h,p}$ is the functional dependency between past and future observations for the inner and external temperature at the given line and package configuration, and $w_{h,l,p}$ accounts for the modeling error, disturbances, and noise for the forecasting model.

$$f_{h,p} = (1 - \beta_{i,p}) \tag{3.12}$$

$$w_{h,l,p} = \beta_{h,i} * X_{t+h,l} \tag{3.13}$$

3.7 Methodology Summary and Conclusions

This section summarizes the methodology framework to implement a forecasting practice to predict inner temperatures for the passive temperature controlled packaging solutions across the sponsor company's supply chain. The process starts with the data collection and preparation of the summer and winter EPS 11 Box dataset. Next, select the relevant features that best relate to the predicted variable, such as the number of internal temperature lags and external temperature. Then a series of data transformations are applied to the data such as normalization and logarithm, creating split data sets for training and test sets. Lastly, each model is trained and tested for each season's data set, comparing their performance and selecting the best model based on lowest error. Further improvements are made iteratively by tuning each model and cross-validating the training data set to obtain the best possible model.

The results of these methodologies will be discussed in Chapter 4, which will be followed by an in-depth discussion in Chapter 5.

4. RESULTS AND ANALYSIS

This section reports on and analyzes the applied models by comparing the resulting forecasts to the inner temperatures from the sponsor company's prequalification testing. The first section will cover data normalization and feature selection. Then, classical methods including autoregression and ARIMA, will be compared and will further serve as basis for comparison to the machine learning forecasting models in Section 4.3. Next, the performance of all the models will be discussed and summarized. Lastly, the application of the models in new simulated external temperature conditions will be evaluated and discussed.

Before developing the models, training and test sets were split from the complete set of historical values. This split is necessary to have a reference to evaluate the forecast models. The training set is used as an input to be fed into the forecasting models, while the test set evaluates the performance of the models by comparing the predicted values against the actual values. To validate the accuracy of the initial training models from the dataset, a Time Series Cross-Validation was performed on the data, averaging the error by 10 k-folds.

The summer package temperature profile contains 277 data intervals, with a 10-minute time interval between each observation. For our analysis, 245 observations were used in the training set and 22 observations for the test set. In the winter package dataset, we initially had 322 observations. For the winter analysis, 290 observations were used in the training set and 32 observations for the test set.

Three indicators were used to gauge the performance of the forecast model. These indicators are defined as follows:

1. Root Mean Square Error (RMSE) – A statistical measurement to evaluate the size of the residual errors, by showing the distance between the actual data points and the best-fit model.
2. Mean Absolute Error (MAE) – Simplest measure of forecast accuracy that calculates the absolute value of the difference between the actual and the forecast value.
3. Mean Absolute Percentage Error (MAPE) – A statistical measurement to assess the forecast accuracy by assessing the model's performance versus the data used for the analysis.

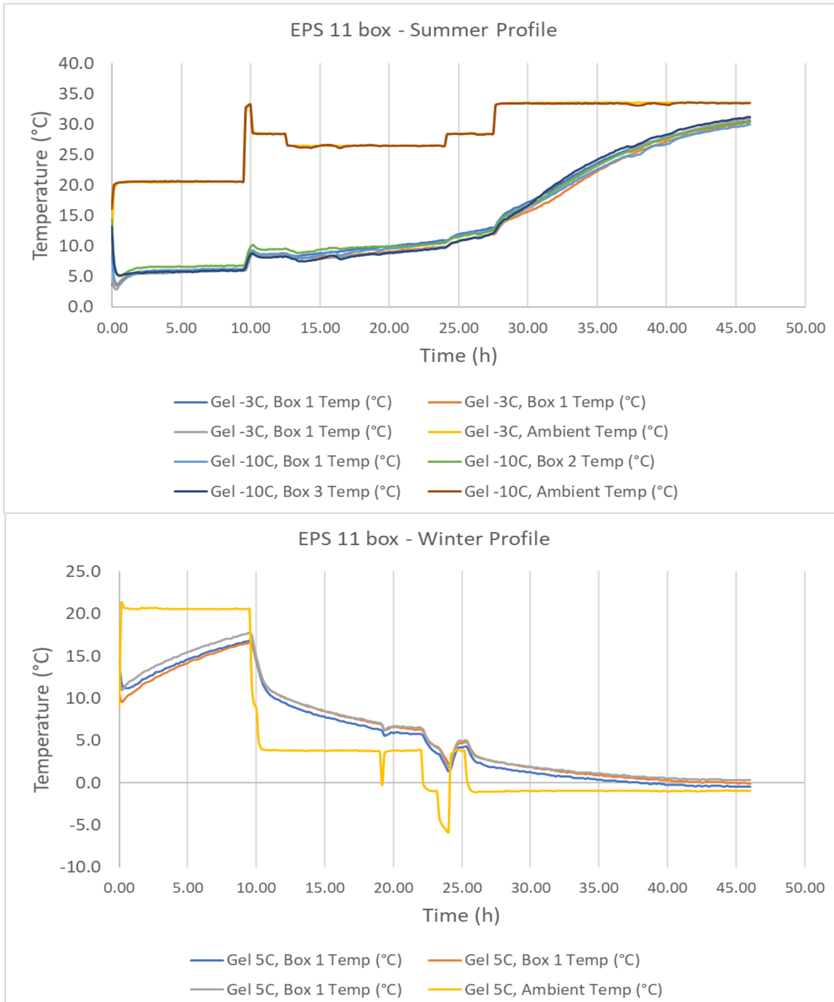
To find the best-fit model, different forecasting models were applied to determine which model produces the smallest RMSE, MAE, and MAPE. The results include the error comparison of the three configurations available of the EPS 11 box; the summer configuration with gels conditioned at -3°C, the summer configuration with gels conditioned at -10°C, and the winter configuration with gels conditioned at 5°C.

4.1 Data Normalization and Feature Selection

An initial exploration of the data is shown in Figure 4.1. The level and trend for inner temperature vials and ambient temperature are presented for the summer and winter package.

Figure 4.1

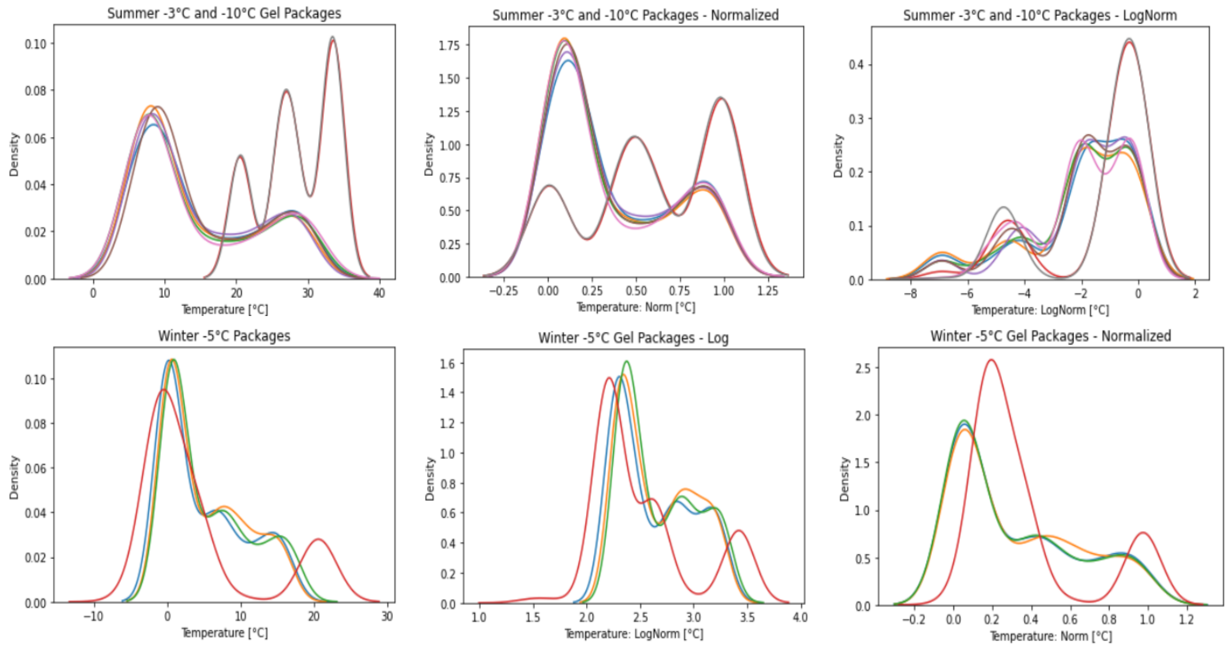
EPS 11 Box Summer (top) and Winter (bottom) Temperature Trend over Time



Normalization and Logarithmic transform were performed for all summer and winter package datasets to find the best fit for the models. Additionally, we also transformed our dataset into a logarithmic scale. The correlation analysis presented in Section 3.1 found a high positive correlation between internal and external temperatures in each profile setting. To further validate the correlation analysis, a density plot was implemented to compare the distribution curves of the temperatures across the features in the dataset. Figure 4.2 shows the density plots for internal and external temperature measurements.

Figure 4.2

Density Plots for Internal and External Temperature Measurements



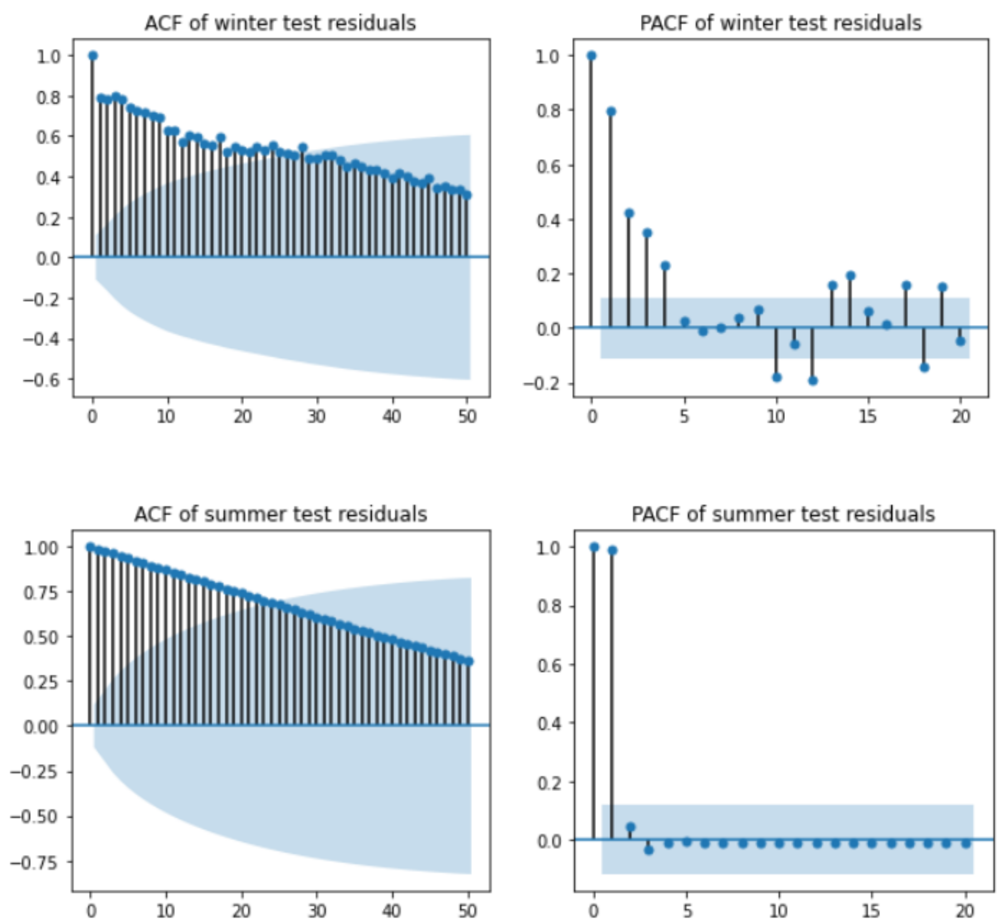
The initial normalization shows that the internal and external temperatures share a similar distribution when transformed into a logarithmic scale. This outcome confirms our initial assumption that external temperature is highly correlated with the internal temperature measures for each test.

Additionally, an autocorrelation analysis of the residuals was conducted to identify the critical lag features of the dataset. Autocorrelation applied to the internal temperature describes the similarity of historical measures against a time-shifted version of itself. This method can help identify hidden patterns in our time data.

From Figure 4.3, we can see that the first 24 lags of our data could be relevant for predicting the behavior of the internal temperature over time. The effect of initial lags is accumulating as we move to higher values on the *x*-axis. In order to avoid overfitting the model a partial correlation analysis was conducted to isolate each residual lag's impact.

Figure 4.3

Autocorrelation (ACF) and Partial Correlation Analysis (PACF)



By reviewing the autocorrelation analysis for the summer and winter data, we can conclude that for the summer packages, considering one lag is enough to capture the historical interdependencies of our data with a 95% confidence interval. On the otherhand for the winter packages, more time lags seem significant at a 95% confidence level in the data, indicating that the winter data has more interrelated lags.

Since it has already been confirmed that internal temperature is affected mainly by external temperature conditions, and to avoid overfitting, the number of lags will be reduced to one for simplicity.

4.2 Statistical Models

Autoregression and ARIMA models were used to forecast the internal temperature of the summer and winter EPS profiles. We evaluated the models using cross-validation and tuning to create a training model that will be used as a benchmark to be compared to other approaches that incorporate the external temperature as an exogenous factor.

Starting with an autoregressive analysis, both univariate and multivariate forecasts were evaluated for the summer and winter datasets. Forecast errors were then compared across both methodologies. Additionally, external temperature features were included as an extension of the model to evaluate its effect on the predicted package internal temperature. The results are summarized in Table 4.1

Table 4.1

Summary of Autoregression Forecasting Errors

		Test Set		
Data	Model	RMSE	MAE	MAPE
Summer	Autoregression, One-step	0.23	0.20	0.02
	Autoregression, Multi-step	0.76	0.66	0.04
	Autoregression, One-Step + External Temperature	0.31	0.22	0.02
	Autoregression, Multi-step + External Temperature	0.81	0.70	0.05
Winter	Autoregression, One-step	0.25	0.22	4.4E+12
	Autoregression, Multi-step	1.33	1.09	1.7E+13
	Autoregression, One-Step + External Temperature	3.85	2.70	6.6E+12
	Autoregression, Multi-step + External Temperature	1.17	0.96	6.4E+12

Autoregressive multi-step models performed worse in all metrics compared to their one-step counterparts when using ten k-folds time series cross-validation. The results appear to be in line with what other researchers have reported: multi-step models tend to suffer from error accumulation problems when the prediction period is long due to the bias and variance from previous time steps are propagated into future predictions. We can conclude that a one-step approach is sufficient to compare and validate each model's performance.

MAPE in the winter models yield values beyond 100% for each test. Considering the MAPE formula, the residuals per step are divided by the actual values of the temperatures in the container. Since some of the internal package's values are close or equal to zero, the values produced approach infinity. Thus, only RMSE and MAE will be considered for the rest of the models.

After building a time series forecast based on autoregressive lags, the external temperature data was incorporated as the exogenous variables in the model. The exogenous variables did not improve the model's accuracy significantly, regardless of the number of lags or number of future multi-step used. Based on our correlation analysis and empirical knowledge, external temperature gradients significantly impact packages' delivery under different ambient conditions. Thus, we conclude that simple autoregressive models in this research only serve as benchmark for comparison, not for predicting the internal temperature based on new conditions.

The second classic approach implemented in our datasets was the Autoregressive Integrated Moving Average analysis on summer and winter profiles. ARIMA models are time series methods that incorporate a linear regression of lags from autoregression (*p*), an integrated part from the differences of raw observations (*i*), and moving average used to measure the dependency between an observation and residual errors from the moving average model applied to lagged observations (q).

An array of ARIMA models on different hyperparameters and comparing the RMSE values across runs helps identify the best configuration for getting the least error for both training and test sets.

Initially the model was tested only considering the linear regression of lags from autoregressive (p) to compare the results to the autoregressive model from Table 4.1. It was found that an autoregression model ARIMA (3,1,0) returns the best RMSE when autocorrelation between residuals is not significant at p-value > 0.05. The ARIMA model now can serve as a reference for comparing other complex models.

Table 4.2

11 EPS Box Summer -10°C ARIMA Analysis on Multiple Hyperparameters

	Training Set			Test Set			Residuals
Model	RMSE	MAE	MAPE	RMSE	MAE	MAPE	p-value
ARIMA (0, 1, 0)	0.17	0.1	1.04	0.81	0.76	2.7	0
ARIMA (0, 2, 0)	0.15	0.09	0.92	0.7	0.64	2.29	0.036
ARIMA (1, 1, 0)	0.13	0.09	0.84	0.82	0.77	2.73	0.424
ARIMA (2, 1, 0)	0.13	0.09	0.86	0.79	0.74	2.61	0.323
ARIMA (3, 1, 0)	0.13	0.09	0.85	0.73	0.67	2.4	0.512
ARIMA (4, 1, 0)	0.13	0.09	0.85	0.74	0.69	2.46	0.710
ARIMA (2, 2, 0)	0.15	0.09	0.9	1.35	1.28	4.51	0.027
ARIMA (3, 1, 0)	0.13	0.09	0.85	0.73	0.67	2.4	0.512
ARIMA (3, 2, 0)	0.14	0.09	0.91	1.31	1.24	4.36	0.263
ARIMA (4, 1, 0)	0.13	0.09	0.85	0.74	0.69	2.46	0.710
ARIMA (4, 2, 0)	0.14	0.09	0.89	1.34	1.26	4.45	0.264
ARIMA (1, 2, 1)	0.13	0.08	0.8	1	0.7	2.41	0.545
ARIMA (2, 2, 1)	0.13	0.08	0.81	1.2	0.87	2.99	0.546
ARIMA (3, 2, 1)	0.13	0.08	0.79	1.4	1.05	3.62	0.918

The same analysis was applied to ARIMA models including average component q as a feature to improve the forecast model. In this case, ARIMA (1,2,1) got the best results and its predictors AR1 and MA1 were statistically significant with p-value > 0.05. Nonetheless, compared to ARIMA (1,1,0), ARIMA (1,2,1) performed worse in out-of-sample RMSE, getting a value of 1 °C, almost 30% higher. Thus, a simple ARIMA model (1,1,0) will be enough to benchmark the rest of the ML forecast models.

4.3 Machine Learning

As previously applied, machine learning algorithms were used to learn from both the internal and external temperature of the payload across time. One of the main benefits of using these approaches is that these models can identify non-linear patterns of data for new predictions. The five different machine learning approaches utilized in this project are the following: K-Nearest Neighbors, Support Vector Machines, Random Forest, Quantile Regression, and LSTM Neural Network. Each model's hyperparameters were tuned using a Grid search technique, aiming to set an optimal configuration for the testing set. The summary of these methods is shown in Table 4.3.

Table 4.3

Comparison of Performance Metrics of the Machine Learning Models

Data	Model	Test Set RMSE	Test Set MAE
Summer	K-Nearest Neighbors	1.72	1.52
	Support Vector Regression	0.45	0.43
	Random Forests	2.15	2.01
	Quantile Regression	0.45	0.57
	Long Short-Term Memory Neural Network	0.86	0.82
Winter	K-Nearest Neighbors	0.84	0.76
	Support Vector Regression	0.41	0.33
	Random Forests	1.19	1.09
	Quantile Regression	0.32	0.47
	Long Short-Term Memory Neural Network	2.46	2.34

From the machine learning methods used, Supported Vector Machines and Quantile Regression showed the best performance on the test dataset from the summer and winter profiles. As expected, K-NN performance is worse, as this methodology struggles with historical data containing trends. Lastly, the Random Forest and LSTM Neural Network approach gives the worst performance in both profiles.

4.4 Model Comparison

From the previous analysis we can conclude that many methods can be used to forecast temperature time series data with varying levels of accuracy. We compared several models, highlighted in Figure 2.2.

The results demonstrated in Table 4.1 from Section 4.2 indicate that one-step autoregression performs better than multi-step for both the summer and winter profiles. In addition, the external temperature did not lead to a better model; moreover, it slightly increased the error metrics in the summer models and changing them dramatically in the winter models. This can be explained by the fact that the lag feature has a high impact and the fact that not enough variability is present in the dataset. From these results we can confirm a one-step auto regression model without external factors is the best approach. From an implementation point of view, this means that the simpler modeling approach the better for the purpose of benchmarking our models.

In addition, the machine learning models used did not perform well due to insufficient data points to accurately predict the behavior of the packaging solution. In Section 4.3, Table 4.3 includes a comparison of the forecast errors on the test set. These models usually perform better with large datasets, especially the nonlinear algorithms. Machine learning models do not perform well at predicting data far ahead in the future, usually performing better at predicting values earlier in the time series. However, an exception was found in quantile regression, for the following:

1- Quantile regression's main objective is to optimize the loss function. In the model, the RMSE was selected as the error metric.
2- The inner temperature lag and external temperature coefficients of the regression are bound by the optimization constraints, allowing a non-linear assessment of the impact of external conditions for the package.
3- A set of regression models for each step was created instead of one regression for the entire horizon. Each external temperature step change leads to a new regression model with its own set of regression coefficients.

For each activity in the delivery of the package across the sponsor company's supply chain, it can be observed from Figure 4.1 that there is a significant step-change in the external temperature conditions impacting the internal temperature, either as cooling or a heating process. Therefore, six processes for cooling and heating were identified for splitting the data into smaller subsets, each with its own regression model. Table 4.4 revisits the UPS environment temperature profile conditions for winter and summer with each temperature process considered.

Table 4.4

UPS Subprocesses and Ambient Temperature Conditions for Winter and Summer

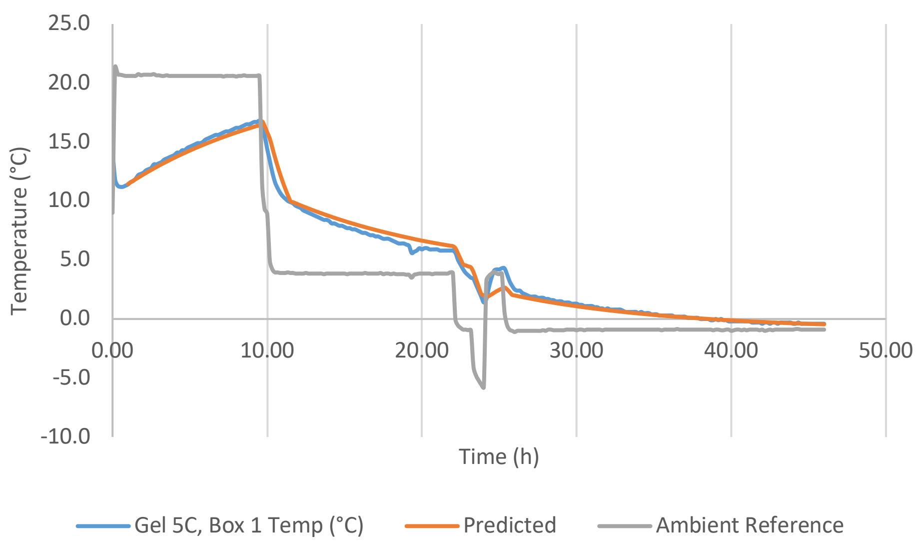
Step	*Mode of Transport*	*Total Time (hours)*	*Winter (°C)*	*Winter Process*	*Summer (°C)*	*Summer Process*
Stage in DC	Storage	9.5	22	1	22	1
Pick up at DC to Local Hub	Truck	0.5	10	2	35	2
At Local Hub	Storage	2.5	5	2	30	3
From Local Hub to Sort Hub	Truck	5.0	5	2	28	4
At Sort Hub	Storage	1.5	5	2	28	4
Stage on Tarmac	Storage	0.25	0	2	28	4
In transit to Airport Gateway	Aircraft	2.75	5	2	28	4
At Airport Gateway	Storage	1.25	0	3	28	4
In transit to local UPS facility	Truck	0.75	-5	4	28	4
At destination	Storage	1.25	5	5	30	5
With Courier	Truck	2.25	0	6	30	5
With Courier	Truck	3	0	6	35	6

Moreover, as time progresses for each step change, the internal temperature becomes asymptotically closer to the external temperature. Thus, each temperature process was further divided into two sub-steps to better capture the non-linearity of each cooling and heating process.

For the winter dataset, it was assumed that the effect of phase change of the gel packs is nonexistent considering the initial preconditioning of the gel packs at 5°C, which is above their melting point. Furthermore, since all packages were tested inside a controlled temperature chamber, we can infer that only the external temperature influences the internal temperature of the vials inside the package. Applying these assumptions to the quantile regression allows us to generalize the regression coefficients to be the same across all the model subsets. Figure 4.4 shows the predicted values of the entire EPS 11Winter Box dataset.

Figure 4.4

Quantile Regression Model Predictions on EPS 11 Winter Box Dataset



Coefficients obtained for the winter quantile regression were 0.0765 for the initial subprocess and 0.01525 for the second portion of the subprocess. From Equation (3.10), substituting the coefficients, we get:

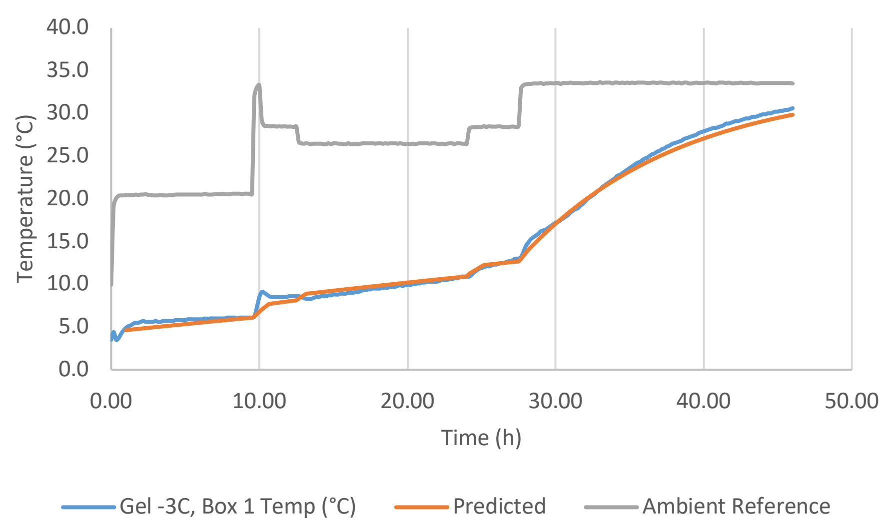
$$y_{t+h,l,p} = (1 - 0.0765) * \ y_{t+h-1,l,p} + 0.0765 * X_{t+h,l} \quad (4.1)$$

$$y_{t+h,l,p} = (1 - 0.0152) * \ y_{t+h-1,l,p} + 0.0152 * X_{t+h,l} \quad (4.2)$$

On the contrary, the effect of phase change of the gel packs is remarkably prevalent during the initial 26 hours for the summer package. The latent heat of fusion of the melting gel packs is present, having an increased capacity to absorb heat that does not change the temperature of the material. Furthermore, considering the initial preconditioning of the gel packs at -3°C and -10°C, respectively, the gel packs used in the laboratory testing are below their melting point, explaining the internal chamber's additional capacity to resist temperature changes. Therefore, compared to the winter quantile regression, the data must be partitioned into two main sections, first the model with the latent heat of fusion, and a second regression model only considering the external temperature factor. Figure 4.4 shows the predicted values of the entire EPS 11 Summer Box dataset.

Figure 4.5

Quantile Regression Model Predictions on EPS 11 Summer Box Dataset



Coefficients obtained for the winter quantile regression were 0.01020 and 0.00189 latent heat change process first and second subprocesses respectively, and 0.01804 and 0.01538 for the subprocess considering mostly external temperature influence.

4.5 Testing Models on New External Conditions

In Section 4.4, we discussed the implementation of classical and machine learning models on the winter and summer profiles of the EPS 11 box samples tested in controlled conditions. Both ambient temperatures and process times are assumed to be fixed across each step of the distribution lane, which follows the International Safe Transit Association (ISTA) lane data for ambient temperature profiles. The temperature profiles used by the sponsor company are the extreme ISTA temperature profiles across the United States. However, in real situations, variability must be considered when assessing the models' prediction performance.

Assessing the variability of the ambient temperature at the origin and destination during delivery is critical for creating representative temperature profiles. Table 4.4 from Section 4.4 defines a typical product delivery from the sponsor company based on extreme temperature conditions. As the models developed in this research give predictions based on the changes in temperature across time, an understanding of the ambient temperature variability is relevant to evaluate our model's performance under new conditions.

The temperature variability used in this section is extracted from the ISTA Lane Data Study for Pharmaceutical Cold Chain Strategies. In their database, information about daily averages, maximum, and minimum temperatures are published for expedited parcel shipping across the United States. To simplify the analysis, the variability of ambient temperatures is taken from the 72-hour average temperature range values, assuming a normal distribution of temperatures across all lanes. Table 4.5 summarizes the estimated statistics for the normal distribution profiles for summer and winter.

Table 4.5

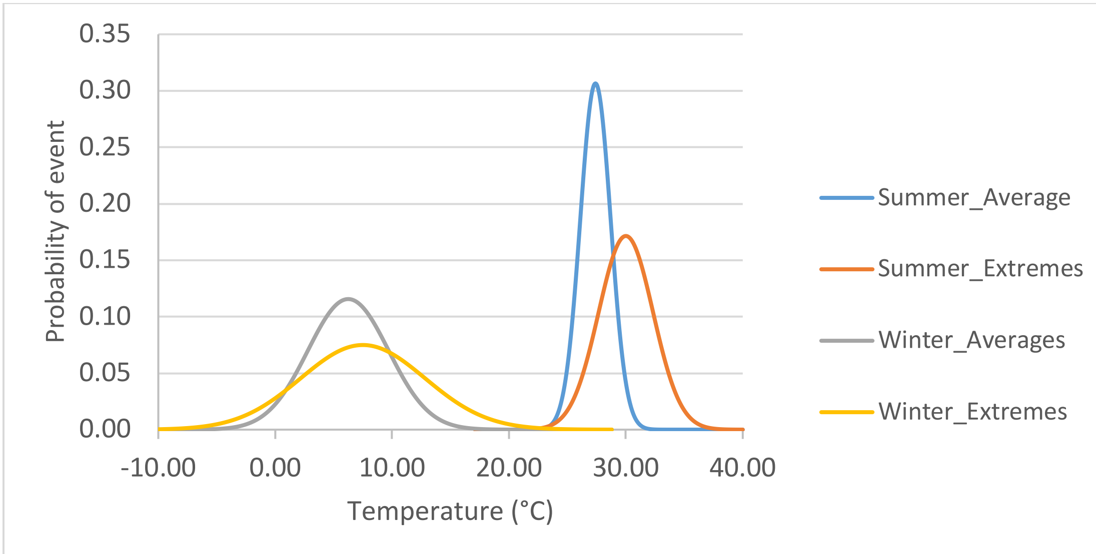
Normal Distributions for Summer and Winter 72-hour Averages and Extremes

	Summer Average Lanes	Summer Extreme Lanes	Winter Average Lanes	Winter Extreme Lanes
Maximum Average Temperature	30.7	35.9	15.0	21.0
Minimum Average Temperature	24.1	24.1	-2.5	-6.0
Mean Average Temperature	27.4	30.0	6.3	7.5
Standard Deviation	1.3	2.3	3.5	5.3
Coefficient of Variation	4.8%	7.8%	55.2%	71.0%
95% Confidence Interval, Lower bound	24.8	25.4	-0.5	-2.9
95% Confidence Interval, Upper bound	30.0	34.6	13.0	17.9

Figure 4.6 shows the normal distribution of daily average values for all the locations for summer's highest temperatures and winter's lowest temperatures. Given the high variability of the lanes studied, a profile of extreme temperature locations is added for comparison to the average normal curve.

Figure 4.6

Normal Distribution of Daily Average Temperatures for Summer and Winter Profiles across the US



A further calculation of the confidence intervals for our distribution profiles helps to define a framework of potential scenarios to analyze our models. The final estimations for the analysis of confidence interval at 95 percent of significance show that ambient temperature variability is bound within a range of 24.8°C to 34.6°C and -6°C to 21°C for summer and winter profiles, respectively.

In addition to the variability of ambient temperature conditions for each profile, the impact of variability in lead times for each process was evaluated, as this variability can affect the predicted temperature of the payload. Hence, a simple approach is proposed by assuming each step follows a normal distribution with a coefficient of variance of 15%. The results of the confidence intervals for each step are summarized in Table 4.6.

Table 4.6

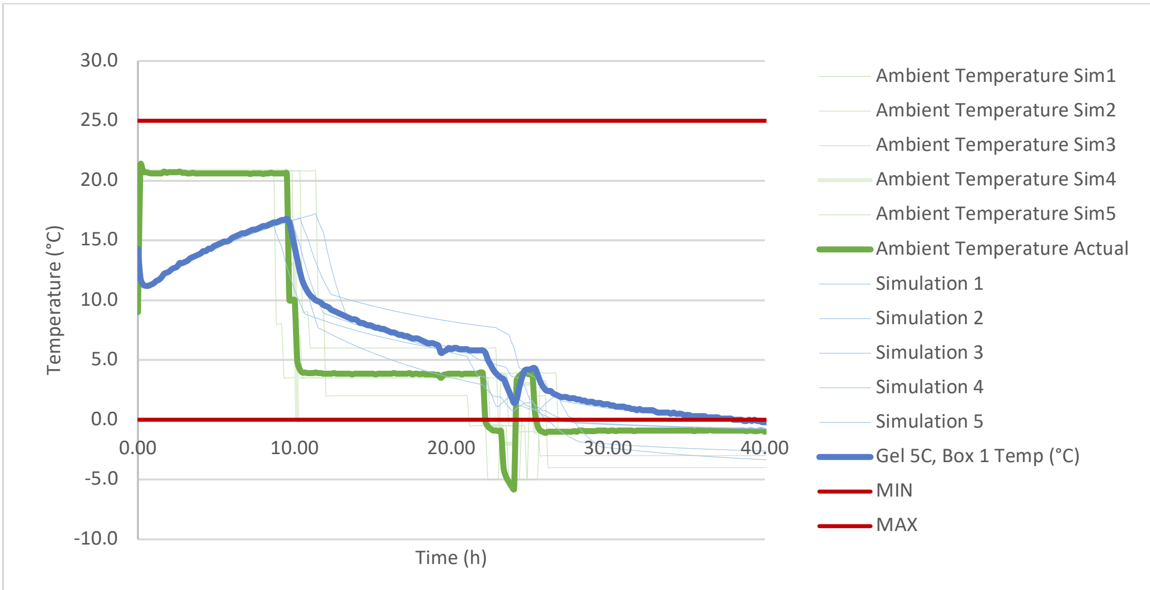
Normal Distributions for Summer and Winter Lead Times per Process Step

Step	*Mode of Transport*	*Average Time (hours)*	*Standard Deviation* (hours)*	*95% Confidence Interval, Min (hours)*	*95% Confidence Interval, Max (hours)*
Stage in DC	Storage	9.50	1.43	6.71	12.29
Pick up at DC to Local Hub	Truck	0.50	0.08	0.35	0.65
At Local Hub	Storage	2.50	0.38	1.77	3.23
From Local Hub to Sort Hub	Truck	5.00	0.75	3.53	6.47
At Sort Hub	Storage	1.50	0.23	1.06	1.94
Stage on Tarmac	Storage	0.25	0.04	0.18	0.32
In transit to Airport Gateway	Aircraft	2.75	0.41	1.94	3.56
At Airport Gateway	Storage	1.25	0.19	0.88	1.62
In transit to local UPS facility	Truck	0.75	0.11	0.53	0.97
At destination	Storage	1.25	0.19	0.88	1.62
With Courier	Truck	2.25	0.34	1.59	2.91
With Courier	Truck	3.00	0.45	2.12	3.88
*Assuming a Coefficient of Variation of 15%	Total	30.50	4.58	21.53	39.47

Combining the variability of ambient temperatures presented in Table 4.5 with the distribution of lead times per step proposed in Table 4.6 allows the model to capture how the uncertainty of external temperatures impacts the internal product temperature as it is delivered through the sponsor's typical supply chain. For instance, Figure 4.7 shows the impact of five simulated ambient temperature scenarios for the Winter Season on the inner package temperature, with variable temperature and lead times using the Quantile Regression model.

Figure 4.7

Internal Temperature Prediction for EPS 11 Winter Box on Simulated Profiles



As observed in the simulated runs, the variability of temperature and lead times dramatically impacts the predicted inner temperature on the long term. For instance, the payload in the five simulations evaluated reached the minimum threshold of 0°C within 26.1 to 37.2 hours, compared to the initial 38 hours in our sample laboratory box dataset. Furthermore, simulations of extreme lower temperatures and extended step lead times shows that the Box decreases the ability to safely deliver the product within the required temperatures for the analyzed horizon. Thus, understanding the ambient conditions at each location is vital to assess the exposure risk of the product vials in the packaging solution.

The proposed forecasting models in this section provide the sponsor company a framework for assessing future EPS 11 Boxes deliveries across their Supply Chain, considering the uncertainty of their delivery processes and ambient temperatures. Both classical and machine learning models were adequate to predict training and test set values with considerable high accuracy. However, from the methodologies implemented in this research, Quantile Regression is the only successful model that predicts the inner temperature of the payload given new ambient conditions. Despite the poor performance of the machine learning models tested under new ambient temperatures studied in this research, the models have the potential to increase their forecast performance as more laboratory testing datasets with different temperature profiles become available.

5. DISCUSSION

In this project we implemented multiple models, from classical to machine learning, to help the sponsor company develop predictive capabilities for their temperature sensitive products delivered in passive control shipped solutions for their last mile delivery. We analyzed the patterns of three configurations for the EPS 11 Box that were tested in controlled temperature chambers, based on extreme conditions defined for the summer and winter seasons. In this section, we summarize the business insights derived from our results, management recommendations for the implementation of the models developed in this research, and lastly the limitations and challenges of the study.

5.1 Insights and Management Recommendations

This research shows that the inner temperature of a payload can be modeled as a function of the external ambient temperature conditions. The data needed to make predictions with high accuracy can significantly vary depending on the forecasting technique selected. The general rule of thumb is that the availability of more useful and relevant data translates into better forecasting. Increasing the amount of data for training the models decreases the probability of overfitting the models. It also allows the trained model to extract insights into the features interrelations provided. However, selection of the models must also consider the practicality and interpretability of its coefficients, and business problem context. For example, if more data samples were available and were input into our Artificial Neural Network until the model has the lowest RMSE, it does not necessarily imply that the sponsor company should replace existing models and use ANN model for all their predictions. In the case of LSTM-NN, deep learning models are known for being difficult to interpret and require multiple iterations to reach a satisfactory result. Considering this high resource requirement, LSTM-NN might not be responsive enough for the volume of daily payloads. Hence, making this decision in the business environment depends on how frequently temperature curves are being measured, the response time needed to test a new solution, and ease of implementation.

The applied models in this report can be further improved by identifying the frequency of failed deliveries per lane from the sponsor company's historical ERP data. This information will help identify the ambient temperatures and conditions that took the payload out of bounds. From these failed deliveries, a frequency analysis of failures per lane can help tune the models in line with the temperature bounds for the product.

Given that the sponsor company has operations worldwide, the insights and learnings from the models developed with operation data in the United States can be extrapolated to any country where the sponsor company has cold chain operations utilizing passive temperature-control solutions. Additionally, the models are product agnostic, as long as the shipped product is of a similar amount, the same package insulation material, and similar ambient temperature ranges. The sponsor company could better monitor their operations, focusing on potential delivery outliers and failed shipments, and potentially improve productivity by focusing on the vials likely to fail by documenting and communicating their findings to their peers globally.

The models provided could also be used on any perishable product sensitive to external ambient conditions, not implicitly related to the pharma industry. As the company portfolio increases, the demand for using predictive models for their payloads will increase.

Using these models to reconsider how passive temperature-control packaging works in tandem with active temperature-control infrastructure can lead to more efficient and sustainable use of resources. Reduction of the use of passive packaging solutions, which are often single use, would help the sponsor company reach its sustainability goals. If the sponsor company takes over the delivery for example, buffer zones could be put at different stages in the supply chain, leveraging the power of using both active and passive temperature-control throughout the supply chain. In addition, If the sponsor company decides to take transportation and delivery inhouse, the company will have more control on the temperature conditions the product is subject to.

Lastly, developing a framework forecasting temperature can enhance predictive practices for the sponsor company. Equivalent examples include continuous improvement programs in many facilities or how predictive maintenance works in manufacturing sites. Establishing a recurrent program that monitors future package deliveries based on existing demand could be highly valuable to the sponsor company. This gives visibility and transparency to all departments involved in the payload deliveries. In addition, it may help to raise when potential disruptions are present in our projected deliveries and take a course of action, that could be increasing the number of gels, adding a thermal blanket, changing the material insulator, and more. Predictive methods used in this study could highly enhance inventory management policies which would positively affect the sponsor company's working capital management leading to more efficient operation and freeing up resources for research and design: the primary driver of growth.

Based on the results from Chapter 4, and building on the discussion of the models, our recommendations are as follows:

1. For the "EPS 11" box, the differences between the summer gel pack preconditioning temperatures are minor. Based on our modeling and results, preconditioning the gel packs at -3°C and -10°C does not significantly affect behavior. The sponsor company

should consider the extra costs accrued by preconditioning gels at -10°C instead of -3°C. Preconditioning the gel packs closer to -3°C could lead the sponsor company to significant cost savings.

2. The summer and winter profiles need to be considered and modeled separately due to the heat capacity and phase change of the gels under the different ambient temperatures for both profiles. The sponsor company has been aiming to unify the solutions in an effort to reduce complexity. The models and data confirm that this is not a viable option for this case.
3. The models provided by this research are a low-cost initial testing tool to prescreen packaging solutions. As the sponsor company always does prequalification testing for its shipping solutions, a less resource-intensive way to screen their solutions would be to use these models as a powerful tool to reduce the number of lab tests conducted, saving costs and time.
4. The sponsor company could benefit from establishing a process that automates the calculation of the temperature predictions from our model coefficients based on a known destination ambient temperature. This system could extract future weather conditions with an API connection and simulate runs based on these profiles. In this manner, quality and supply chain teams could assess if the package conditions for the product selected will arrive at the destination.

5.2 Limitations

This section will discuss the limitations faced by this research. This research has only considered ambient temperature and no other meteorological factors. Considering the meteorological factors and due to the nature of the prequalification testing and the design of the testing chamber, we do not have data on how the different ambient weather conditions affect the internal temperature of the payload. The prequalification testing is done in isolated chambers in which the only controlled variable is temperature. In the real world, many factors are at play, including but not limited to:

- Time of day and solar radiation. In many delivery locations and especially at touchpoints, the shipping boxes are exposed to a level of direct sunlight. The amount of solar radiation is always present during the daylight hours but not at night for example. Situations where cloud cover is abundant would have different heat transfer mechanisms to situations where direct sunlight is hitting the shipping boxes. This affects the shipping container through radiation heat transfer, an increase in the proportion of radiation heat transfer would lead to difference in the temperature behavior, more sunlight leads to more heat being absorbed by the shipping solution.
- Humidity. The United States has many different climates ranging from arid, dry climates to humid, tropical ones. This research was unable to consider these effects and incorporate them into our analysis. Increasing relative humidity affects the heat transfer process by decreasing the heat transferred. This is due to lower air thermal conductivity with the higher proportion of humidity.
- Wind speed. Increasing wind speeds increases the heat transfer rates when the boxes are exposed to air currents. Stronger air currents lead to an increase in heat transferred by convection.

In addition, variations of solution setups used by the sponsor company were not included in the prequalification testing data. Different boxes, insulation types, and number of gel packs all affect the behavior of the passive cooling solution. These factors could not be accounted for in the models provided, even though they have considerable implications on heat transfer to the payload. This research is unable to deliver on the following points:

1) Due to having only three configurations of gel conditioning for one container, we were unable to create a model that predicts the success of the container due to changes in its components and estimate its performance against various ambient temperatures.

2) We are unable to suggest optimal configurations for the containers due to not having data on real weather conditions, actual behavior during shipping, and considering that data on types of boxes, gel packs, and insulators for different shipping solutions is unavailable. Feeding our models with data on the configuration of the shipment solutions and the cost of the component parts could lead to a prescriptive optimization model that would help the sponsor company reduce costs and better serve their customers.

6. CONCLUSION

Through this project, we looked at the passive cooling solutions used by the sponsor company. We assessed the prequalification testing data and the sponsor company's supply chain setup. The fact that the only variables available in the testing data were the ambient and internal temperatures posed a challenge to deliver a predictive model to forecast the internal temperature changes.

The key aim for this project was to enable the sponsor company to evaluate the performance of their passive containers along the supply chain and whether the internal temperature stays within the required bounds during transit to the end customer through a predictive model. Through this research, two types of approaches were used to model the behavior of the "EPS 11" box's shipping configuration: classical methods and machine learning. The output of the models was then compared and discussed. Section 6.1 includes recommendations for future research.

To summarize, of the original deliverables, we were able to deliver on the following points:

1) Initial testing models for the sample case data have been created to predict the inner container temperature along a typical representation of the company's supply chain.

2) A validated model that predicts the likelihood of solution success against different ambient temperature profiles was also created.

3) Recommendations were given in Section 5.1 based on the research findings, including preconditioning the summer gel packs closer to -3°C, using these models to prescreen solutions and lanes, and separate consideration of the summer and winter profiles.

6.1 Future Research

Reflecting on our research, we see many promising avenues to build on this research and take this work further. Considering the importance of the temperature-sensitive logistics field and the limitations of this project, many further opportunities and promising ventures exist.

The research has been focused on the “EPS 11” box, as the sponsor company uses several boxes to package their goods for shipping. Though the sponsor company is increasing its use of this box, others are also still in use, providing other dimension options and characteristics. Future research into these solutions would provide the sponsor company a chance to reduce the complexity of their solution catalogue.

Research into the effects of the different types of insulation on the final payload would also be valuable. Insulation is the main protector of the payload shielding it from most of the ambient temperature conditions. Research into which insulators provide better protection and how that protection varies between different weather conditions would prove valuable to the entire perishable supply chain market, not just pharmaceuticals.

In addition, the internal configuration and setup of the passive packaging solutions play an important role in the behavior of the package as a whole. Further research could study the effects of the different types of gel, the amount in each package, and their positioning within the physical space would not only benefit the sponsor company for its specific problem but all industries working with active cooling solutions.

Further study can be done with a focus on the setup of the supply chain from a network design standpoint. Further research focused on this topic could consider the number of stops, the length of travel between points, and any idle times. This would be particularly beneficial to the sponsor company as their courier offers lower prices for off-peak package pickup times, which would lead to a more cost-effective approach. UPS, for example, has pick-up slots with lower prices for specific times throughout the day. Changing the schedule could lead to a longer delivery time and necessitate that the package withstands more exposure to ambient temperature conditions. Though the sponsor company always does prequalification testing for the solutions before making changes, accurately predicting the success of different scenarios is now possible without prequalification testing.

Only ambient temperature was considered for this study. More comprehensive research could consider atmospheric conditions other than temperature, e.g., humidity, solar radiation, and windchill. Since a considerable amount of time is spent at or near air terminals, multivariate meteorological data from airports could be used. This data is usually of high quality due to the meteorological monitoring systems present in all airports. This data could be considered with more complex machine learning or neural network techniques that preform best with large datasets.

This is increasing in importance as more dramatic weather patterns are brought on by the effects of climate change.

This project has helped the sponsor in its objective of delivering lifesaving, temperature-sensitive pharmaceuticals to patients with no deterioration in quality, while complying with quality control regulations, and in a cost-effective manner. Modeling the temperature of goods in the supply chain is of value to pharmaceutical companies and more broadly, any other actors with temperature-controlled supply chains.

Humanitarian actors also depend heavily on passive packaging to deliver temperature sensitive medications vaccines to vulnerable communities. To support humanitarian efforts, this approach can be applied to predict the success of temperature-controlled deliveries to remote areas or when active cold chain infrastructure does not exist.

REFERENCES

Accorsi, R., Manzini, R., & Ferrari, E. (2013). A comparison of shipping containers from technical, economicand environmental perspectives. *Transportation Research, Part D* (26), 52-59.

Ben Taieb, S., & Hyndman, R. (2014). Boosting multi-step autoregressive forecasts. *Proceedings of the International Conference on Machine Learning*, in *Proceedings of Machine Learning Research* 32(1):109-117.

Ben Taieb, S., Bontempi, G., Atiya, A., & Sorjamaa, A. (2011). A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. *Expert systems with applications*.

Berthou, T., Stabat, P., Salvazet, R., & Dominique, M. (2014). Development and validation of a gray box model to predict thermalbehavior of occupied office buildings. *Energy and Buildings, Volume 74*, 91-100.

Grand View Research. (2021). *Pharmaceutical Logistics Market Size, Share & Trends Analysis Report By Type (Cold Chain Logistics, Non-cold Chain Logistics), By Component (Storage, Transportation), By Region, And Segment Forecasts, 2021 - 2028.* San Francisco: Grand View Research.

Kharfan, M., & Chan, V. W. (2018). Forecasting Seasonal Footwear Demand Using Machine Learning [Master's capstone project, Massachusetts Institute of Technology]. DSpace@MIT. http://hdl.handle.net/1721.1/117612

Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLoS ONE*, 13(3).

Mateo, F., Carrasco, J., Sellami, A., Millán-Giraldo, M., Domínguez, M., & Soria-Olivasa, E. (2013). Machine learning methods to forecast temperature in buildings. *Expert Systems with Applications, Volume 40*(4), 1061-1068.

Mentink, T. (2018, November). Optimization of packaging options in the cold supply chain. *Eindhoven University of Technology*.

Mikulic, M. (2021, April). *Revenue of the worldwide pharmaceutical market from 2001 to 2020*. Retrieved from Satista: https://www.statista.com/statistics/263102/pharmaceutical-market-worldwide-revenue-since-2001/

Mustafaraj, G., Lowry, G., & Chen, J. (2011). Prediction of room temperature and relative humidity by autoregressive linear and nonlinear neural network models for an open office. *Energy and Buildings, 43*, 1452 - 1460.

Narvekar, M., & Fargose, P. (2015, July). Daily Weather Forecasting using Artificial Neural Network. *International Journal of Computer Applications, Volume 121*(No. 22), 9-13.

O'Neill, A. (2021, July 30). *Global gross domestic product (GDP) at current prices from 1985 to 2026*. Retrieved from Statista: https://www.statista.com/statistics/268750/global-gross-domestic-product-gdp/

Pillai, B. S. (2014, August). Selection of Cold Chain Packaging Solution Using an Interpretive Structural Modeling (ISM) approach [Master's capstone project, HEC Montreal].HEC Montreal Biblioteque. https://biblos.hec.ca/biblio/memoires/2014NO62.PDF

Rong, A., Akkerman, R., & Grunow, M. (2011). An optimization approach for managing fresh food quality throughout the supply chain. *International Journal of Production Economics, 131*(1), 421-429.

Shmilovici, A. (2009). *Data Mining and Knowledge Discovery Handbook.* Boston, MA: Springer.

Singh, M. (2005, May 5th). Doctoral Dissertation. *The Pharmaceutical Supply Chain: a Diagnosis of the State-of-the-Art*. Cambridge, Massachusetts, USA: Massachusetts Institute of Technology.

Smadi, M. M., & Mjalli, F. S. (2007). Forecasting Air Temperatures Using Time Series Models and Neural-based Algorithms. *Journal of Mathematics and Statistics, Volume 3*(No. 2), 44-48.

Suradhaniwar, S., Kar, S., Durbha, S., & Jagarlapudi, A. (2021). Time Series Forecasting of Univariate Agrometeorological Data: A Comparative Performance Evaluation via One-Step and Multi-Step Ahead Forecasting Strategies. *Sensors, Volume 21*(Issue 7), 2430. Retrieved from https://doi.org/10.3390/s21072430

Tan, M., Wang, J., Li, P., & Xie, J. (2020). Storage time prediction of glazed frozen squids during frozen storage at different temperatures based on neural network. *International Journal of Food Properties*, 1663-1677.

Xu, W. Z. (2014). Neural Network Model for the Risk Prediction in Cold Chain Logistics. *International Journal of Multimedia and Ubiquitous Engineering, Vol. 9*(No. 8), 111-124.

Xu, W., Zhang, Z., Gong, D., & Guan, X. (2014). Neural Network Model for the Risk Prediction in Cold Chain Logistics. *International Journal of Multimedia and Ubiquitous Engineering*, pp. 111-124.

Zaharia, A. (2021, June 14). *Passive Cooling vs. Active Cooling – What's the Difference?* Retrieved from https://www.simscale.com/: https://www.simscale.com/blog/2017/01/active-and-passive-cooling/

Zao Zhang, Y. D. (2020). Temperature Forecasting via Convolutional Recurrent Neural Networks Based on Time-Series Data. *Complexity, 2020*, 1-8.

Zwierzycki, W., Bieńczak, K., Bieńczak, M., Stachowiak, A., Tyczewski, P., & Rochatka, T. (2011). Thermal damage to the load in cold chain transport. *Procedia Social and Behavioral Sciences 20*, pp. 761–766.