Temperature controlled shipments in temporary storage location choice of KLM Cargo

Master Thesis

Zhao Kang

Temperature controlled shipments in temporary storage location choice of KLM Cargo

By

Zhao Kang

to obtain the degree of

Master of Science
in Civil Engineering

at the Delft University of Technology.

Student number: 4706463

Thesis committee:	Prof.dr.ir. L.A. (Lóránt) Tavasszy	TU Delft, supervisor, committee chair
	Dr. M. (Milan) Janic	TU Delft, supervisor
	Dr. W.W.A. Beelaerts van Blokland	TU Delft, supervisor
	Rutger-Jan Pegels	KLM Cargo, external supervisor
	Dr. Xiao Lin	Daily advisor

Preface

This master thesis is the final academic project for the master program Civil Engineering (Track: Transport & Planning) at Delft University of Technology. The topic of this thesis is the temperature-controlled shipments in temporary storage location choice of KLM Cargo, which is performed in collaboration with Air France KLM Martinair Cargo and TU Delft.

I am grateful for this opportunity to undertake such a great project related to the supply chain industry. I would like to thank Prof.dr.ir. L.A. (Lóránt) Tavasszy for his help and encouragement in completing this project over the past seven months. I would also like to thank Dr. M. (Milan) Janic for his advice. His feedback and advice were always insightful, which gave me confidence to finish the thesis better. Next I would like to thank Dr. W.W.A. Beelaerts van Blokland. Thank you for taking time out of your busy schedule to guide me and give me valuable advice to complete this subject. In addition, I would like to thank Rutger-Jan Pegels, my company supervisor, who gave me this intern position. Many thanks go to all other KLM employees who helped when necessary. I would like to thank Dr. Xiao Lin as well. I enjoyed weekly discussion with him, who gave me lots of insights into this project.

I would like to thank my family. Thank you for your support and help over the past 24 years. I also want to thank my friends. Thank you for your company and help when I left my hometown these two years.

Zhao Kang

Delft, November 8, 2019

Summary

Ambient temperature has been identified as the most common reason for quality loss of most high value perishable goods. In actual operation, it is not cost- effective to cool perishable products in active and/or passive containers when their storage time in the warehouse is short. This thesis aims to develop a control strategy to reduce quality loss of perishable goods (pharmaceuticals and flowers) due to temperature in the stage of warehouse storage location choice. To achieve this objective, we propose the following main research question and sub-questions:

Research question:

"Given information about perishable goods (pharmaceuticals and flowers) and outside weather forecasts, in what ways KLM employees can better make decisions in the storage location choice?"

Sub-questions:

1. What is the relationship between the outside weather (temperature, humidity, wind, rainfall, etc.) and the indoor temperature at div storage locations?

2. How to design a control strategy to reduce total quality loss?

3. Given the information of perishable goods and weather forecasts, how much improvement can the control strategy provide according to the simulation results?

The first sub-question aims to predict the indoor temperature of different locations in the warehouse according to the future weather forecast. Here, the multiple linear regression model and the neural network model were introduced. The k-fold cross-validation approach was used in the neural network model to prevent overtraining and in the linear regression model to find the optimal independent variables. The prediction performance of the two models was evaluated by RMSE values. It was found that when the heating was turned off, the fitting degree of the two models is basically within the reasonable range. When the heater is turned on, the multiple linear regression models of half of the detectors have a large RMSE value. In general, the result of the neural network model has smaller error and higher fitting degree.

The second sub-question aims to establish a control strategy to reduce the total quality loss of perishable goods in the location choice procedure. The control strategy consists of three parts: indoor temperature prediction model, quality loss model and location choice model. The indoor temperature prediction model has been answered in the part of the sub-question 1.

For the quality loss model, it is composed of two parts. Firstly, the hourly vase life/shelf life loss at a given temperature can be estimated by the temperature-time quality loss function. Then, the total vase life/shelf life loss to money are converted to money by the time-money quality loss model.

For location choice model, we proposed a flowchart approach, which is shown below. For each shipment, we need to check the storage capacity of each location first. Then, the location with the least quality loss value will be selected. After that, the capacity information for each location will be updated.

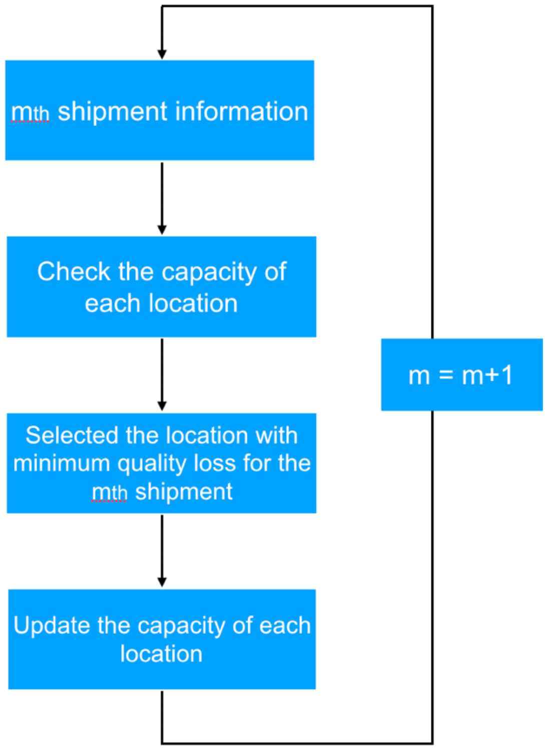


Figure 0. 1 The flowchart approach for storage location choice.

For this method, we need to decide in advance the sequence in which the goods are allocated. Therefore, the simulation experiment was provided to find out the best sort criteria as well as to answer the sub-question 3. We introduced three sort criteria of the flowchart approach and found that the average quality loss of the shipment at all storage locations was the best sort criteria based on simulation results.

We also showed the percentage of each location selected in the simulation of strategic level and tactical level, which provides the basis for KLM cargo to make adjustments in the future. In addition, we compared the simulation results based on the proposed control strategy and the actual case. The results show that with the increase of indoor temperature, the proposed control decision can better reduce the quality loss of products.

List of Figures

List of Tables

List of Abbreviations

Abb.	Meaning
ANN	Artificial Neural Network
IATA	International Air Transport Association
KLM	Koninklijke Luchtvaart Maatschappij (Royal Dutch Airlines)
KPI	Key performance indicator
MLR	Multiple linear regression model
MSE	Mean squared error
RMSE	Root mean square error
SCCAS	Smartest Connected Cargo Airport Schiphol
TIACA	International Air Cargo Association
ULD	Unit Load Device
WSN	Wireless Sensor Network

Contents

Chapter 1 Introduction

1.1Background information

This section is intended to provide some background information on the perishable supply chains as well as general information on the transportation of KLM Cargo. This information contributes to a better understanding of the purpose and significance of the subject.

1.1.1 Perishable goods supply chain

There is an obvious change of lifestyles among consumers as the pace of modern industrialized society has accelerated (Olsson, 2004). With the consideration of health, environment and convenience, today's consumer preferences are shifting toward fresh, and healthy consumable products, thus the demands of perishable goods are gradually growing (Diop and Jaffee, 2005). Perishable goods, generally with short shelf-life, now can be found in every aspect of people's lives: consumers all over the world can drink milk from New Zealand, decorate their homes with freshly picked tulips from the Netherlands, and eat delicious fruit from South Africa. Research reveals that perishable products accounted for 38% of total-store sales while they made up about 65% of total store shrink for 43 countries in 2011 (Cognizant, 2015). This lifestyle change creates new business opportunities for logistics service providers, who need to provide more convenience to consumers through perishable goods that can be prepared, transported and supplied quickly and reliably.

Perishable supply chain business can be broadly divided into the following four categories: Food, including fruits, vegetables, live fish, shellfish, meat and meat products; Flowers and plants; Pharmaceutical products, including pharmaceuticals, vaccines, plasma; Live insects and mammals (Sales, 2016).

For the floral industry, the global consumption of flowers is estimated to vary from $40 billion to $60 billion a year. Eighty percent of floral consumption comes from six countries, including Germany, the US, the UK, France, the Netherlands and Switzerland (Vega, 2008).

A typical fresh flower supply chain consists of the following steps shown in Fig.1.1, covering the entire logistics process from picking flowers on farms to delivering them to consumers. This process involves transhipment at multiple locations and coordination of multiple modes of transport, which usually takes about 10 to 12 days (Eric, 2016).

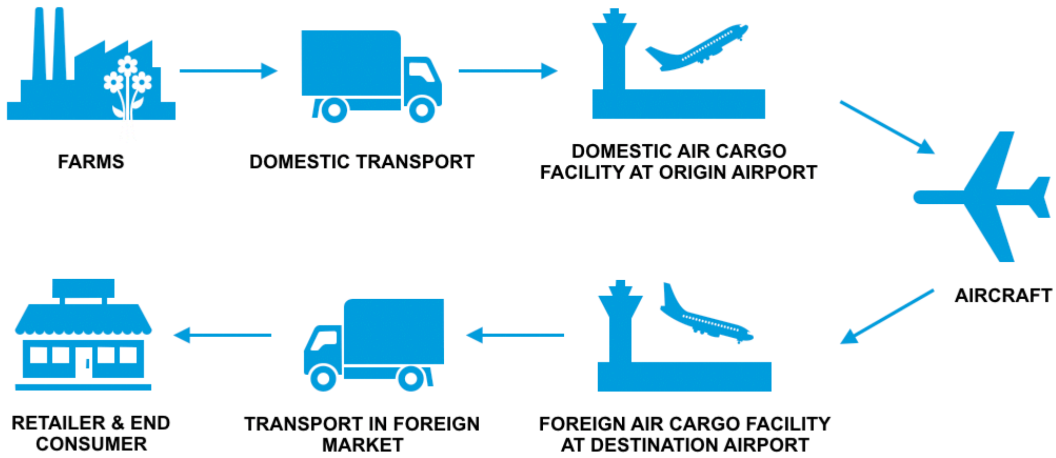


Figure 1. 1 Typical floral supply chain (Vega, 2008).

In the growing cool chain market, the pharmaceutical industry is the largest and most profitable single industry as well as the most temperature-sensitive. Compared with the floral supply chain, the pharmaceutical supply chain is more complex, including raw materials sourcing, manufacturing, distribution and dispensing. At present, health care products are mainly manufactured and exported by Switzerland, Germany, the United Kingdom and the United States. Besides, more than 30 airlines serve the pharmaceutical supply chain and are investing in expanding their cold-chain facilities.

The quality of perishable products is easily affected by environmental factors, such as chemical changes, ambient temperature and atmosphere (Katsaliaki, Mustafee, & Kumar, 2014). Ambient temperature has the greatest impact on affecting the shelf-life of perishable goods among all of the particular factors (Mc Carthy et al., 2018). It is estimated that due to the poor processing technology after harvest, including poor temperature management, the total waste of agricultural products before the consumption stage accounts for about 60% in the supply chain (Gustavsson, Cederberg, Sonesson, Van Otterdijk, & Meybeck, 2011). For pharmaceutical products, it is estimated that as about 15 to 20 percent of the products are wasted during transportation due to storage temperature and have to be discarded upon delivery (Sales, 2016). Thus, temperature management is important to reduce the loss of perishable products.

Perishable supply chains are more complex than other product supply chains due to the short shelf-life of goods, which requires the cooperation of product growers, retailers, carriers and logistics companies. The cold chain association was established in 2003 to develop global industrial standards and rules for perishable products. It now works closely with aviation agencies such as the International Air Transport Association (IATA) and the International Air Cargo Association (TIACA) to build more efficient perishable supply chains.

Ensuring the quality of these perishable products and providing the freshest products to customers is also one of the challenges that researchers are struggling with. The earliest research in this field can be traced back to (Straughn & Church, 1909), who studied the quality decline of agricultural products from the first harvest to the warehouse (Paterson, Tanksley, & Sorrells, 1991). Thanks to the development of technology, some new technologies are constantly emerging to control the quality of perishable products. For instance, the Wireless Sensor Network (WSN) has been proposed to provide real-time environmental information about perishable goods to decision-makers (J. Wang et al., 2015). Radio Frequency Identification (RFID) can reliably gather environmental information through sensors and help reduce waste (Lorite et al., 2017). These technologies bring new insights to the temperature-based quality control of perishable products.

1.1.2 KLM Cargo

This research was initiated by the Performance Management department of KLM Cargo as part of the Smartest Connected Cargo Airport Schiphol (SCCAS) collaboration. The Performance Management department of KLM Cargo has been established to structurally measure the performance of all units and processes of KLM and from there identify possibilities and needs for improvements.

KLM Cargo (Air France KLM Martinair Cargo) is the specialized air cargo business of the KLM Group, which is a key player in the air cargo industry. It offers a worldwide network of 457 destinations from our two hubs, Paris Charles de Gaulle and Amsterdam Airport Schiphol (KLM, 2019). KLM Cargo is dedicated to providing air transport services for perishable supply chains, mainly including floral supply chains and pharmaceutical supply chains. In 2018, it transported more than 84,000 tonnes of flowers from Africa and Latin America to Amsterdam Airport Schiphol and delivered pharmaceuticals to nearly 300 stations around the world.

As shown in Figure 1.2, currently there are three freight buildings for KLM Cargo at Schiphol. [redacted]

[redacted]

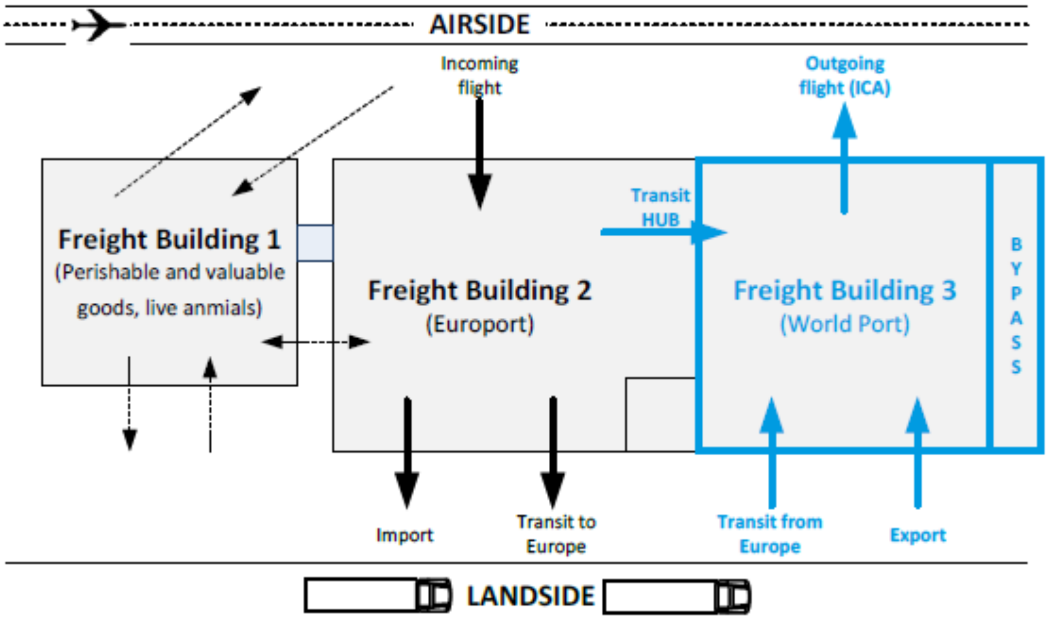


Figure 1. 2 The map of KLM Cargo at Schiphol.

The typical air cargo supply chain of KLM Cargo is shown below in Fig. 1.3, which consists of multiple commercial parties. There are five different steps in which shipments are transported from place of origin to destination, with KLM cargo serving as the airline positioned in the middle of the supply chain between the sending and receiving end of the chain. Production of the shipment is completed at the beginning of the supply chain (grower). The forwarders are traditionally KLM's most important customers, they are an agent that collects and bundles shipments and then provides it for transport by the airline to the correct destination.

Figure 1. 3 The general air cargo supply chain of KLM Cargo.

1.1.3 Warehouse storage process

Before introducing the warehouse storage process, we will introduce the classification of delivered shipments. The standardized cargo units are called Unit Loading Devices (ULDs) and can be classified into three different types:

Through/T-ULD: is filled shipments destined for one destination and ready for flight.

Mixed/M-ULD: should be broken down as it contains shipments to multiple destinations.

Bijbouwer/BB ULD: there is spare space on the ULD, so shipments with the same destination can be added.

T-ULDs are send to the PCHS output lanes on the airside of the building. They do not need to be stored at the temporary storage locations shown in Figure 1.4, nor need they be broken down or built up. This project mainly focuses on M-ULDs, because the temporary storage locations are mainly set up to serve M-ULDs.

The transportation process of perishable goods in the warehouse is mainly divided into three types.

[redacted]

[redacted]

[redacted]

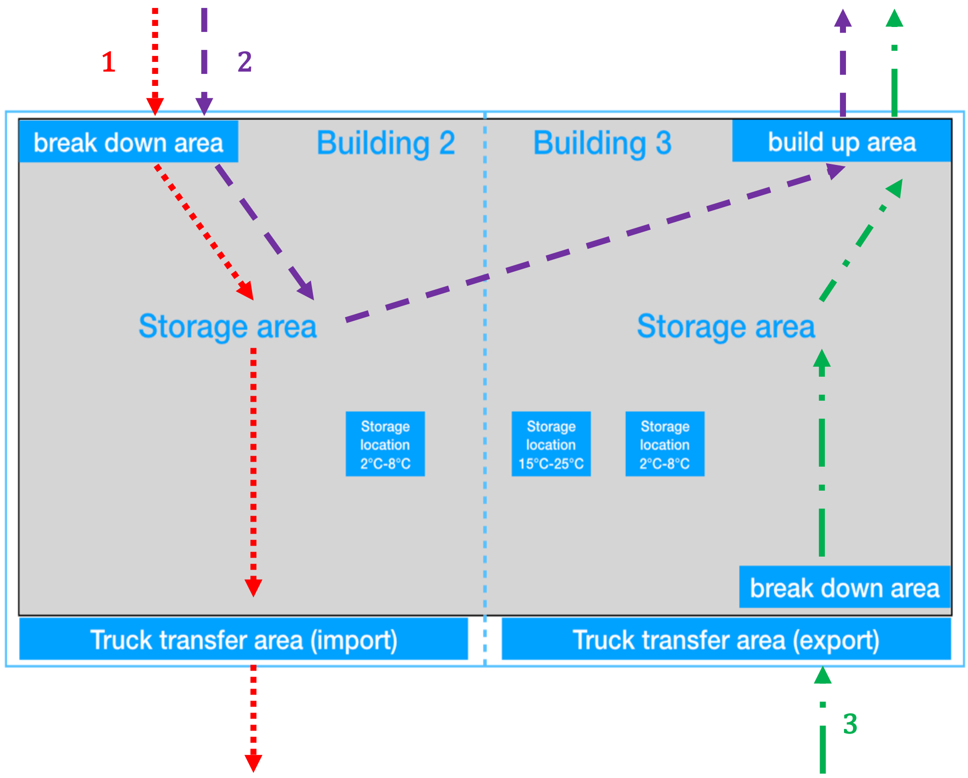


Figure 1. 4 The storage process in the warehouse.

1.2 Problem description

Temperature is one of the key environmental factors affecting perishable goods' quality. It is important that these products are not unnecessarily exposed to too high or too low temperatures for too long. For KLM Cargo, different perishable products need to be stored in different temperature ranges. There are three main types of pharmaceutical products: COL, CRT and PIL, and their corresponding storage temperature range is 2 °C to 8 °C, 15 °C to 25 °C and 2 °C to 25 °C respectively. For fresh products such as flowers, they only have the two product categories of COL and PIL. If the storage time in the warehouse for COL or CRT products is more than eight hours, they will be stored in active and/or passive containers with active temperature control. The temperature inside the container can be adjusted to maintain the specified temperature range.

However, it is not cost-effective to cool perishable products in active and/or passive containers when they are broken down into cases and stored in temporary storage locations. Besides, although these temporary storage locations have heating equipment to prevent the temperature from getting too low in winter, there is no temperature control equipment to ensure the temperature is within the storage temperature range of goods. When the ambient temperature exceeds the storage temperature range, the quality of fresh products in these locations may be affected.

For the actual operation, shipments information is usually available one to three days in advance. KLM employees typically need to decide where to temporarily store incoming shipments before they arrive at the warehouse. If considering the influence of temperature, it is necessary to predict the temperature of different storage locations in the next 1-3 days. Although we now have historical temperature data for different locations of warehouse, no model has been developed to predict indoor temperature based on outdoor weather forecasts for KLM Cargo.

If KLM employees want to reduce the loss of product quality due to temperature, it is important to look for intuitive and easy-to-use metrics to measure the loss of quality. Key Performance Indicators (KPIs) are quantifiable metrics used by organizations to measure the performance of products. For KLM Cargo, the KPIs currently used are time-dependent for all products, so only the impact of time is taken into account. Besides, the KPIs are used differently in the flower and pharmaceutical industries to measure quality loss. Unfortunately, there is no widely used temperature-related KPI to measure the quality loss caused by temperature for different types of products.

The storage location of the warehouse is currently equipped with multiple temperature detectors, which are used to monitor temperature changes in different locations. According to the previous records, the temperature at different locations has obvious differences. Therefore, it is necessary to design a control strategy to find suitable temporary storage locations in advance for goods with different storage conditions, so as to reduce the quality loss caused when the temperature is higher than the recommended storage temperature.

1.3 Research objectives and questions

The main objective of this research is to develop a control strategy to reduce the quality loss when perishable goods (pharmaceuticals and flowers) are outside the storage temperature range in the stage of warehouse storage location choice. To achieve this objective, three sub-objectives are listed:

1. The relationship between weather forecasts (outside temperature, humidity, wind, rainfall, etc.) and the indoor temperature at different storage locations should be established.

Since its better for KLM employees to plan the temporary storage locations of the incoming goods one to three days in advance, we need to predict the temperature of different locations in the warehouse according to the future weather forecast.

2. To establish a control strategy to reduce the total quality loss of perishable goods in the location choice procedure.

Firstly, we need to find a quality loss model that can be used for both pharmaceuticals and flowers. Then, a location choice control strategy will be provided, which is able to measure the quality loss of storing products in temporary storage locations within a specified time. The main objective of the control strategy is to reduce the quality loss of perishable goods outside the storage temperature range in the stage of temporary storage in the warehouse.

3. To do simulation experiments about the control strategy and compare results with the current case.

Simulation experiments will be carried out to show the advantage of the method developed in this project. Here, different storage location choice scenarios are tested to discover the effect of external weather on location choice results. By comparing the total quality loss of perishable goods based on the actual case and the proposed approach, the improvement of the proposed model will be further discussed.

These objectives result in the following research question:

"Given information about perishable goods (pharmaceuticals and flowers) and outside weather forecasts, in what ways KLM employees can better make decisions in the storage location choice?"

The information about perishable goods includes the type and quantity of perishable goods, the optimal storage temperature range and the temporary storage duration. Outside weather forecasts include hourly temperature, humidity, wind, rainfall and other information for the next three to seven days.

To answer the main research question, we defined the following sub-questions.

1. What is the relationship between the outside weather (temperature, humidity, wind, rainfall, etc.) and the indoor temperature at div storage locations?
   - How to do data preparation?
   - What is the result of the analysis?

- Which temperature prediction model will be used?

2. How to design a control strategy to reduce total quality loss?
   - What are the inputs, outputs in the control strategy?
   - How to describe the loss of quality of goods?
   - Based on the answers given above, how to choose the locations of goods to reduce the loss of quality?

3. Given the information of perishable goods and weather forecasts, how much improvement can the control strategy provide according to the simulation results?
   - What assumptions are made in the simulation?
   - What is the best sort criteria for decision support?
   - What is the loss of quality based on the actual case and based on the proposed control strategy?

1.4 Research Outline

The research outline of this paper is shown in Figure 1.5.

The first chapter mainly introduces the research background, describes the problems to be solved, and puts forward the research objectives and questions.

Chapter 2 is a review and summary of past research. In this chapter, we will summarize KPIs commonly used in the flower and pharmaceutical industries as well as the classic quality loss model. They help us to establish the general quality loss model in Chapter 3. We will also review available temperature prediction models and common mathematical programming models in logistics.

In Chapter 3, we will discuss the temperature prediction models used in this project, including the multiple linear regression model and the neural network analysis model. At the same time, according to the characteristics of flowers and drugs, we will design a new quality loss model based on previous studies. Then, control strategy for storage location choice and the simulation method will also be discussed. The control strategy model is based on the quality loss model obtained in Chapter 3 and the indoor temperature prediction model obtained in Chapter 4. Two approaches, namely the operations research model and the flowchart approach, will be introduced as decision support for the control strategy.

In Chapter 4, we will first filter the obtained data after the data collection step. After modelling, we will compare the results of different temperature prediction models, and choose a suitable model for the control strategy design and simulation. Based on the actual operation of KLM cargo, we will simulate the quality loss based on the real situation, and compare it with the proposed control strategy.

Simulation experiments will be carried out in Chapter 5 to show the advantage of the proposed control strategy. In this chapter different storage location choice scenarios are tested to discover the effect of external weather on location choice results. The flowchart method will be used as the decision support for the simulation experiment.

Chapter 6 will answer the research questions mentioned in Chapter 1 and provide recommendations for future research.

For sub-questions, the temperature prediction model in Chapter 4 aims to answer the sub-question 1, the quality loss model and the control strategy design in Chapter 3 aims to answer the sub-question 2, and the simulation results in Chapter 5 will answer the sub-question 3.

Therefore, the following research question is established for this chapter:

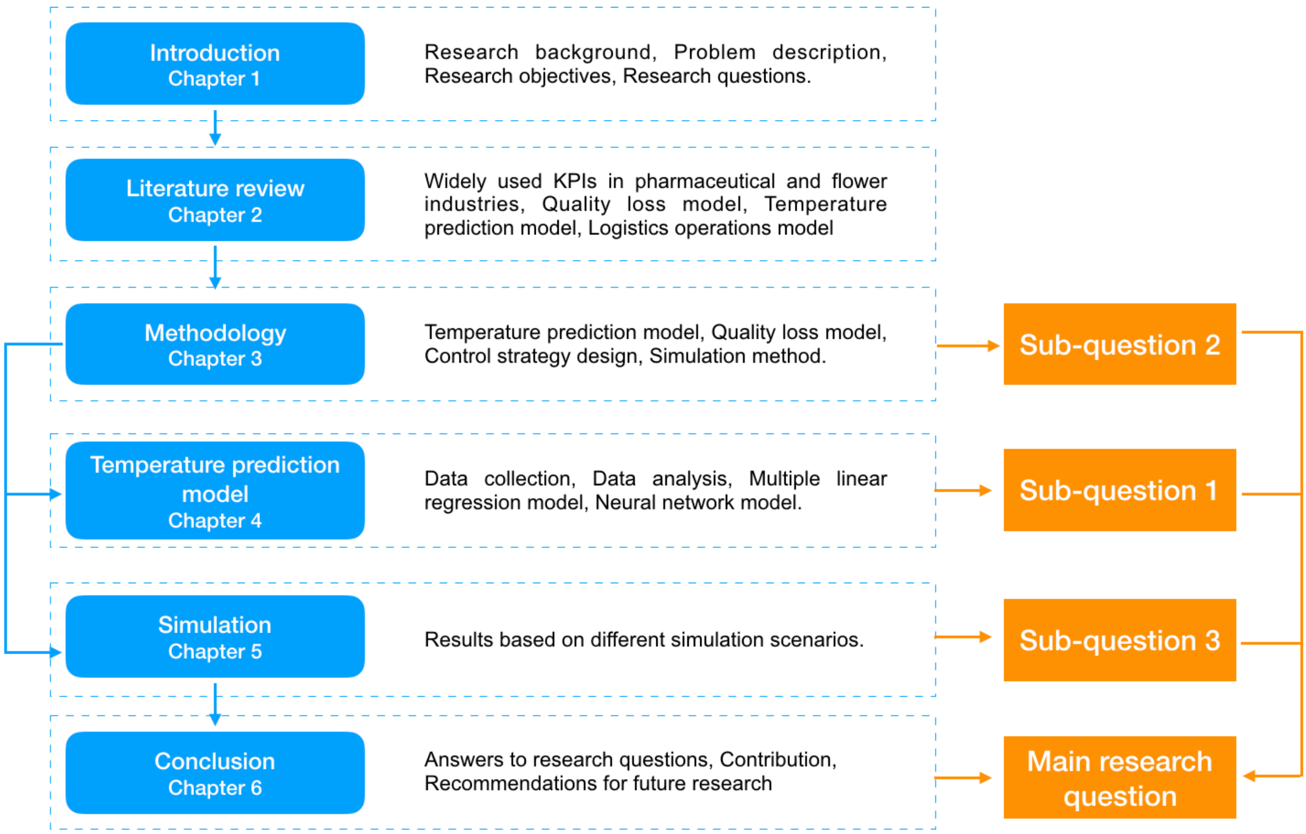


Figure 1. 5 Research approach.

Chapter 2 Literature review

In the literature, it is observed that temperature is one of the key environmental factors affecting perishable goods' quality (Aung and Chang, 2014). For flowers, on the one hand, storage temperature can affect the physiological process, such as accelerating the senescence of flowers and influencing the wound responses of the cut stems; on the other hand, some physical characteristics are also affected by temperature, such as water loss, condensation and drying (Van Meeteren, 2008). For pharmaceutical products, factors such as temperature, time, humidity have a significant impact on the final quality (Kausar Shafaat, Kumar, Hasan, Prabhat, & Yadav, 2013). This section summarizes the commonly used KPIs of quality loss for perishable goods, reviews literature that discusses how flower quality (often measured in terms of "vase life" for flowers) and pharmaceutical quality (often measured in terms of "shelf life")are influenced by temperature, how weather forecast of future temperature may provide insights for indoor temperature, and how this relationship could be used for logistics processes of perishable goods.

2.1 KPIs of quality loss for perishable goods

Key performance indicators (KPIs) are quantifiable metrics used by organizations to measure the performance of products over time. These metrics are used to determine a company's progress and performance regarding its set goals and objectives (Marr, 2012). By presenting KPIs on a live dashboard, every employee in the organisation has insight into the KPIs of the organisation and its current performance. If performance would not be up to the expected standards, employees can adjust their activities immediately. This part discusses the KPIs commonly used in previous studies on flowers and medicines, which can help us to select appropriate indicators when establishing the quality loss model.

2.1.1 Floral KPIs

Some existing KPIs have been widely used to evaluate quality losses for the floral industry. Vase life is leading as a quality indicator and is defined as follows: "the remaining vase life is defined as the time that flowers can be kept on the vase at temperature, which is regularly assumed to be equal to 20 degrees." (Tromp, van der Sman, Vollebregt, & Woltering, 2012). Much previous research has used vase life as an indicator of changes in the quality of flowers. To explore the influence of long-distance transportation on the quality of red anthurium cut flowers, Hettiarachchi and Balas compared the changes of vase life under different storage environments and revealed that floral preservatives and stem-end treatments could better guarantee the quality of flowers (Hettiarachchi & Balas, 2003). Zheng and Guo studied the influence of cerium on the quality of carnation (Zheng & Guo, 2018). They found that cerium helps to regenerate ascorbate and glutathione, thus extending the vase life.

The literature review of KPIs commonly used in flowers was provided in Table 2.1, which has never been done before.

Table 2. 1 Summary of KPIs in previous research.

Variety	Research year	KPIs	Objective	Reference
Carnation, Chrysanthemum, Daffodil, Freesia, Iris, Lily, Rose, Tulip.	1991	Vase life.	To build the mathematical model on the keeping quality of cut flowers.	(van Doorn & Tijskens, 1991)
Aster	2000	Duration of flower development, Flowering, Flowerhead diameter, Number of ray florets of flowerheads.	To study the effect of growth temperature on aster flower development.	(Oren-Shamir, Shaked-Sachray, Nissim-Levi, & Weiss, 2000)
Aster	2002	Height, Flowerhead diameter, Number of buds, Inflorescences.	To study the effect of day length, growth regulators and fertilization on growth and development of Michaelmas Daisy.	(Vršek, Karlović, & Židovec, 2002)
Asteraceae	2002	Respiration, Vase life, Bending score.	To study the effect of storage temperature on cut flowers from the Asteraceae.	(Fisun G. Çelikel & Reid, 2002)
Anthurium	2005	Fresh weight, Vase life, Water uptake, Transpiration, Osmolality, oBrix Values,	To study the effect of stem-end treatments and floral additives on quality.	(Hettiarachchi & Balas, 2003)

		Colour.		
Gerbera	2003	Vase life, Fresh weight, Water uptake, Flowerhead diameter,	To study the effect of storage solution and temperature on quality of cut gerbera. flowers	(Yoo & Kim, 2003)
Carnation, Rose	2004	Vase life, Flowerhead diameter, Fresh weight, Bending score.	To study the effect of red light on longevity and quality of cut flowers.	(Heo, Chakrabarty, & Paek, 2004)
Anthurium	2007	Colour, Fresh weight, Water uptake, Vase life.	To study the effect of ethanol on cut flowers of anthurium.	(Kanlayanarat, Thawiang, & Buanong, 2007)
Unspecified	2008	Senescence rate, Vase life,	To analyse the causes of quality loss of cut flowers.	(Van Meeteren, 2008)
Tree peony	2009	Vase life, flower opening index, fresh weight loss, electrical conductivity (EC), malondialdehyde (MDA).	To study the effect of storage temperature on quality of cut flowers.	(J.-X. Wang, Xia, Zhang, & Zhou, 2009)
Dahlia	2015	Vase life, Appearance of leaves on the stem	To study the effect of preservation solutions on morpho-decorative characteristics of Dahlia.	(CIOBANU, PUI, BAIA, HUSTI, & CANTOR, 2015)

2.1.2 Pharmaceutical KPIs

There are some existing industry-wide key performance indicators, which have different applications among different stakeholders in the pharmaceutical industry. According to previous research by KLM Cargo, three temperature-related KPIs that are either in use or under development. These KPIs are:

-Time out of Refrigeration;

-The Mean Kinetic Temperature;

-Kelvin hours.

According to KLM Cargo's documents, all pharmaceutical producers indicated that they are currently using the time out of refrigeration to ship or develop their products. During transport, it is not immediately an issue when a product is out of refrigeration, as long as it does not exceed the maximum time out of refrigeration determined by the shipper.

As for the mean kinetic temperature, Tong defined it as a single temperature at which the total amount of degradation of a drug over a given period of time is equal to the sum of the individual degradation that occurs at different temperatures (Tong & Lock, 2015). In the calculation of mean kinetic temperature, the stability data of specific pharmaceutical products play an important role and need to be obtained by pharmaceutical manufacturers.

The kelvin hour a currently under development parameter, which combines the consideration of both time and temperature. The calculation of this indicator is as follows:

$$Kelvin\ hours = (T_A - T_{\mathrm{i}}) * D \tag{2.1}$$

Where T_A is the (average) ambient temperature during transport; T_i is the internal temperature of the shipment (median of the booked temperature product); D is the duration of the transport in hours.

Besides, shelf life is an important KPI of pharmaceutical products, which can be defined as the time when pharmaceutical products can maintain its essential performance characteristics under specific storage conditions (Anderson & Scott, 1991). This index has been widely used in previous studies for various types of pharmaceutical products, such as N-acetylcysteine, vitamin C, ascorbic acid, metronidazole infusion and isotretinoin products (Jaworska, Szulińska, Wilk, & Tautt, 1999; Moussa & Haushey, 2015; Shaheen, 2005; Taylor & Keenan, 2006).

2.2 Quality loss model

This part studies the product quality loss models commonly used in the past, which mainly monetize quality loss. We also found temperature-based quality loss functions for flowers and medicines. By combining and adjusting these models, we can find a quality loss model that can be used in this project.

2.2.1Quality loss function

Goalpost approach is the most primitive quality loss model. Its use can be traced back to the quality control of television products by Sony-USA (Ealey, 1988). For this company, whether the products produced to exceed the specification is an indicator of the target quality loss. Products are good when their performance is within the limits of the upper and lower specification, and bad when they are outside the range of specifications (Fig2.1). However, from the customer's point of view, products that meet the specification are as good (or bad) as those that slightly exceed the specification. It indicates that for goalpost approach, the measuring system of quality loss is quite different from reality and cannot describe the quality loss of products well.

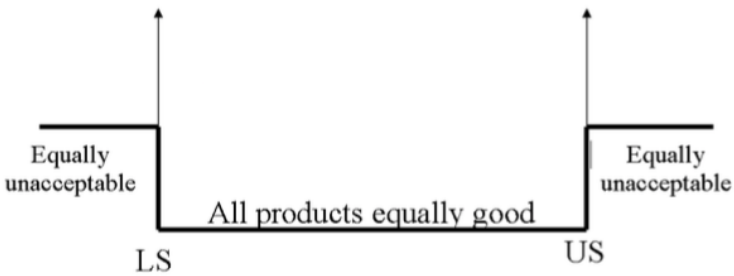


Figure 2. 1 Goalpost approach (Sharma, Cudney, Ragsdell, & Paryani, 2007)

To make the quality loss model consistent with the facts, Japanese engineer Genichi Taguchi introduced a new concept of a loss function, which is an effective method for quality engineering. Taguchi proposed three types of quality loss functions. These three types of functions are all based on the continuous quadratic relation, and the formula is as follows:

Nominal-is-best loss function:

$$L(y) = k(y - m)^2 \tag{2.2}$$

Where $L(y)$ is the quality loss related to a particular value of the critical performance parameter y, m is the nominal value of the parameter specification; k is the average loss coefficient, and its value is a constant, which depends on the cost at the specification limits and the width of the specification (e.g., $\mathrm{m} \pm \Delta$).

Smaller-is-better loss function:

$$L(y) = k.(y)^2, k = A/\Delta^2 \tag{2.3}$$

Higher-is-better loss function:

$$L(y) = k/(y)^2, k = A.\Delta^2 \qquad (2.4)$$

where A is the total quality loss of products, Δ is the customer's tolerance. All other variables are the same as those in the nominal-is-best loss function.

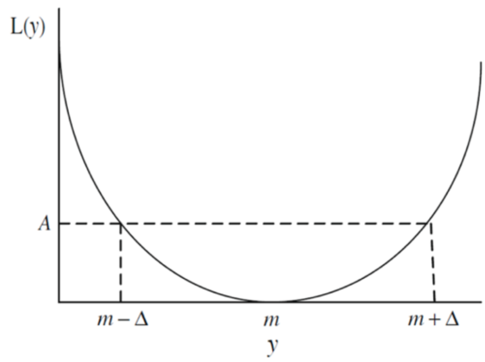


Figure 2. 2 Nominal-is-best loss function (Kumar et al., 2012).

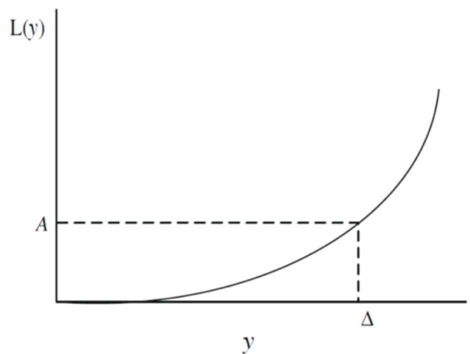


Figure 2. 3 Smaller-is-better loss function (Kumar et al., 2012).

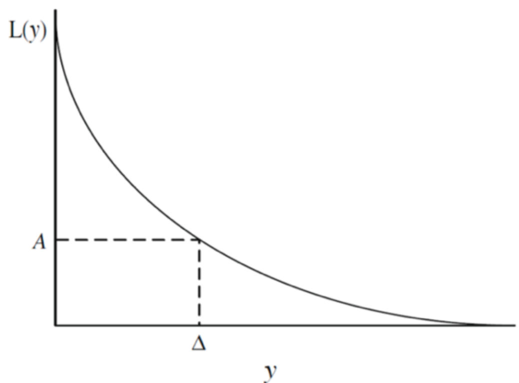


Figure 2. 4 Higher is-better loss function(Kumar et al., 2012).

Taguchi methods have been widely used to evaluate and improve the quality of products and have achieved great success in many industries.

2.2.2Foral quality loss model

In Section 2.1.1, we discussed KPIs commonly used to measure the quality of flowers. Among them, the vase life is one of the most commonly used in research and practical applications.

The relationship between temperature and vase life of flowers has been mentioned in some previous studies. (Cevallos & Reid, 2001) tested the relationship between vase lives and transportation temperature in seven varieties of flowers: carnations, daffodils, iris, Kilian daisies, paperwhite narcissus, roses and tulips. He found the decreased vase lives with increasing storage temperature. Besides, the effect of storage temperature was different even for closely related species. To study the effect of temperature on vase life and respiration rate, Çelikel and Reid conducted a series of experiments on five common flowers over a decade, starting in 2002 (F. G. Çelikel & Reid, 2004; Fisun G. Çelikel, Cevallos, & Reid, 2010; Fisun G. Çelikel & Reid, 2002). The results showed that the respiration of all flowers increased exponentially with the increase of storage temperature. Also, there was a significant (negative) linear correlation between respiration rate and vase life. In their experiments, all flowers stored at 0°C have the largest vase life. The relationships between the three variables were expressed as functions shown below.

Gypsophila：

$$Respiration\ rate = 14.333 \cdot e^{0.1092 \cdot Storage\ temperature} \tag{2.5}$$

$$Vase\ life = -0.1916 \cdot Respiration\ rate + 15.422 \tag{2.6}$$

Snapdragon:

$$Respiration\ rate = 29.168 \cdot e^{0.1009 \cdot Storage\ temperature} \tag{2.7}$$

$$Vase\ life = -0.0827 \cdot Respiration\ rate + 12.190 \tag{2.8}$$

Some other KPIs have also been used in research to illustrate the relationship between flower quality and temperature. By parameterizing for specific flowers based on data from previous literature, Van Meeteren from Wageningen University set up a model of the relationship between flower senescence-rate and temperature (Van Meeteren, 2008). This model can generally be described as follows:

$$k_s = \frac{k_{max}}{1+10^{(T_{half}-T_s).slope}} , \tag{2.9}$$

in which T_s is the storage temperature (°C), k_s is the senescence rate at temperature T_s, k_{max} is the maximum senescence rate, T_{half} is the temperature at which the senescence rate is half of k_{max}, the slope is the steepness of the curve.

These previous studies can help us better understand the relationship between temperature and the quality of flowers, and provide an important basis for us to determine the storage temperature of common varieties of flowers.

2.2.3Pharma quality loss model

For pharmaceutical products, previous academic studies have focused on shelf life as an indicator to measure drug stability (Jaworska et al., 1999; Taylor & Keenan, 2006). Van't Hoff and Lehfeldt proposed the Q10 model, which formulates the relationship between pharmaceutical shelf life and storage temperature (Van't Hoff & Lehfeldt, 1899). It can generally be described as follows:

$$S_{(T2)} = \frac{S^{(T1)}}{Q10^{\left(\frac{T2-T1}{10}\right)}} \tag{2.10}$$

In which $T1$ is the optimal storage temperature, $S(T1)$ is the shelf-life at $T1$, $T2$ is the actual storage temperature, $S(T2)$ is the estimated shelf life at $T2$ and $Q10$ is the reaction rate. The value of Q10 is usually evaluated at 2, 3, or 4. Generally, if Q10 is 2, the results tend to be conservative, that is, assume that the stability of pharmaceutical products is sensitive to changes in temperature; when Q10 is 3, the results are considered relatively reliable; while a pessimistic estimate of Q10 equals 4 (London & East, 2001).

Q10 model has been widely used to predict the shelf life of different drugs in previous studies. Anderson did an accelerated stability test and evaluated the shelf life of five major drugs: morphine, 11-nor-i-9-tetrahydrocan-nabinol-9-carboxylic acid, amphetamine, phencyclidine, and benzoylecgonine (Anderson & Scott, 1991). With the help of Q10 method, he found out the optimal activation energy values to lower the risks inherent in projecting product shelf life. To study the accelerated stability of metronidazole injection (per 100ml), Shaheen evaluated its physical and chemical properties at 37 °C and 50 °C for 90 days (Shaheen, 2005). Using the Q10 method, he estimated the shelf life of metronidazole injection at different temperatures. Finally, he concluded that the Q10 method was significantly effective in evaluating the shelf life of the drug substances in solution because their degradation was a temperature-sensitive process. Moussa compared the shelf life of vitamin C/Ascorbic Acid at 37 °C and 45 °C based on Q10 model (Moussa & Haushey, 2015). He concluded that this model is considered to be a cornerstone in predicting the storage life of heat-sensitive degraded drugs.

2.3 Temperature prediction model

The temperature prediction of storage locations is also one of the focuses of this project, which includes the temperature data filtering and the establishment of the temperature relation model.

2.3.1 Data filtering

Data filtering step is required due to the consideration of checking outliers and missing temperature data. An outlier was defined by the statistician Hawkins as an observation so different from other observations that it was suspected to be produced by a different mechanism (Hawkins, 1980). The ROADIDEA institute of the European Commission has been working on temperature data filtering for many years (Cordis, 2019). In their report, the mean and standard deviation value of 20 temperature samples over for 55 days were calculated. Samples exceeding the threshold that greater than or less than the daily average value five times the standard deviation will be considered outliers and removed from the data set. This threshold is consistent with the study of Asumadu-Sakyi, who introduced a temperature data filtering threshold in his data preparation stage, in which abnormal temperature data was defined as five times or less the standard deviation than the mean outdoor temperature of the day (Asumadu-Sakyi et al., 2019).

2.3.2 Temperature prediction

Previous studies have provided three main methods for temperature prediction, namely, time series analysis model, regression analysis model and neural network analysis model.

According to Reppo, the relationship between indoor-outdoor temperature in winter was fitted by quadratic polynomial formula, while the relationship in summer was fitted by exponential function formula (Reppo, Mikson, & Vaarak, 2004). His model simply includes two parameters: indoor and outdoor temperature.

By calculating the Pearson correlation coefficients between indoor and outdoor daily temperature averages and comparing these coefficients with three mathematical models, Nguyen found the association was strong at warmer outdoor temperature while it was weak at a cooler temperature (J. L. Nguyen, Schwartz, & Dockery, 2014). This finding was consistent with a study done in Japan that same year (Saeki et al., 2014). Taking into account the impact of seasons and region differences, Nguyen identified a sample from eight regions around the world in 2016. The result of the new sample was consistent with his previous study in 2014, and he pointed out that the relationship between indoor temperature and outdoor temperature varied across seasons and locations (Nguyen and Dockery, 2016).

To study the temperature prediction in the greenhouse, Kim did a comparison of the Artificial Neural Network (ANN) model, the Multiple Regression Model and the Recurrent Neural Network model (S. Y. Kim et al., 2018). By comparing the root mean squared errors, it was found that the result from the ANN model was more accurate than other prediction models.

With the help of the Support Vector Machine (SVM) method, Xing built a data-driven model for soil temperature prediction (Xing et al., 2018). Air temperature, solar radiant, wind speed and relative humidity and time are selected as the main inputs after calculating the correlation coefficient indicates. The innovation of this model is that it combines the daily soil temperature variations (short-term impact) and annual average ground temperatures predictions (long term impact).

Asumadu quantified the impact of outdoor on indoor temperature during the warm season and cool season, and established the Multiple Regression Model, which is described as follows (Asumadu-Sakyi et al., 2019):

$$I_{i,t} \sim \beta_{i,0} + \beta_{i,1} O_{i,t} + \sum \alpha_j X_{i,j} + \varepsilon_{i,t} \tag{2.11}$$

in which $I_{i,t}$ is the indoor temperature (°C), $\beta_{i,0}$ is random intercept (°C), $\beta_{i,1}$ is the random slope, $O_{i,t}$ is the outdoor temperature (°C), α_j is the estimated effect of the building characteristics, $X_{i,j}$ is building characteristic, $\varepsilon_{i,t}$ is error. This model takes into account the effects of other modifying factors on indoor temperatures, such as the heating system, the building material, the roof material and the age of the house.

2.4 Operations research model for logistics

In recent years, more and more attention has been paid to the logistic scheduling for perishable products in the literature due to its strategic importance. According to the climate change, a mixed-integer linear programming model is formulated by Accorsi to plan the production, storage and distribution operations of perishable goods (Accorsi, Gallo, & Manzini, 2017). Due to the consideration of responsiveness and environmental impacts, Musavi designed a hub scheduling model to plan the distribution chains of perishable goods (Musavi & Bozorgi-Amiri, 2017). The main advantage of this model is that it simultaneously achieves three prominent objectives: optimizing the total transportation cost, the total perishability of the product and the total carbon emission of vehicles. Kim proposed a two-phase location-allocation model to find the optimal locations of biomass storage facilities (S. Kim, Kim, & Kiniry, 2018). To minimise the total weighted inventory cost and transport cost, Ji built a three-stage batching and scheduling model involving both batch supply and batch delivery (Ji et al., 2018). His model fully considered the three stages of a supply chain: a supplier, a manufacturer, and a customer. However, the input parameters in the above model did not have the temperature, and most of them considered the entire stage of a supply chain to minimise costs. This project only focuses on the scheduling for internal logistics, which includes one party or actor of supply chains.

2.5 Research gaps

Although the pharmaceutical and flower industries have some existing KPIs to measure quality loss. There is no universal standard for measuring different kinds of products. As for flowers, many KPIs cannot be quantified, such as flower colour and bending score. Therefore, in this project, it is necessary to find a genetic temperature-based KPI that can be used for both drugs and flowers to measure their quality loss.

A lot of previous research has focused on predicting temperatures, but their models' temperature predictions are only accurate to the day. For instance, the temperature features of Afroz's and Xing's model are chosen as the average daily temperature or the highest and lowest temperature (Afroz, Urmee, Shafiullah, & Higgins, 2018; Xing et al., 2018). However, our project needs to predict the temperature at hourly level. Moreover, in previous studies, the prediction models were mainly applied in residential buildings or greenhouses. No paper has predicted the temperature of industrial buildings such as warehouses.

For the temperature prediction models used in the past, although the results of different prediction models (mainly neural network model and multiple linear regression model) have been compared, there was no study on the validation of different prediction models with the same approach.

Although previous literature has discussed some temperature-based quality loss models for perishable supply chain, these models are usually limited to one kind of products. No generic model has been applied to measure the quality loss for different types of perishable products of different industries. In addition, most of location choice models are aimed at reducing storage and transportation costs, while few studies have involved reducing product quality losses.

Chapter 3 Methodology

The control strategy consists of three models: indoor temperature prediction model, quality loss model and location choice model. This chapter introduces all the models and the simulation approach used in this project. The aim of this chapter is to answer the second sub-question and provide methodological support for the first sub-question. The second sub-question is as follows:

How to design a control strategy to reduce total quality loss?

- What are the inputs, outputs in the control strategy?
- How to describe the loss of quality of goods?
- Based on the answers given above, how to choose the locations of goods to reduce the loss of quality?

For the indoor temperature prediction model, Section 3.1 introduces and compares the traditional multiple linear regression model with the neural network model. Section 3.2 presents a new temperature-based quality loss model for both flowers and medicines. Section 3.3 describes the input and output parameters of the control decision and describes the decision support for optimizing the storage location choice of the goods. This chapter also introduces the model validation method and the simulation experiment approach.

3.1 Inputs and outputs

The whole control strategy consists of three models, including the indoor temperature prediction model, product quality loss model and location choice model. The inputs and outputs for each model are shown in Table 3.1.

Table 3. 1 Inputs and outputs of the control strategy

Model	Inputs	Outputs
Indoor temperature prediction model	Outdoor temperature, Wind speed, Precipitation amount, Air pressure, Cloud cover, Humidity, Rainfall or not, Thunder or not.	Hourly indoor temperature at different locations of the warehouse.

Quality loss model	Product category (flower or pharmaceutical products), Arrival and departure time, Storage temperature range, Product value and quantity, Hourly storage temperature.	Estimated quality loss of product per hour at different locations.
Location choice model	Estimated quality loss per hour for each product at different locations, Simulation time, The capacity of each location (in skids)	Storage location-allocation results, Estimated total quality loss of all goods.

Input parameters for the indoor temperature prediction model can be obtained from mainstream weather prediction websites. For example, https://www.wunderground.com, https://www.yr.no. The output of this model provides data for the quality loss model.

The input parameters of quality loss model mainly come from KLM cargo. For the actual operation, shipments information is usually available 1-3 days in advance. The information includes the estimated time of arrival and departure at Schiphol airport, the type and quantity of shipments. There are three main types of pharmaceutical products: COL, CRT and PIL, and their corresponding storage temperature range is 2 °C to 8 °C, 15 °C to 25 °C and 2 °C to 25 °C respectively. For fresh products such as flowers, they only have the two product categories of COL and PIL. The value of the product will be estimated based on past claims records related to pharma and fresh goods.

When making the storage location selection, the time length of the model needs to be set. In this case, we set it as the next 72 hours. The capacity of each storage location is estimated by KLM Cargo.

After combining the three models, the final output is the storage location of each item and the total product quality loss.

3.2Temperature prediction model

In this chapter, the indoor temperature prediction models that have been used for this project are described from a technical point of view. The multiple linear regression model and the neural network model for indoor temperature prediction will be introduced.

3.2.1Multiple linear regression model

A multiple linear regression model is carried out to describe the relationship between more than one independent variable X and a dependent variable Y. The relationship can be represented by the following equation:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + e \tag{3.1}$$

In this equation β_0 is the constant term; β_1 to β_2 are the coefficients which represent the slopes of a plane in three-dimensional space (X_1 , X_2 , Y).

R is used in this project based on the stepwise regression approach. Stepwise regression is a combination of forward and backward selection techniques. After each step of adding variables, all candidate variables in the model are examined to see if their significance has dropped below the specified tolerance level. Non-significant variables are removed from the model step by step. Finally, only the significant variables are brought into the model.

3.2.2 Neural network analysis model

With the development of artificial intelligence, machine learning has been applied in more and more subjects in recent years. As a subset of machine learning, Artificial Neural Network model can simulate complex and non-linear relationship between input and output streams based on mathematical regression (Yang & Chen, 2015).

Layer-wise organization

A typical ANN architecture is shown in Fig.3.1, which consists of three layers: input layer, hidden layer and output layer. Input parameters are stored in the input layer and fed into the system for processing. The role of hidden neurons is to intervene in a useful way between external input and network output. This layer cannot be accessed by developers during processing (Haykin, 1999). According to previous studies, when the number of input data is n, the number of hidden layers is usually set as n/2, n, 2n, 2n+1 (S. Y. Kim et al., 2018). Result parameters are included in the output layer depending on the purpose of the system.

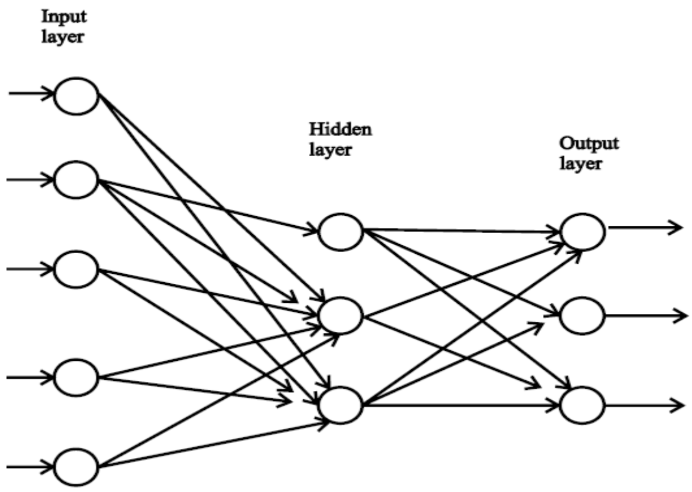


Figure 3. 1 A typical 2-layer Neural Network(Mehta, Mehta, Manjunath, & Ardil, 2008)

Mackay and Foresee introduced a function description of two-layer neural network, which is given by (Foresee & Hagan, 1997; MacKay, 1992):

$$y_i = g(x_i) + e_i = \sum_{k=1}^{s} w_k g_k (b_k + \sum_{j=1}^{p} x_{ij} \beta_j^{[k]}) + e_i \, , i = 1, \dots, n \qquad (3.2)$$

Where:

$e_i \sim N(0, \sigma_e^2)$,

s is the number of neurons,

w_k is the weight of k-th neuron, $k = 1, \dots, s$.

b_k is a bias for the k-th neuron, $k = 1, \dots, s$.

$\beta_j^{[k]}$ is the weight of the j-th input to the net, $j = 1, \dots, s$.

$g_k(.)$ is the activation function, in this implementation $g_k(x) = \frac{\exp(2x)-1}{\exp(2x)+1}$.

Based on this function, the “brnn package” of R programming software can be used to build neural networks. This package was developed based on Nguyen algorithm to assign initial weights and was optimized by Gauss-Newton algorithm (D. Nguyen & Widrow, 1990). The optimization function of R will minimize:

$$F = \beta E_D + \alpha E_W \qquad (3.3)$$

Where:

$E_D = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$, i.e. the error sum of squares,

E_W is the sum of squares of network parameters (weights and biases),

$\beta = \frac{1}{2\sigma_e^2}$,

$\alpha = \frac{1}{2\sigma_\theta^2}$, σ_θ^2 is a dispersion parameter for weights and biases.

3.2.3 Comparison of two models

After modelling, the overall significance of the model should be checked to select the most suitable one. Root mean squared error (RMSE), which is used to measure the deviation between the observed value and the true value. RMSE is expected to be close to 0, which means good model selection and fitting, which means good model selection and fitting. The calculation method of this indicator is as follow:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y' - y)^2}, \qquad (3.4)$$

where y' and y are predicted and observed value.

3.2.4 K-fold cross-validation

In the process of machine learning, the final model usually overfits the training data. Especially for the model with a large number of parameters, the overfitting problem is more pronounced. As a result, the estimated error rate is usually too optimistic (lower than the true error rate). Therefore, the cross-validation method is introduced to prevent overfitting due to model complexity and estimate the skill of machine learning models.

K-fold cross-validation method is to divide a given data set into K folds, where each fold is used as a testing set at a certain point (James, Witten, Hastie, & Tibshirani, 2013). The general procedure of this method is as follows:

1) Shuffle the dataset randomly.

2) Divide the dataset into k groups.

3) For each unique group: The group is divided into test data set and training data set. Then the model was fitted on the training set and evaluated on the test set. After that, the estimated error (MSE) is retained.

4) Repeat the process k times and calculate the average MSE.

As the test samples are independent of the training data, the results derived from this 10-fold cross-validation are reliable. For the number of folds, in practice, the choice of the number of folds depends on the size of the dataset. Typically, k =10 is chosen in the field of applied machine learning (Refaeilzadeh,2009). In order to determine the accuracy range of the model, the 10-fold cross-validation results can be represented by the box plot.

3.3 Quality loss model

Here, two independent variables: time and temperature are contained in the quality loss model. That is, we assume that product quality loss is only related to these two factors. According to previous studies, temperature and time are the two main factors that have the greatest influence on perishable products. Besides, the main purpose of our project is to reduce the quality loss caused by temperature, so the quality loss caused by other factors is ignored in this study.

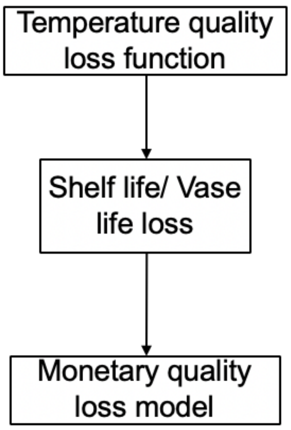


Figure 3. 2 The process of quality loss calculation.

The process to evaluate the quality loss is described in Figure 3.2. The final quality loss model is composed of the time-money quality loss model and the temperature-time quality loss functions. For the temperature-time quality loss functions, according to the relationship between temperature and flower vase life/ medicine shelf life, the hourly vase life/shelf life loss at a given temperature can be estimated. After that, we end up converting total vase life/shelf life loss to money, based on the time-money quality loss model.

3.3.1 Time-money quality loss model

Based on the smaller-is-better loss function developed by Taguchi, a time-based quality loss model has been built.

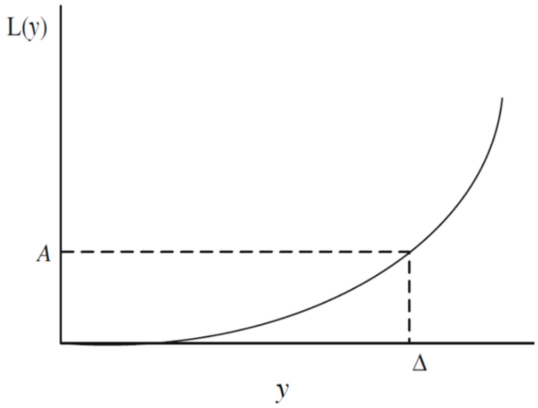


Figure 3. 3 Smaller-is-better loss function(Kumar et al., 2012).

$$L(y) = k.(y)^2, k = A/\Delta^2 \tag{3.5}$$

Where $L(y)$ (€/unit) is the monetary loss caused by 100% product quality loss per unit of products; y (hours) is the loss of vase life for flowers and the loss of shelf life for pharma; Δ (hours) is the maximum vase life for flowers or maximum shelf life for pharma and A (€) is the price per unit of products.

3.3.2 Floral vase life loss function

According to Çelikel's experimental conclusion mentioned in Section 2.2.2, we can calculate the quality loss caused by deviation from the optimal storage temperature.

Take gypsophila as an example. In her experiments, flowers stored at 0°C have the largest vase life. The average maximum vase life for gypsophila is known to be 304.22 hours (F. G. Çelikel & Reid, 2004). When the storage temperature deviates from the optimal storage temperature ΔT,

The corresponding vase life is：

$$Vase\ life_{\ (\Delta T)} = 370.13 - 65.91 \cdot e^{0.1092 \cdot \Delta T}\ (hours) \tag{3.6}$$

The corresponding vase life loss is:

$$Vase\ life\ loss_{(\Delta T)} = \text{Maximum vase } life - vase\ life_{\ (\Delta T)}$$

$$= -65.91 + 65.91 \cdot e^{0.1092 \cdot \Delta T}\ (hours) \tag{3.7}$$

The hourly vase life loss is:

$$Hourly\ vase\ life\ loss_{(\Delta T)} = Vase\ life\ loss_{(\Delta T)} / Vase\ life_{\ (\Delta T)} \tag{3.8}$$

3.3.3 Pharma shelf life loss function

The pharma shelf life loss function was established according to Q10 model. Here, a drug is assumed to be moderately sensitive to temperature. Its shelf life is $S(T1)$ (unit: hour) when the storage temperature is $T1$. When the storage temperature is $T2$ ($\Delta T = T2 - T1$), the corresponding shelf life is:

$$S_{(\mathrm{T2})} = \frac{S^{(T1)}}{3^{\left(\frac{T2-T1}{10}\right)}}\ (hours) \tag{3.9}$$

The corresponding shelf life loss is:

$$Shelf\ life\ loss_{\ (\Delta T)} = S(T1) - S(T2)\ (hours) \tag{3.10}$$

The shelf life loss per hour at $T2$ is:

$$Hourly\ vase\ life\ loss_{(\Delta T)} = Shelf\ life\ loss_{\ (\Delta T)} / S_{(\mathrm{T2})} \tag{3.11}$$

3.3.4 Example of the quality loss model

By combing the time-money quality loss model and vase life/ shelf loss function, we can calculate the quality loss (€) of shipments in different temperature deviation.

Here we assume that the optimal storage temperature of a drug is 8 degrees Celsius and the corresponding shelf life is 180 days (4,320 hours). The value of the pharmaceutical products per skid is assumed as 100,000 € and the value of flowers per skid is assumed as 5,000 €.

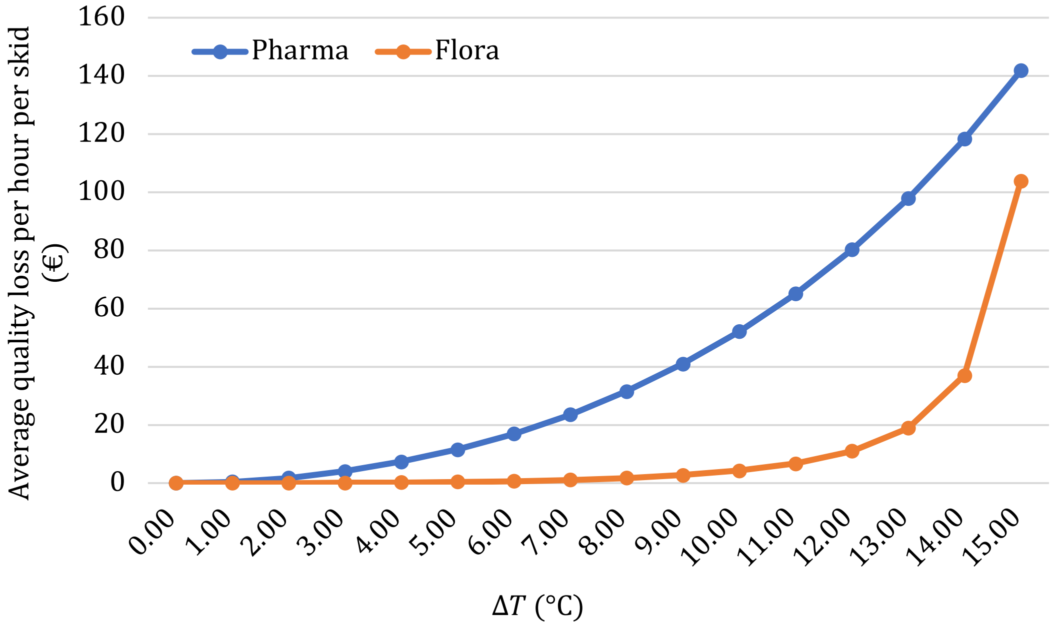


Figure 3. 4 Average monetary quality loss per skid per hour.

The figure above shows the monetary vase life loss per hour per skid as the difference from the optimal temperature increases. It can be found that when the temperature deviation is large, the quality loss per hour increases faster. The quality loss for ΔT=10 °C (0.03 €/skid/hour) is ten times as great as that for ΔT=1 °C (0.001 €/skid/hour). For floral products, although their value is much smaller than that of pharmaceutical products, the temperature has a greater impact on their quality loss due to their short vase life. Especially when ΔT is higher than 10°C, the quality loss per skid per hour for flowers increases exponentially. While the quality loss curve of pharmaceutical products is close to is closer to the exponential growth. The biggest difference between the two types is 0.28 €/skid/hour at ΔT=15 °C.

3.4 Control strategy & Simulation approach

This section will introduce two temperature-based location choice models. One is the operations research l model, which can quantify the total quality loss of products and find out the optimal solution. The other is the flowchart approach for decision making, in which products are ordered according to a certain sort criterion.

3.4.1 Operations research model

This section proposes a quality loss model for perishable goods storage at different locations in the warehouse. This model is an operations research approach. It requires a lot of modelling and solving efforts, which is out of the scope of my contribution. I will only introduce the possibility and potential benefit of this method, but will not go into details. Instead, I propose a more straightforward method with a flowchart in section 3.4.2. And that method will be tested with the Monte Carlo simulation in Chapter 5.

Before formulating the model, we first introduce relevant parameters.

n: *locations for storage in the warehouse*, $n = 1,2,3 \dots N$;

m: *shipments*, $m = 1,2,3 \dots M$;

q_m^n: *loss of quality when the* m_{th} *shipment is stored at the* n_{th};

l: *the amount of shipments*;

t: *simulation time* (*divided by hours*), $t = 1,2,3, \dots T$;

TEM_t^n: *indoor temperature at time t for the location* n;

$TEMM_m$: *maximum storage temperature for shipments* m;

The storage location of shipment m is represented by a series of binary variables x_m^n, denoted as:

$$x_m^n = \begin{cases} 0 & m \text{ is not stored at location } n \\ 1 & m \text{ is stored at location } n \end{cases} \tag{3.12}$$

For instance, x_{12}^3=1 represents that shipment 12 is stored at location 3. It is logical to have the constraint that each shipment can only be stored at one location of the warehouse. For instance, the first shipment should have the constraint that:

$$\sum x_1^n = 1, \forall n \in N \tag{3.13}$$

The q_m^n (hourly quality loss of pharmaceutical products and flowers) can be calculated according to formula (3.10) ~ (3.15). The total quality loss of the first shipment can be written as follows:

$$Q_1 = \sum_{t=1}^{T} \sum_{n=1}^{N} x_1^n q_1^n(t) \tag{3.14}$$

The total quality loss for all shipments is:

$$Q_{total} = \sum_{t=1}^{T} \sum_{n=1}^{N} \sum_{m=1}^{M} x_m^n q_m^n(t) \tag{3.15}$$

The constraints of this mathematical model include the maximum capacity (in skids) for each storage location in addition to formula (3.17). It should be noted that the model does not take into account labor costs and other limiting factors. As more practical factors are taken into account, the model will be closer to the actual operation.

In order to find the optimal location-allocation through this model, we need to minimize the total quality loss:

$$Min\ Q_{total}$$

3.4.2 Flowchart for decision making

Another location choice approach with a flowchart is proposed. The flowchart for selecting storage locations of shipments is shown below. For the m_{th} shipment, after checking the storage capacity of each location, we will screen out all available storage locations. The quality loss at each location for this shipment is then calculated and the one with the least value is selected. After this, the model updates the storage capacity information for each location and begins to select storage locations for the $(m+1)_{th}$ shipment.

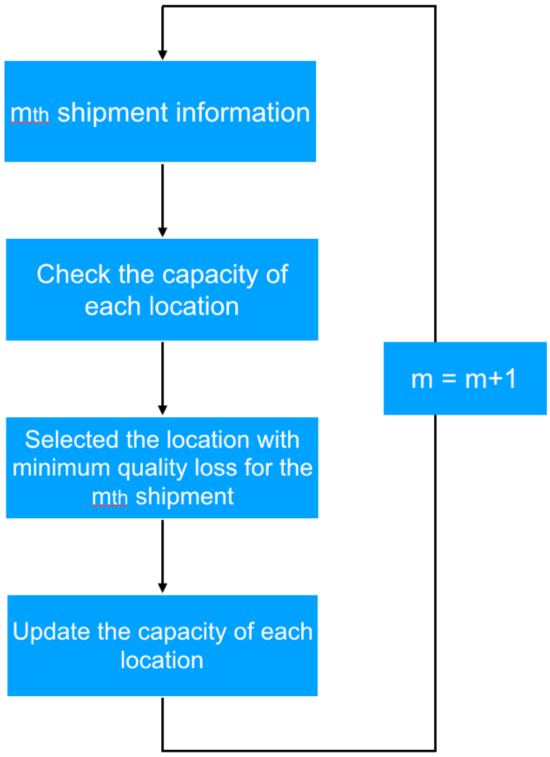


Figure 3. 5 The flowchart approach for storage location choice.

For this model, we need to determine in advance the order in which shipments are selected for storage. Sorting rules and simulation results of this model are discussed in Chapter 5.

3.4.3 Monte Carlo simulation approach

The Monte Carlo approach involves repeated simulations of samples. By using probability distributions, we can understand the range of potential outcomes of a given variable and the likelihood that different values represent the true value.

Before running the model, it is necessary to specify the number of simulations. Simulations can be stopped when the measurement of interest (median or mean) becomes stable and the frequency distribution histogram is smooth. According to previous research, the general rule of thumb is to use 10,000 simulations (McMurray, Pearson, & Casarim, 2017).

3.5 Discussion

For temperature prediction model, multiple linear regression model and neural network analysis model have their advantages and disadvantages. The use of multiple linear regression models needs to satisfy certain basic assumptions. For non-linear factors, they cannot be included in the model. But it has the advantage of being easy to interpret and understand, especially for model correction and updating. In contrast, the neural network analysis model is like a black box but it can capture complex relationships between inputs and outputs as well as doesn't have as many usage conditions as traditional models.

The quality loss model is based on the vase life of flowers and the shelf life of pharmaceutical products. It mainly consists of two steps: first, calculate the vase life loss/shelf life loss results from the storage temperature, and then convert the total vase life loss/shelf life loss into money.

The operations research model describes the functional relationship between each variable and the total quality loss. With the objective function, we can select the best element from some set of available alternatives. This model is widely used and is an effective tool to find the optimal solution. However, factors taken into account in the model have a great influence on the reliability of the results. In this project, labour costs, warehouse operating costs and other factors were not taken into account. Therefore, this model is too simple and worthy of further study and optimization in the future. In this project, we will use the flowchart approach in Section 3.3.3 as the decision support and conduct simulation experiments in Chapter 5 with different sort criteria to find the best ones for control strategy.

Chapter 4 Temperature prediction model

In this chapter, we will use multiple linear regression model and neural network analysis model to predict the temperature of different points in the warehouse. Before modelling, we need to process and filter the collected data. After that, the RMSE value will be used to evaluate the fitting degree of models. Finally, we will select the best model and apply it to the simulation in Chapter 5. Therefore, the purpose of this chapter is to answer the following questions:

What is the relationship between the outside weather (temperature, humidity, wind, rainfall, etc.) and the indoor temperature at div storage locations?

- How to do data preparation?
- What is the result of the analysis?
- Which temperature prediction model will be used?

4.1Data preparation

This section introduces the basic information of indoor and outdoor weather data collection, data processing, and the final data set.

4.1.1 Data collection

This section includes two parts: data collection and processing. This section aims to filter and group the given data.

The temperature data inside the warehouse used here are formally authorised by the operator KLM Cargo, who has standard temperature measuring equipment at every storage location of the warehouse. Weather data outside the warehouse was obtained from the official website of the Royal Netherlands Meteorological Institute (KNMI), which provides the official Dutch weather service (KNMI,2019).

This initial data set covering from 1st June 2018 to 31st May 2019, contains hourly temperature data from 12 detectors inside the warehouse and hourly weather data outside the warehouse. This collected database contained 8,760 sets of data. For each of them, the following information is given: a dependent variable (temperature of each detector) and independent variables (corresponding outdoor temperature, wind speed, dew point value, sunshine duration, precipitation amount, precipitation duration, air pressure, cloud cover, humidity, fog or not, rainfall or not, snow or not, thunder or not, ice or not and so on.

Table 4. 1 Collected weather data.

Attribute of data (unit)	Contents
101.021.343 Stellinggebied VG3 locatie loods muurstelling 7 [1] (°C)	Inside temperature of detector, its number and location description.
TEMPERATURE (°C)	Outside temperature at 1.50 m at the time of observation
WIND_SPEED (m/s)	Hourly mean wind speed
DEW_POINT (°C)	Dew point temperature at 1.50 m at the time of observation
SUNSHINE_DURATION (hour)	Sunshine duration during the hourly division, calculated from global radiation (-1 for <0.05 hour)
RADIATION (J/cm2)	Global radiation during the hourly division
PRECIPITATION_DURATION (hour)	Precipitation duration during the hourly division
PRECIPITATION_AMOUNT (mm)	Hourly precipitation amount (-1 for <0.05 mm)
AIR_PRESSURE (hPa)	Air pressure reduced to mean sea level, at the time of observation
CLOUD_COVER	Cloud cover (in octants), at the time of observation (9=sky invisible)
HUMIDITY (%)	Relative atmospheric humidity at 1.50 m at the time of observation
FOG	0=no occurrence, 1=occurred during the preceding hour and/or at the time of observation
RAINFALL	0=no occurrence, 1=occurred during the preceding hour and/or at the time of observation
SNOW	0=no occurrence, 1=occurred during the preceding hour and/or at the time of observation

THUNDER	0=no occurrence, 1=occurred during the preceding hour and/or at the time of observation
ICE	0=no occurrence, 1=occurred during the preceding hour and/or at the time of observation

4.1.2 Data processing

The data provided is raw and it requires to be filtered. Following steps are taken to sort and filter the data as per the required use:

Step 1: Delete data sets without inside temperature.

Due to the installation time of some temperature detectors later than the start time of the database, or the detector failure caused by some unknown factors during the use, both of these results in the absence of indoor temperature data of some data sets. Therefore, the data sets without inside temperature are deleted.

Step 2: Delete temperature outliers.

According to the definition of temperature outliers mentioned in 2.2.1, the inside temperature exceeding the threshold that greater than or less than the daily average value five times the standard deviation will be considered outliers and removed from the data set.

Step 3: Classify the data according to whether the heating is turned on or not.

The KLM Cargo warehouse is equipped with heating facilities to reduce the quality loss of goods transported in winter when temperatures are too low. The operation of the heating system is a dynamic process: it is only turned on when the indoor temperature is below 16°C from September to April of the next year. When the heating system is turned on, the temperature in some parts of the warehouse may be quite stable and not significantly affected by the outdoor climate. Therefore, before building a temperature prediction model, it is necessary to divide the data into those with and without heating. The data filtering method for this step is described in Appendix B.

Step 4: Delete meteorological parameters that did not change over the data time range.

Since there was no snow or ice weather in the statistical time range, the corresponding values of these two parameters were all zero. This does not contribute to the creation of the model, so these two parameters (snow or not & ice or not) will be removed from the data set.

Step 5: Delete meteorological parameters that are not available in future weather forecasts.

For future model use, the input parameters of the indoor temperature prediction model (ie, outdoor meteorological data) need to be obtained through weather forecasts. However, the parameters in Table 4.1 are not fully available. By comparing the meteorological data of several commonly used weather forecast websites, we remove the meteorological parameters such as sunshine duration, radiation, precipitation duration and fog or not. To understand the impact of these deleted parameters on the fitting and accuracy of the model, we also established temperature prediction models including all meteorological parameters, and the evaluation results are shown in Appendix C.

After completing the above steps, each data set contains the following information: a dependent variable (temperature of each detector) and independent variables (corresponding outdoor temperature, wind speed, dew point value, precipitation amount, air pressure, cloud cover, humidity, rainfall or not, thunder or not).

4.2 Multiple linear regression model

R programming was used for constructing multiple regression models and all relevant statistical analyses. Stepwise regression method was accepted to ensure that all variables contained in the model were significant. Here we used the RMSE value the evaluate the model performance. For the RMSE value, we use 1.5 °C as a threshold to measure the accuracy of the model. The value of 1.5°C is derived from the temperature accuracy of commonly used temperature sensors (Instruments, 2013). For models with RMSE value higher than 1.5 °C, it is assumed that they have weak prediction effect. These models will be abandoned and the average hourly value will be calculated for prediction.

Non-heating data sets:

The RMSE values for all models are shown in Figure 4.1. For non-heating data sets, The RMSE value of most detector models is around 1 °C, among which the model of detector

2 and detector 22 is the best, with an error of 0.8 °C and 0.77°C respectively. The RMSE values of models for detector #1 and #5 are higher than 1.5 °C (1.57 °C and 1.62 °C). This shows that for these three locations, the multiple linear regression model cannot well describe the trend of indoor temperature change based on outside weather.

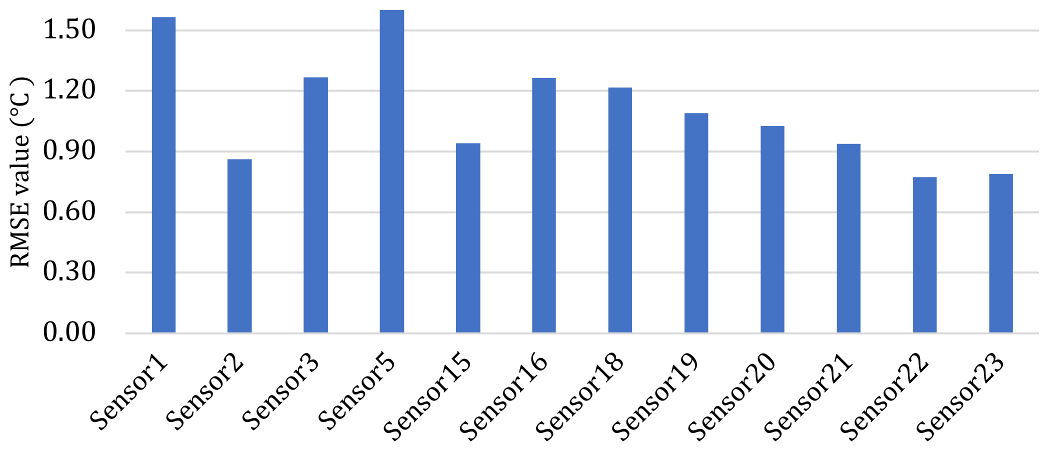


Figure 4. 1 The RMSE values for all sensors without heating

One of the reasons for the low fitting degree is that the correlation between indicators and results is not obvious. It is related to the limitations of available meteorological parameters of the model. When the results are compared with the all-parameter model shown in Appendix C, the RMSE values of all the models are lower than the latter. This is because in the fifth step of data processing, radiation value and precipitation duration, two important parameters that have a significant influence on indoor temperature, are deleted. Another reason is that the indoor temperatures are particularly stable at certain times of day and not easily affected by outdoor weather conditions. This can lead to greater errors in the overall model.

The parameters of each model are shown in Appendix E. By analysing the coefficients of each model, it is found that the coefficient of outdoor temperature and humidity is positive in all the models, which indicates that these two variables are positively correlated with indoor temperature, that is, higher outdoor temperature and humidity will lead to the increase of indoor temperature for all locations of the warehouse. On the contrary, independent variables such as cloud cover, dew point, wind speed and air pressure have negative effects on indoor temperature.

Heating data sets:

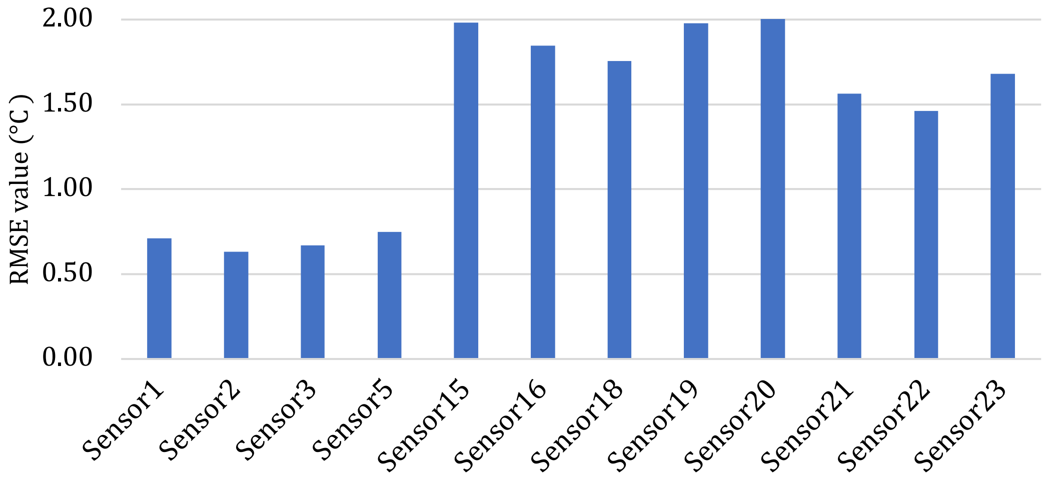


Figure 4. 2 The RMSE values for all sensors with heating

The RMSE values for all models with heating are shown in Figure 4.2. It can be found that for sub-data sets with heating system, the fitting degree of half the models is very low (RMSE values are higher than 1.5 °C). RMSE values of detector 15, 16, 18 and 19 are all close to 2°C. The model with the largest RMSE value is that of detector 20, about 2.075 °C. This is because in cold days, the heating system will turn on to stabilize the temperature in most locations of the warehouse, so that the influence of external weather on the indoor temperature is significantly reduced. Therefore, these multiple linear regression models with large RSME values will not be used to predict the indoor temperature, but the average hourly temperature of each detector will be used as the predicted value. The RMSE value for detectors #1, #2, #3, #5 are much smaller, all around 0.6 °C. It indicates that the fitting degree of these models is ideal.

To reduce the error of the model as much as possible, we will compare the predicted result with the real value to improve the established neural network model.

4.3 Neural network model

This part includes the establishment of the neural network model and the analysis of network model results.

4.3.1 Variable selection for input and output layer

Before modelling the neural network, it is necessary to remove meteorological factors that have no significant effect on indoor temperature. P-value test is an effective means to screen variables. The null hypothesis of p-value test is that the coefficient is zero, that is, the predictor has no effect on the response variable (Hansen, 1997). A low p value (< 0.05) means that the null hypothesis can be rejected, that is, the predictor is likely to be a meaningful addition to the model, and that its changes are related to changes in response variables.

Since the stepwise linear regression was used in Section 4.2, which can screen out outdoor meteorological parameters that have a significant influence on the output according to the p-value coefficient. Here, the variable selection results of multiple linear regression analysis will be used for the parameter selection of neural network analysis, so as to ensure that the selected variables have a significant impact on the results. We assume that when a certain weather index is selected by at least one multiple linear regression model, this factor is likely to have an impact on indoor temperature and is included as an independent variable in the neural network model. The reason the model does not include the dew point parameter is that it is collinear in humidity.

Table 4. 2 Selected factor based on multiple regression model (without heating).

Model coefficients	Model number												Selected
	1	2	3	5	15	16	18	19	20	21	22	23	
TEMPERATURE	√	√	√	√	√	√	√	√	√	√	√	√	YES
HUMIDITY	√	√	√	√	√	√		√	√	√	√	√	YES
CLOUD_COVER	√		√	√	√	√		√	√	√			YES
AIR_PRESSURE	√		√	√								√	YES
WIND_SPEED	√		√	√	√					√			YES
RAINFALL	√	√	√								√	√	YES
THUNDER	√		√	√									YES
DEW_POINT													NO
PRECIPITATION_AMOUNT						√							YES

Table 4. 3 Selected factor based on multiple regression model (with heating).

Model coefficients	Model number												Selected
	1	2	3	5	15	16	18	19	20	21	22	23	
TEMPERATURE	√	√	√	√	√		√	√	√	√	√	√	YES
HUMIDITY		√	√	√		√				√	√	√	YES
CLOUD_COVER	√	√	√	√			√			√	√	√	YES
AIR_PRESSURE	√	√		√	√	√	√	√	√			√	YES
WIND_SPEED	√	√	√	√		√	√		√		√	√	YES
RAINFALL	√				√	√	√		√	√	√	√	YES
THUNDER				√			√						YES
DEW_POINT													NO
PRECIPITATION_AMOUNT			√				√	√					YES

As can be seen from the above two tables, most regression models include temperature, humidity and cloud cover. It shows that these factors have universal influence on indoor temperature. In addition, during heating, we found that the model contained more parameters. In particular, the air pressure, the wind speed, these two factors are basically in the model of all the detectors.

4.3.2 Construction of the neural network

With R programming software, the feedforward neural network is built. Based on Table 4.2 and Table 4.3, the training data set of the neural network model is composed of 9 variables for models without/with the heating system (8 inputs and 1 output). The warehouse internal temperature is the target variable. 80% of the training data set is used for learning, and 20% of the training data set is used for model validation. The code for building the neural network is shown in Appendix D.

The boxplot of the 10-fold cross-validation results shows the error range of the model. The main standard errors of all neural network models are within the error range shown in the boxplots, that is, the models are not over-trained or under-trained.

Based on the research from Kim, the number of neurons in the hidden layer is 2n or 2n+1 (n is the number of independent variables) (S. Y. Kim et al., 2018). For both data sets, the hidden layer number of the ANN was set to 4, 8, 16 and 17. The relevant parameters of the model results are shown in the tables below. Models with the smallest RMSE are selected as final neural networks models for detectors, which are shown below.

Table 4. 4 Selected ANN models.

Detector number	Without heating		With heating	
	RMSE (°C)	Model structure	RMSE (°C)	Model structure
1	1.41	8-4-1	0.54	8-8-1
2	0.82		0.56	
3	1.19		0.56	
5	1.47		0.56	
15	0.87		1.38	
16	1.13		1.37	
18	1.14		1.28	
19	0.96		1.36	
20	0.88		1.17	
21	0.83		1.08	
22	0.71		1.19	
23	0.75		1.32	

Models without the heating system fit better overall than models with the heating system. For all models, all of them have RMSE values less than 1.5. For models without the heating system, RMSE values of detectors #1, #2 are close to 1.5, which indicates that the model cannot well describe the relationship between independent variables and dependent variables.

4.4 Comparison of results for both models

In this part, we compared the RMSE values of two temperature prediction models, which reflected the fitting degree of models.

4.4.1 Without heating system

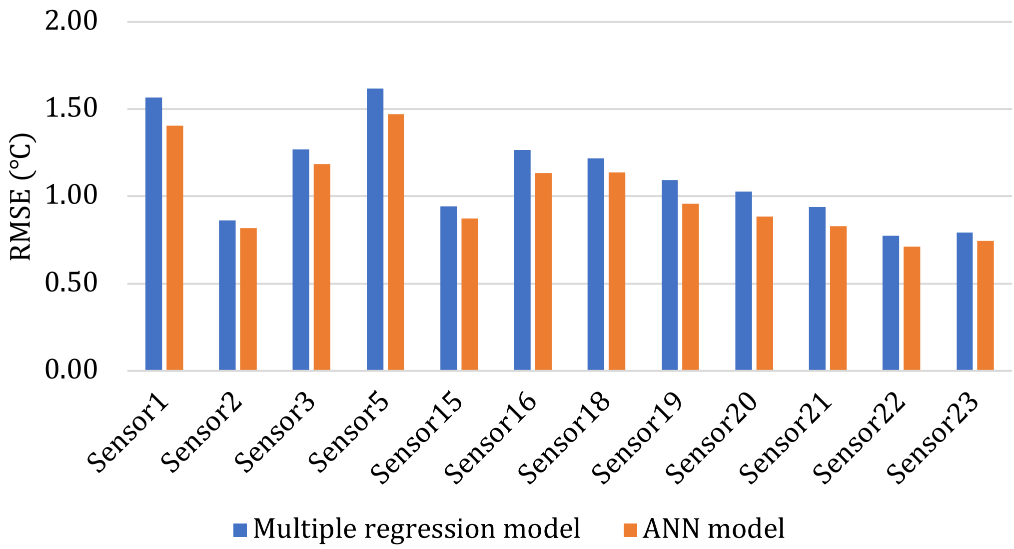


Figure 4. 3 Comparison of RMSE values of ANN model and Multiple regression model.

In general, both models have good predictive power when heating is not turned on. Most models have an error of about 1 °C. The RMSE values of the detector 1 and detector 5 models were slightly larger than those of the others, both close to 1.5 °C. The models of detector 2, detector 22, and detector 23 gave better results, with errors of nearly 0.7 °C. For all sensors, the fitting degree of the neural network model is better than that of the multiple regression model. But RMSE values of the two models are not much different, basically around 0.1 °C.

4.4.2 With heating system

Figure 4.3 shows the comparison of the results of the two models with the heating on. The fitting degree of the neural network model is better than that of the traditional model. For the two models, the results of different detectors are obviously different. RMSE values for both models of detector #1, #2, #3 and #5 are very small (close to 0.5 °C), indicating that both models can play a good role in predicting temperature. For traditional multiple linear regression models, except for sensors #1, #2, #3, #5, and #22 can be accepted, and other models are poorly fitted. The main reason why the model is difficult to predict and the accuracy is too low is that under the condition of the heating, the indoor temperature of some detector areas is relatively stable and will not be affected by the outside weather as obviously as when there is no heating.

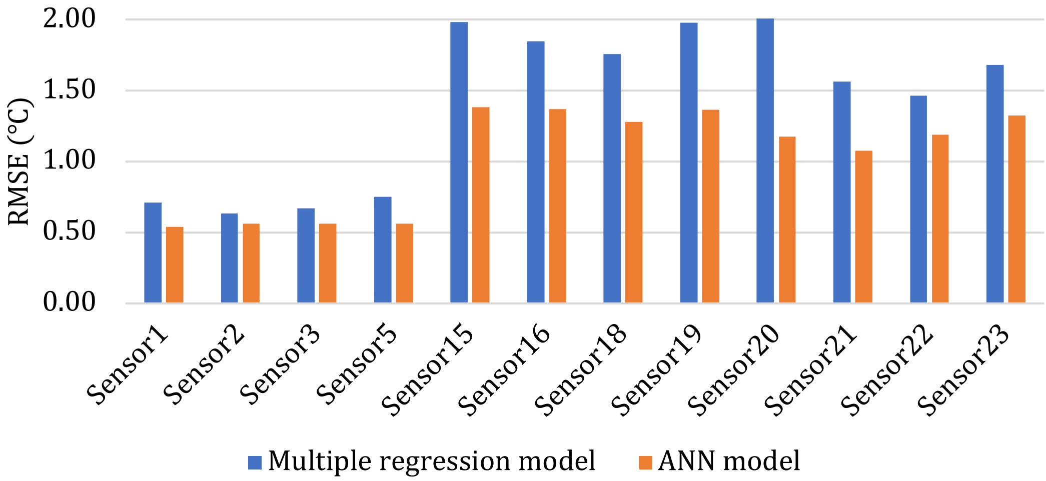


Figure 4. 4 Comparison of RMSE values of ANN model and Multiple regression model.

By summarizing, we listed the advantages and disadvantages of the two models for reference when selecting the final model.

Table 4. 5 Pros and cons of two prediction models.

	Pros	Cons
Multiple regression model	1: The relative effects of one or more predictive variables on the criterion values can be determined. The coefficients of the independent variables of the model can intuitively reflect their influence on the dependent variables. 2: Easy to implement and interpret. As a well-known and widely accepted prediction model, it has been implemented in many previous studies. Based on Table 5 of Appendix E, we can understand which outdoor weather factors affect indoor temperatures in different locations. At the same time, this feature can also help us to understand how to adjust the model when we want to further reduce the	1: This model can only represent linear relationships. Some independent variables may have a significant impact on the model based on the p-value test, but they are excluded due to their non-linear relationship with the independent variables. 2: Relationships that can be learned are restricted and often oversimplify complex realities. Because the use of multiple linear regression models needs to meet multiple assumptions, including parameter independence test, collinearity diagnosis, significance test, linearity test, etc. After meeting all the assumptions, the actual model will contain fewer

	prediction error. 3: Inferential statistics (p-values and RMSE value) are available. P-values helps to figure out the independent variables that actually affect the dependent variable. The RMSE value allows us to evaluate the predictive accuracy of the model. 4: Mathematically, it is straightforward to estimate weights and you have a guarantee to find the optimal weight. After combining with k-fold cross validation approach, we can select the best model parameter and get its corresponding coefficient value. 5: How predictions are produced is transparent. Therefore, when we need to change the selected independent variables of the model in the future, we can directly understand the impact of these changes on the prediction model.	parameters than the nine independent variables obtained from the data processing, which will affect the prediction effect of the model.
Neural network model	1: Very flexible because the basic structure (the number of hidden layers and nodes) is selected by the user. Neural network model does not have many harsh conditions as traditional models. In addition, it can capture complex relationships between inputs and outputs. 2: Ability to work with incomplete knowledge. At present, it is difficult	1: Black box model: difficult to explain, thus hard to justify. When we want to reduce the model error as much as possible, it is difficult to find a clear solution as for the linear regression model. Therefore, we need to make continuous attempts to reduce the prediction errors of neural network models with the help of the experience of traditional

	to describe the relationship between outdoor weather factors and indoor temperature with a simple formula. Neural network analysis can find the relationship between independent variables and dependent variables in the limited knowledge background. 3: Rapid recognition speed. Computers can build complex neural networks in a short time.	models. 2: Space and time complexity: lengthy training times. The building time of neural network will increase with the complexity of the model. Complex model structure does not mean good prediction effect. At the same time, we need to prevent overtraining and undertraining. Thus, it is difficult to find the best neural network structure. 3: Requires user to carefully pre-process predictor variables, experiment with different sets of predictors. Here we introduced k-fold cross approach in model validation.

4.5 Model refinement

In Figure 4.4, if 1.5 of the RMSE value is taken as the threshold, the multiple linear regression model of half the detectors cannot be accepted. At the same time, we hope to reduce model errors by optimizing the model as much as possible. So, we want to optimize the prediction model as much as possible to reduce the RMSE value so that more models can be accepted.

4.5.1 Model validation for multiple linear regression model

For the multiple linear regression model in 4.2, the first obvious disadvantage is that there is no procedure for model validation. Therefore, we cannot determine whether the parameter selection of the regression model is optimal. For multiple linear regression models, model validation can help us select a small number of appropriate independent variables, which can not only avoid over-fitting but also increase the degree of interpretation of the model.

Here, the 10-fold cross-validation approach has been used. The data is divided into 10 folds. Then take the sample of the i fold as the verification set (i =1 to 10), and the others as the training set to train the model. The mean of validation errors of 10 models is the overall validation error of the model.

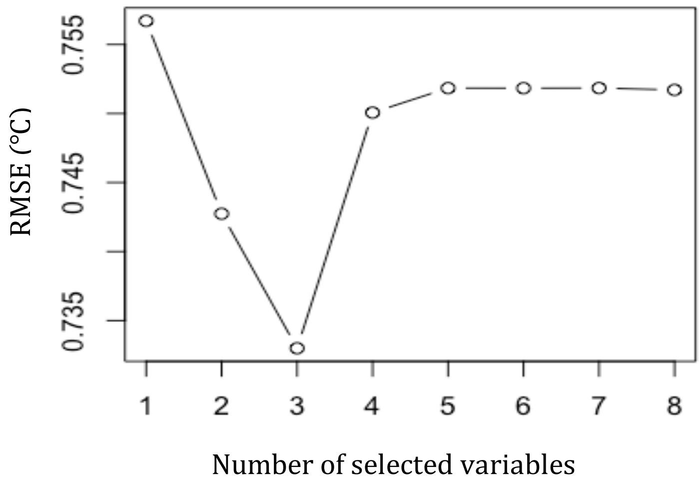


Figure 4. 5 An example of model validation for multiple linear regression model.

For instance, the figure above shows us an example of model validation in this project. According to Section 4.1, there are eight optional independent variables, namely corresponding outdoor temperature, wind speed, dew point value, precipitation amount, air pressure, cloud cover, humidity, rainfall or not, thunder or not. It can be seen that when the model has 3 parameters, the corresponding average verification error is the smallest, which is about 0.731 °C. Based on the code of multiple linear regression model, we can conclude that the best three independent variables are outdoor temperature, humidity and rainfall or not. Then, by performing a full subset regression of the entire dataset, we will build the regression model based on these best 3 variables.

4.5.2 Model refinement

By comparing the predicted results of each sensor's model with the actual results, we found that the data with huge errors were always concentrated at a certain time of day.

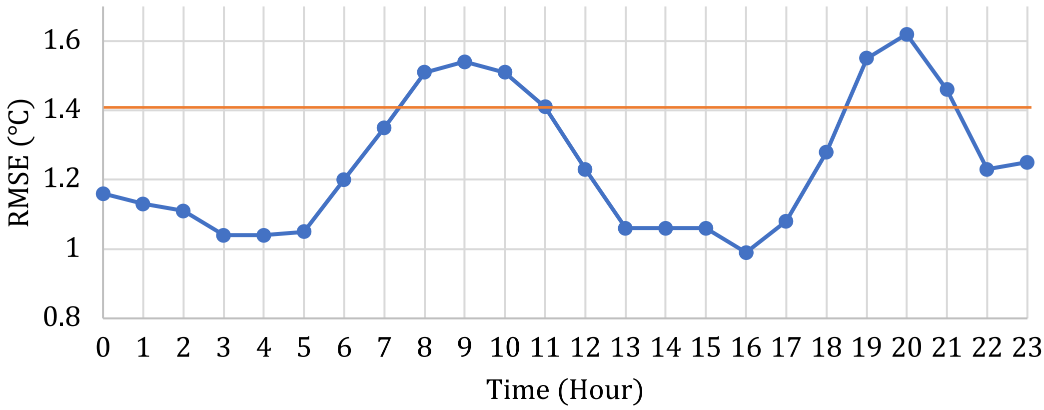


Figure 4. 6 RMSE value per hour of sensor 1(without heating).

The figure above shows the error variation of the regression model of sensor 1 (without heating) at different times. The orange line is the RMSE value of the model (1.41 ° C). It can be found that the errors corresponding to different time points vary greatly. From 1 a.m. to 6 a.m. and from 1 p.m. to 5 p.m. in the afternoon, the RMSE value was about one degree Celsius, far less than the average RMSE value of the model. From 7 a.m. to 11 a.m., and from 6 p.m. to 9 p.m., the error of the model is very large (around 1.5° C). This means that the influence of weather indicators on indoor temperatures is not constant over time. In addition, a previous researcher conducted a series of experiments to understand if time-division affected the accuracy of temperature prediction models(Smith, McClendon, & Hoogenboom, 2006). He compared the prediction errors of the one-hour model, the four-hour model, the eight-hour model, and the 12-hour model. The results show that the one-hour model has the smallest error, while the eight-hour model has the largest error.

Therefore, in order to further reduce the error, we refine the model. According to the time, we will build four models for each sensor: the model from 1 a.m. to 6 a.m., the model from 7 a.m. to 12 a.m., the model from 1 p.m. to 6 p.m., and the model from 7 p.m. to 12 p.m.

The parameters of each model are shown in Appendix E and RMSE values for models are shown in Appendix F. The results show that after refining the model, the RMSE values for most models are smaller than previous models.

In order to see the improvement of the refinement of the model more intuitively, we took the detector 1 (without heating) as an example, randomly selected data for a period of time, and compared the results before and after refinement

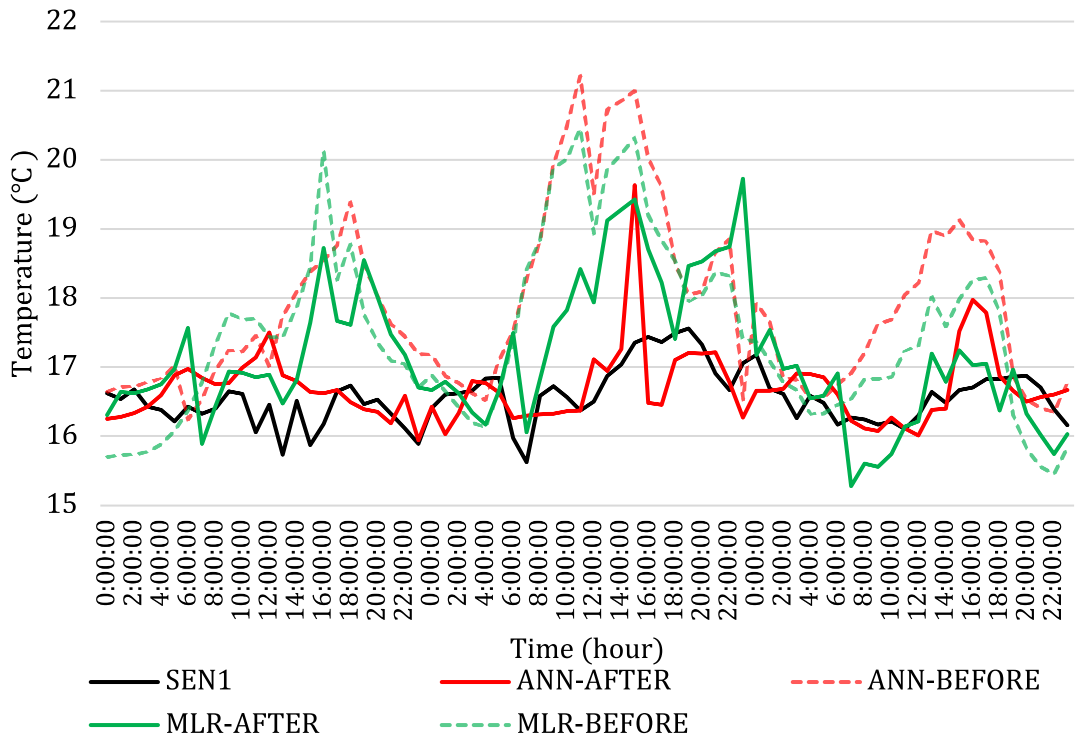


Figure 4. 7 Predicted temperature of two models before and after refinement.

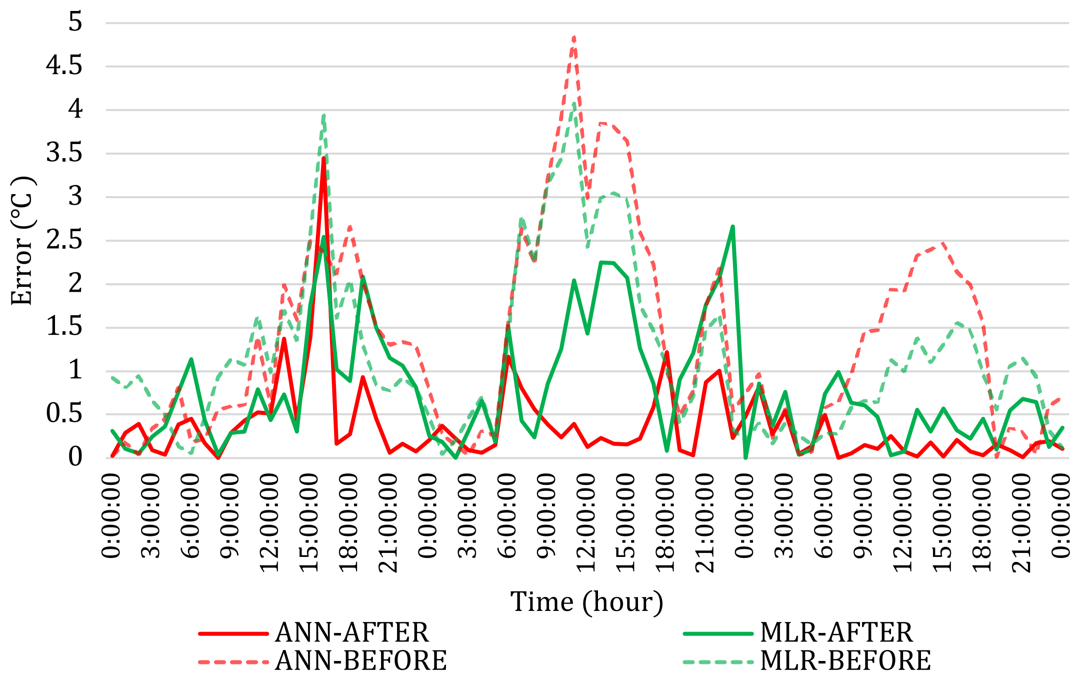


Figure 4. 8 Predicted error of two models before and after refinement.

It can be seen from the figure above that the two models are better close to the real results after refinement, and the neural network analysis model is better than the results of the multiple linear regression model.

4.6 Discussion

In this chapter, the multiple linear regression model and neural network analysis model based on backpropagation algorithm are used to predict indoor temperature. During the data collection phase, the outdoor data came from KNMI and all the indoor temperature data came from KLM Cargo. After data collection, we processed the data.

The k-fold cross-validation approach is used for both models. It is used in the neural network model to prevent overtraining and in the linear regression model to find the optimal parameters.

The prediction performance of the two models was analysed and compared by using the RMSE values. In the case that the heating is not turned on, the fitting degree of the two models is basically within the reasonable range, except the RMSE values of multiple regression model for detector 1 and 5 are greater than 1.5 °C. When the heater is turned on, the multiple linear regression models of half of the detectors have a large RMSE value. This is because when the heating is on, the indoor temperature tends to be stable and less affected by the outside weather.

In the process of model optimization, we found that the error of the model was very large at some time of the day. To reduce the model error, we decided to refine the model. The prediction results showed that the refined model was closer to the real value than the previous one.

In general, the result of the neural network model has smaller error and higher fitting degree. Therefore, it will be applied to the simulation in Chapter 5. But we cannot ignore the advantages of the multiple linear regression model. For instances, in the future when the model is updated, the regression model is more conducive to us to understand and apply.

Chapter 5 Simulation

In this chapter different storage location choice scenarios are tested to discover the effect of external weather on location choice results. The flowchart method in Section 3.3 is the decision support for the simulation experiment. Since the sort criteria in the flowchart have a great influence on the result, we will try different sort criteria in the simulation experiment to find the best one in the control strategy. Besides, we divided the simulation experiment at two levels to explore how to optimize the layout and selection of temporary storage locations from strategic and tactical perspectives. Based on the actual operation of KLM cargo, we simulate the quality loss based on the real situation and compare it with the proposed control strategy. Therefore, the following research question is established for this chapter:

Given the information of perishable goods and weather forecasts, how much improvement can the control strategy provide according to the simulation results?

- What assumptions are made in the simulation?
- What is the best sort criteria for decision support?
- What is the loss of quality based on the actual case and based on the proposed control strategy?

After making some basic assumptions about this experiment in the first section, we proposed three sort criteria in the flowchart method according to the operation of KLM Cargo to found the best one through simulation results. By measuring the quality loss and the proportion of goods in different locations, we can learn whether the current temporary storage locations are optimal and which warehouse locations can be used for temporary storage in the future.

5.1 Assumptions

1. The location of the warehouse is divided according to the temperature detector, and it is assumed that the temperature measured by each detector can represent the average temperature of a certain area. The location of the temperature sensor is shown below.

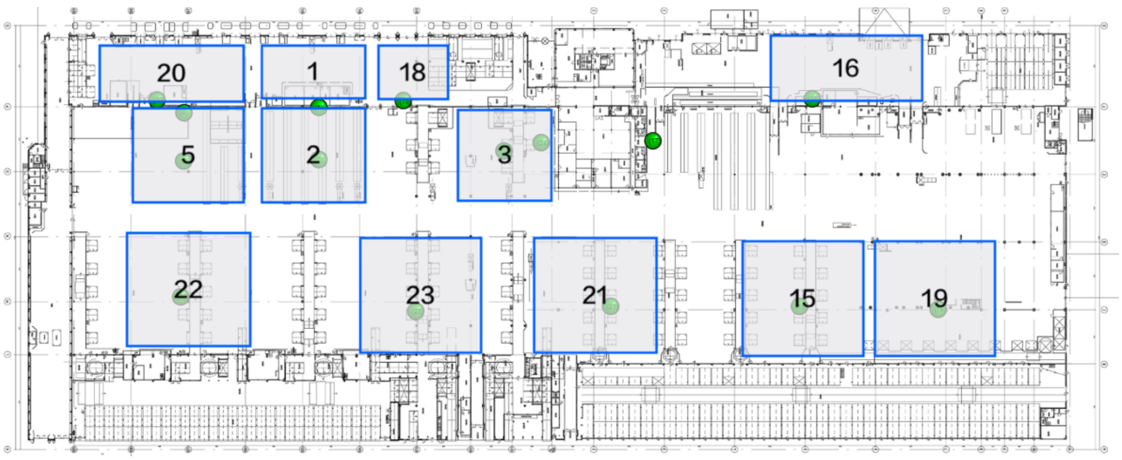


Figure 5. 1 The location of the temperature sensors.

2. In actual shipment, there may be other kinds of goods besides medical products and flowers. Here we assume that only drugs and flowers arrive at the warehouse during the simulation.

3. As the goal of this chapter is to create scenarios and run simulations, it is assumed that the best shelf life of all pharmaceutical products is half a year, and the optimal vase life of all flowers is 304 hours.

4. Since the breaking down and building up time in the warehouse is difficult to estimate, it is assumed that the time spent in temporary storage locations is the period when the goods arrive and leave the warehouse.

5. [redacted]

6. [redacted]

7. For the simulation at strategic level, it is assumed that all locations of the warehouse may be used as temporary storage sites in the future.

8. At present, KLM cargo takes time into account when selecting the location of its shipments. The selection of temporary storage locations is mainly based on the storage time in the warehouse. To make the simulation more realistic, it is assumed that KLM cargo currently uses a sort criterion based on the length of time the shipment is stored in the warehouse. Products that stay in the warehouse for a long time can have the right to choose the storage location first.

9. For the quality loss model, we assume that the goods will not suffer any quality loss within the recommended storage temperature range.

5.2 Sort criteria

The simulation in this chapter is based on the flowchart method described in Section 3.4 for decision making. For this model, we need to decide in advance the sequence in which the goods are allocated. This order is related to the information that KLM Cargo has obtained about the incoming shipments. To test the method, we assume three possible cases.

- Case 1: The information about the shipments arriving in the next 72 hours is constantly updated. KLM Cargo is likely to keep getting orders for the next three days. The information of the incoming goods is confirmed and will not be changed.

- Case 2: Part of the information about shipments arriving in the next 72 hours can be obtained, and part is uncertain. For example, some goods only provide the approximate time of departure from the last airport, which needs to be further updated and confirmed.

- Case 3: All information about the shipments arriving in the next 72 hours is available. All cargo information has been confirmed.

For the first case, we will take the estimated arrival time of shipments at Schiphol as the basis for determining the sequence. Goods arriving early have the right to choose storage locations first. For the second case, we can use machine learning methods to predict the approximate range of goods uncertainty information. However, it is not the focus of this topic, so it will be further studied in the future. For the third case, we will determine the sequence based on the mean and standard deviation of the quality loss of each good at all storage locations.

To select the best sequencing criteria, the Monte Carlo simulation described in Section 3.3.3 will be used to evaluate the final results of the three sequencing methods.

5.3 Scenarios Building

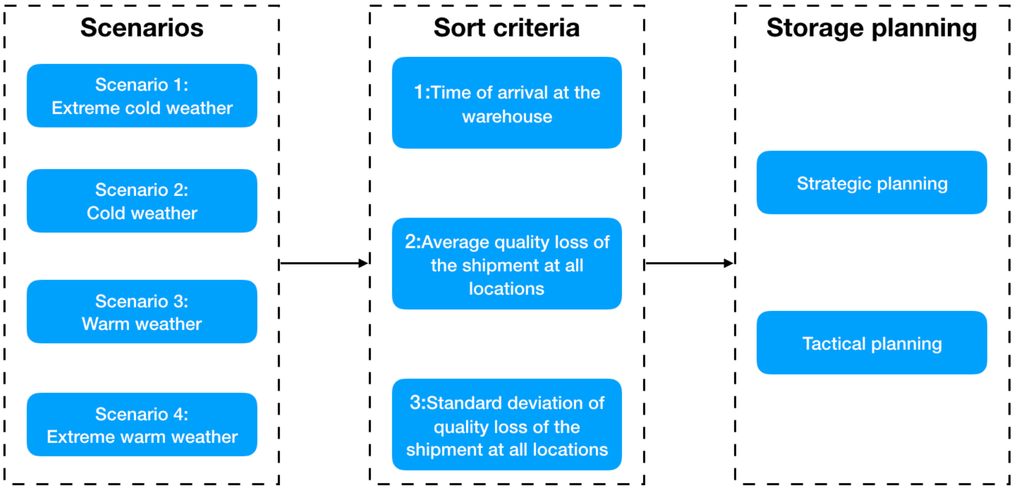


Figure 5. 2 Simulation plan

As shown in Figure 5.2, in order to understand the impact of different weather conditions on the results of storage location selection, we will simulate four different scenarios using three sort criteria. By analysing meteorological data from June 1, 2018 to May 31, 2019, we selected four representative periods. The temperature information corresponding to these four periods is shown in the following table.

Table 5. 1 Weather data for four scenarios.

Scenarios	Time	Temperature range	Average temperature
Cold	1/11/2018—3/11/2018	1.7°C — 12.3°C	8.1 °C
Warm	29/5/2019—31/5/2019	5.2 °C — 20.1 °C	15.6 °C
Extreme cold	26/2/2019—28/2/2019	-8.6 °C — -0.2 °C	-4.0 °C
Extreme warm	25/7/2018—27 /7/2018	18.9 °C — 35.2 °C	27.2 °C

The layout planning of temporary storage locations strategic planning and tactical planning. For KLM cargo, the layout of the warehouse is not likely to change much anytime soon. Therefore, we need to understand the utilization of existing storage locations in different weather conditions. This helps KLM cargo make some minor adjustments. For example, according to the simulation results, they can increase the number of storage racks in high-utilization locations to improve their capacity. Therefore, for tactical planning, we will only simulate the location selection of goods in the existing temporary storage locations. The temporary storage locations currently in use are located near sensor 2, 3 and 5.

For strategic planning, we need to understand where the goods are stored with the least quality loss when all the locations in the warehouse are available. Based on the simulation results, temporary storage locations can be adjusted greatly in the future. Here, all locations can be selected according to hypothesis 7.

5.4 Simulation results

The simulation relies on the supercomputer at the Leibniz Supercomputer Centre. Simulations will be run 10,000 times. The computation time for each model is around three hours and the total computation time is around 24 hours. During the simulation, goods information will be generated randomly. The Settings of simulation parameters are described in Appendix G.

Simulation results are as follows:

5.4.1 Strategic planning:

5.4.1.1 Scenario 1: Extreme cold weather

Because the results based on sort criteria 2 and sort criteria 3 are very close and they are quite different from the results based on sort criteria 1. If the distribution of the three is put together, it is difficult to distinguish the results generated by sort criteria 2 and sort criteria 3. Therefore, here we compare the results of sort criteria 1 with the results of sort criteria 2, and then compare the results of sort criteria 2 and of sort criteria 3.

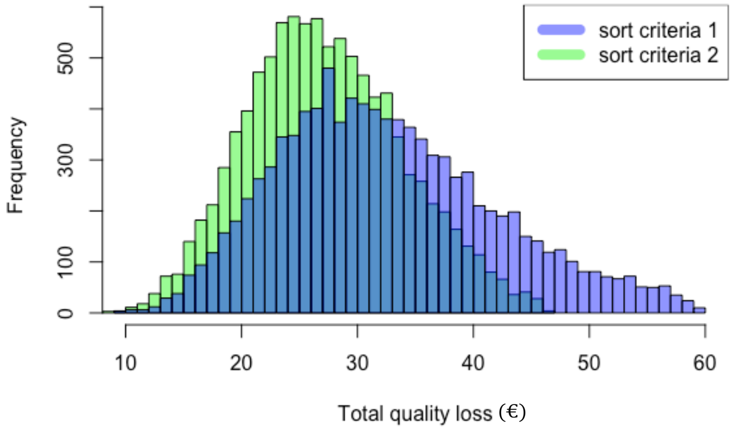


Figure 5. 3 Quality loss distribution based on sort criteria 1 and 2.

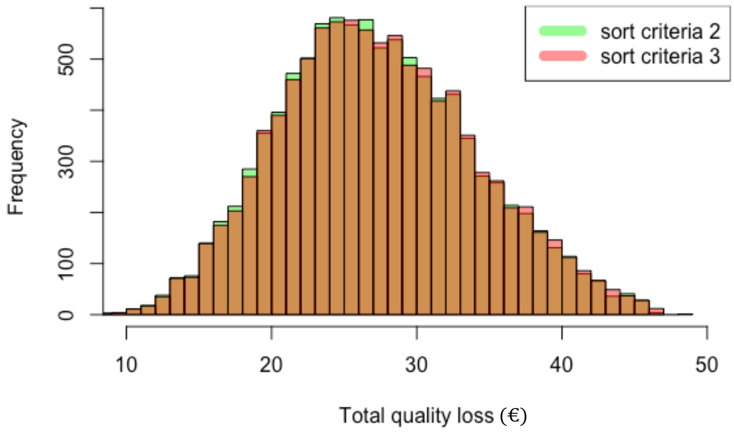


Figure 5. 4 Quality loss distribution based on sort criteria 2 and 3.

In general, when the weather is extreme cold, the total quality loss according to sort criteria 2 and 3 is basically normal distributed. According to Figure 5.3, we can see that the results generated by sort criteria 1 have a wider distribution range and are more skewed to the right, indicating that compared with sort criteria 1, the corresponding total quality loss is generally larger. However, the distribution of simulation results by sort criteria 2 is similar to that by sort criteria 3.

Once we have the final distributions, we first need to assess whether the distribution is normal. Based on the goodness of fit test, it was found that these three distributions are not normal. Thus, we applied the bootstrapping method to calculate confidence intervals of the median of the quality loss value, which should be the median of the distribution. The results are shown in the following table.

Table 5. 2 95% Confidence interval of results based on the Bootstrap approach.

	Sort criteria 1	Sort criteria 2	Sort criteria 3
Lower bound	32.38	27.13	27.23
Upper bound	32.71	27.37	27.51

From the analysis of the above results, it can be concluded that under extremely cold conditions, the quality loss generated by the three sorting rules is not much different. The biggest difference between the three sort criteria is around €5. But the control strategy based on sort criteria is able to reduce more quality loss (around 27.13 to 27.37 €).

5.4.1.2 Scenario 2: Cold weather

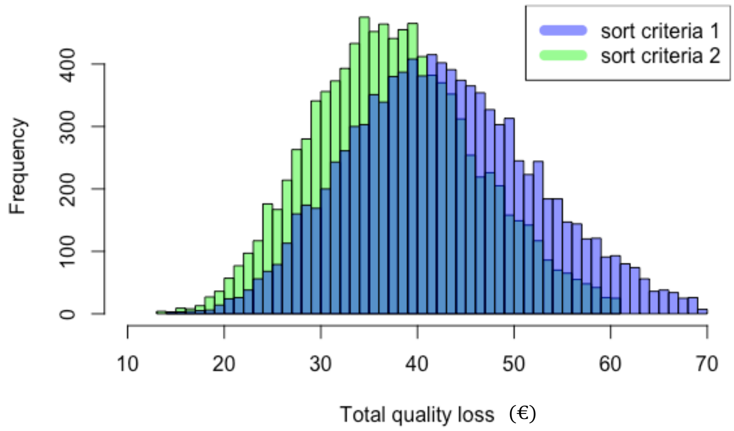


Figure 5. 5 Quality loss distribution based on sort criteria 1 and 2.

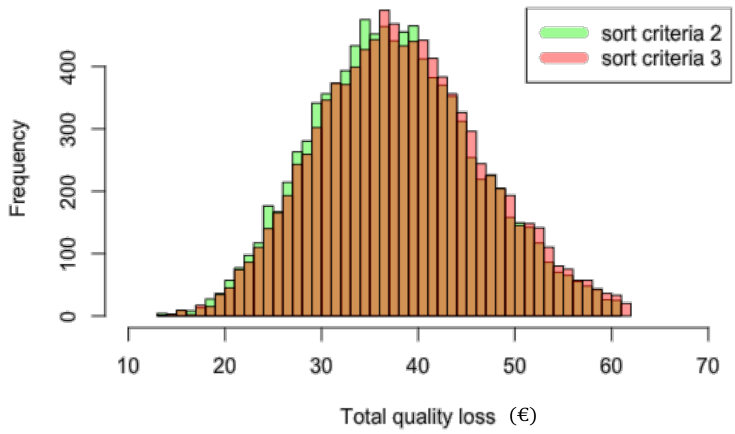


Figure 5. 6 Quality loss distribution based on sort criteria 2 and 3.

The distribution characteristics of simulation results generated by the three sort criteria are basically consistent with scenario 1. As can be seen from Figure 5.6, the simulation results according to the sort criteria 2 are slightly skewed to the left on the whole, so the overall result will be slightly less than that according to the sort criteria 3. The confidence intervals of the three simulation results are as follows.

Table 5. 3 95% Confidence interval of results based on the Bootstrap approach.

	Sort criteria 1	Sort criteria 2	Sort criteria 3
Lower bound	42.37	37.44	38.09
Upper bound	42.76	37.78	38.44

From the analysis of the above results, we can conclude that there is not much difference between the results under cold conditions and those under extremely cold conditions. The biggest difference between sort criteria 1 and sort criteria 2 is still around €5. The simulation results generated by sort criteria 2 are still optimal (37.44 to 37.78 €).

5.4.1.3 Scenario 3: Warm weather

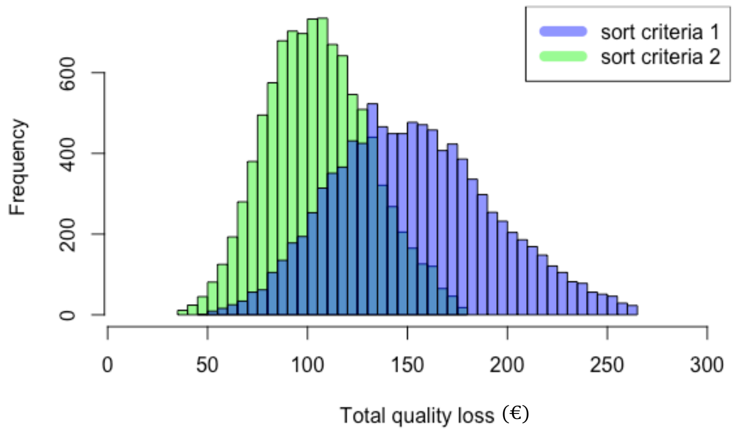


Figure 5. 7 Quality loss distribution based on sort criteria 1 and 2.

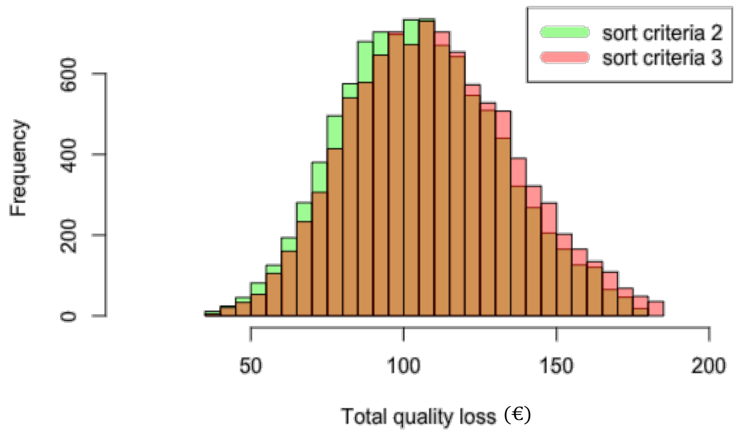


Figure 5. 8 Quality loss distribution based on sort criteria 2 and 3.

As the weather warms, all the simulation results become more widely distributed. Based on Figure 5.7, we can find the distribution of results produced by the two different sort criteria is more different than at low temperatures. As for Figure 5.8, it is more obvious that the simulation results obtained by sort criteria 2 are relatively small. The confidence intervals of the three simulation results are as follows.

Table 5. 4 95% Confidence interval of results based on the Bootstrap approach.

	Sort criteria 1	Sort criteria 2	Sort criteria 3
Lower bound	152.28	105.15	109.12
Upper bound	154.08	106.27	110.33

It can be found that when the temperature becomes warmer, the mass loss of the item increases significantly. The results from different collation rules are also much different. As with the previous two scenarios, sort criteria 2 is still the best.

5.4.1.4 Scenario 4: Extreme warm weather

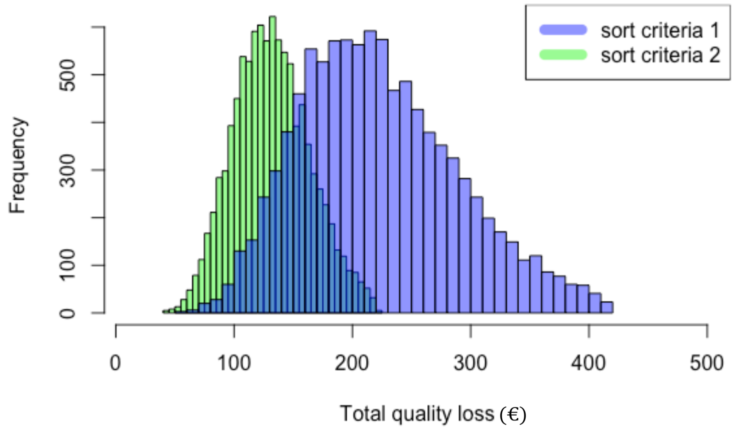


Figure 5. 9 Quality loss distribution based on sort criteria 1 and 2.

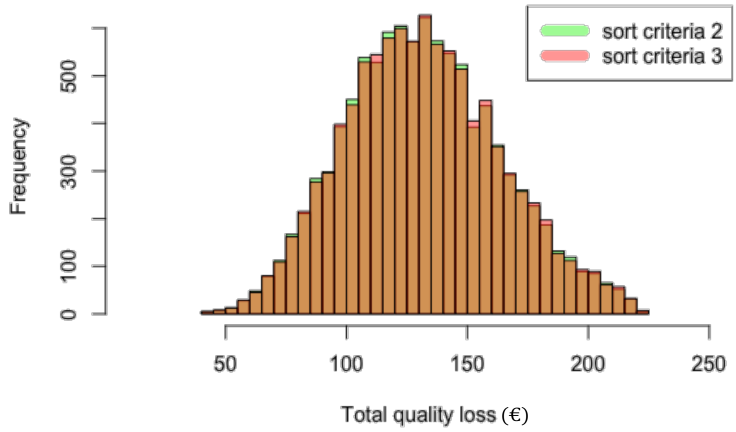


Figure 5. 10 Quality loss distribution based on sort criteria 2 and 3.

The distributions of simulation results in extremely warm weather are similar to those in warm weather. The simulation results produced by sort criteria 1 are the widest of the three, ranging from 50 € to 260 €. The range of simulation results produced by sort criteria 2 and 3 is consistent, from 40 € to 220 €. The confidence intervals of the three simulation results are as follows.

Table 5. 5 95% Confidence interval of results based on the Bootstrap approach.

	Sort criteria 1	Sort criteria 2	Sort criteria 3
Lower bound	221.07	130.91	131.16
Upper bound	223.67	132.32	132.55

The characteristics of the confidence interval are consistent with the distribution shown in above figures. Under extremely hot weather, the quality loss caused by sort criteria 1 is significantly greater than that of the others.

5.4.1.5 Overall performance for strategic planning

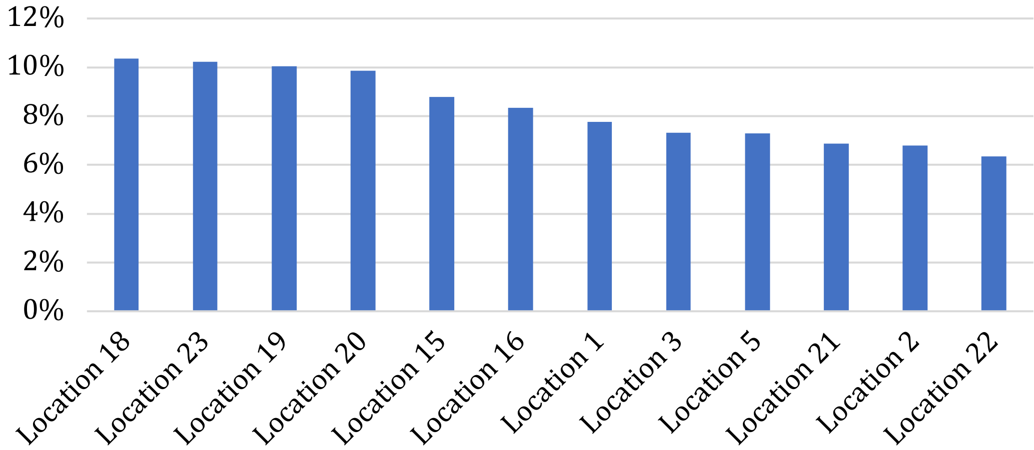


Figure 5. 11 The average percentage of each location selected in the simulation.

It was found that a relatively large percentage of shipments were stored near detector 23, 18, 20 and 19 (around 10%). However, a relatively low proportion of goods are stored in the temporary storage locations currently in use (located near detector 2, 3 and 5). The proportion of detector 2, 3 and 5 is 6%, 7% and 7% respectively. It is indicated that the temporary storage location currently in use is not an optimal storage location in the long run.

5.4.2 Tactical planning:

The tactical planning portion contains only the temporary storage locations currently in use, located near the detector 2, 3 and 5. This part of the simulation experiment is to understand the use of three storage locations in different weather conditions.

5.4.2.1 Scenario 1: Extreme cold weather

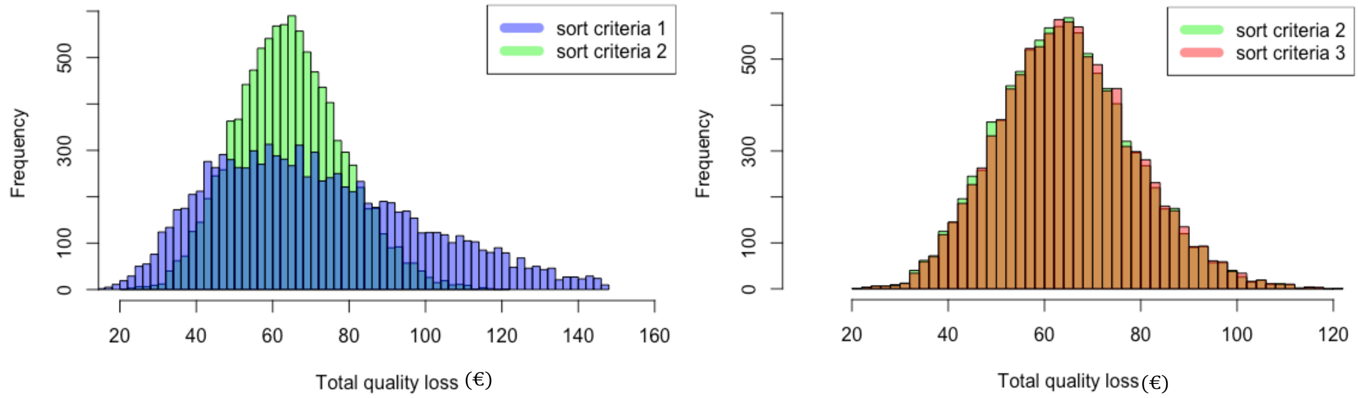


Figure 5. 12 Quality loss distribution of results based on sort criteria 1,2 and 3.

Table 5. 6 95% Confidence interval of results based on the Bootstrap approach.

	Sort criteria 1	Sort criteria 2	Sort criteria 3
Lower bound	67.69	64.22	64.53
Upper bound	68.57	64.75	65.18

Compared with the results of strategic planning (around 30 €), tactical planning produces more quality loss in the same scenario (around 65 €). Although the confidence intervals based on three sort criteria are similar, it can be seen that the value range of the quality loss caused by sort criteria 1 is more widely distributed (from 20 € to 150 €). The range of results produced by sort criteria 2 and sort criteria 3 is basically the same (between 20 € and 120 €), and the difference in confidence points is within 0.5 €.

5.4.2.2 Scenario 2: Cold weather

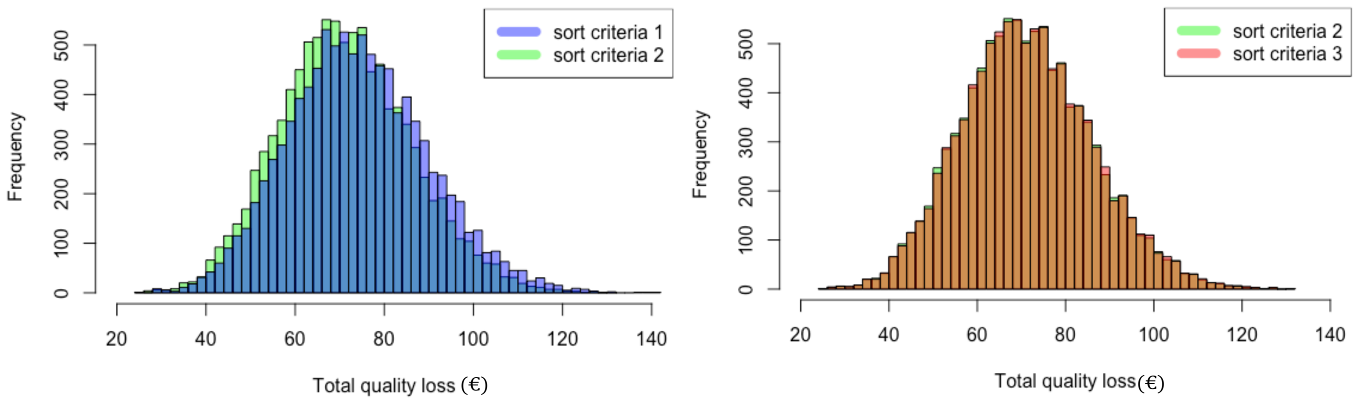


Figure 5. 13 Quality loss distribution based on sort criteria 1, 2 and 3.

Table 5. 7 95% Confidence interval of results based on the Bootstrap approach.

	Sort criteria 1	Sort criteria 2	Sort criteria 3
Lower bound	73.80	70.96	71.13
Upper bound	74.45	71.51	71.58

Compared to the extreme cold scenario, the quality loss of all the models increased by about 5 euros. In cold weather, although the three sort criteria produce similar results, sort criteria 2 is still better than others with less quality loss. Sort criteria 1 still produced the largest result, about 3 € more than the confidence interval of the other two results.

5.4.2.3 Scenario 3: Warm weather

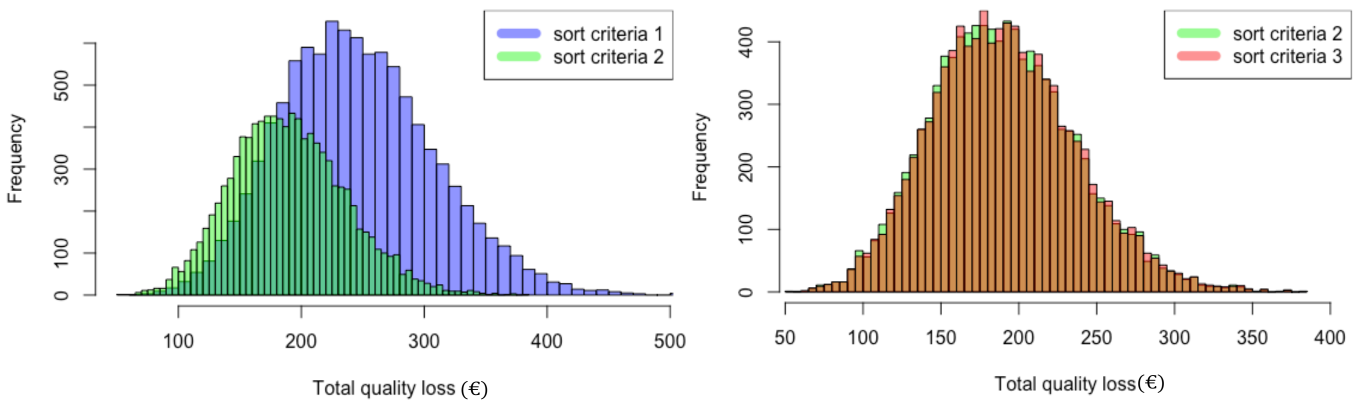


Figure 5. 14 Quality loss distribution based on sort criteria 1, 2 and 3.

Table 5. 8 95% Confidence interval of results based on the Bootstrap approach.

	Sort criteria 1	Sort criteria 2	Sort criteria 3
Lower bound	243.67	188.91	189.38
Upper bound	246.42	190.59	191.30

It can be seen that when the weather gets warmer, the quality loss of goods increases significantly, especially the result simulated based on the sort criteria 1. The maximum value of simulation results based on sort criteria 2 and 3 is about 400 euros, while that value for sort criteria 1 is 500 euros. According to Table 5.9, it is found that sort criteria 2 is slightly better than sort criteria 3 and its quality loss is obviously less than that of criteria 1.

5.4.2.4 Scenario 4: Extreme warm weather

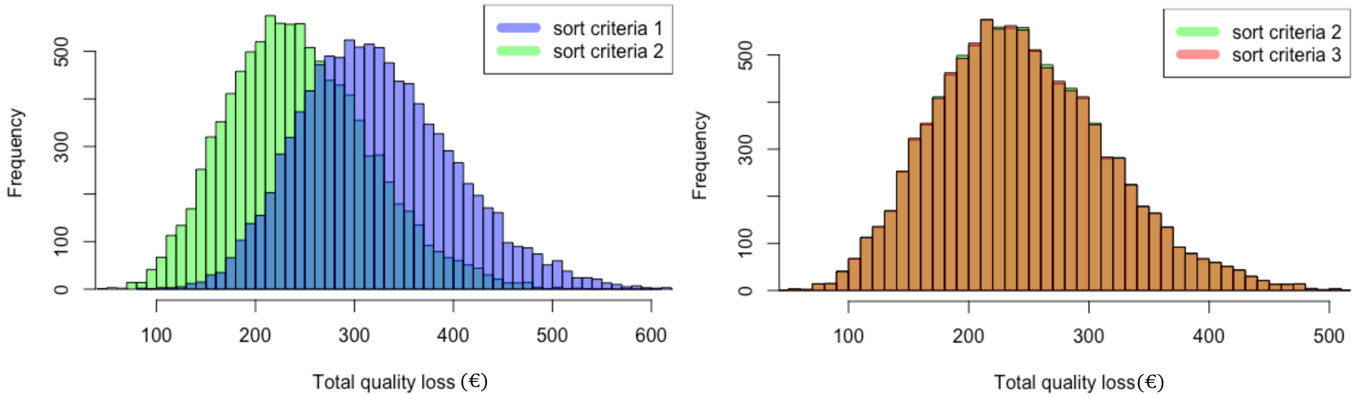


Figure 5. 15 Quality loss distribution based on sort criteria 1 and 2.

Table 5. 9 95% Confidence interval of results based on the Bootstrap approach.

	Sort criteria 1	Sort criteria 2	Sort criteria 3
Lower bound	321.42	242.93	242.85
Upper bound	324.65	245.57	245.78

When the weather is extremely hot, all the simulation results reach a larger range. From the above results, it can be concluded that the quality loss according to the sort criteria 1 is significantly larger than that of the others. Its confidence interval turned out to be about 320 euros, nearly 80 euros larger than the others. The results from the other two sort criteria are very close. By analysing the confidence interval, we can conclude that the quality loss obtained by sort criteria 2 is the minimum of the three (around 240 euros).

In addition, in all four scenarios, the confidence interval obtained by the tactical planning part is larger than that obtained by the strategic planning part, which indicates that the storage location used by the current warehouse is not the best one.

5.3.2.5 Overall performance for tactical planning

To evaluate the performance of each storage location in the tactical planning, we calculated the proportion of goods in the simulation results of different storage locations.

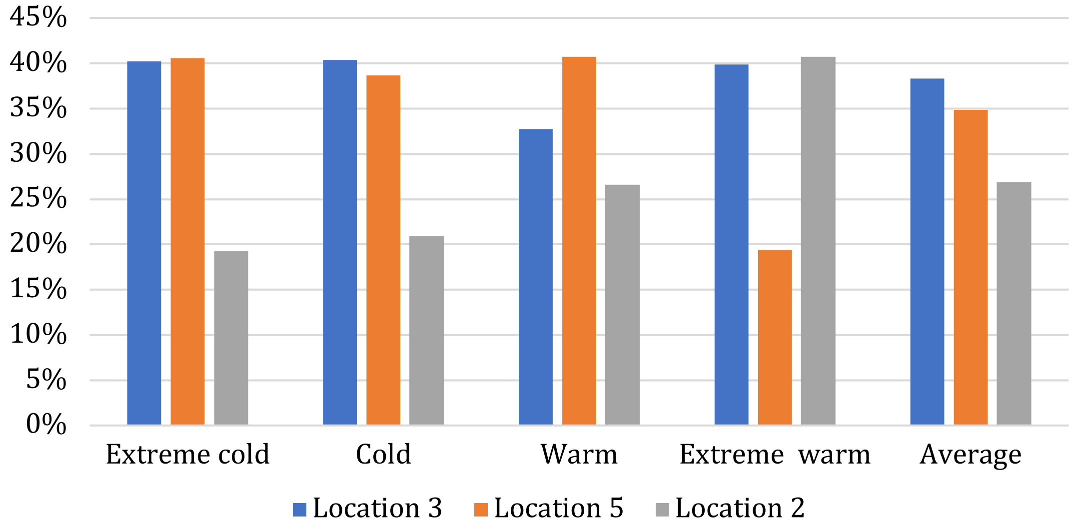


Figure 5. 16 The percentage of each location selected in the simulation.

In general, most of the goods are stored in the location where the sensor 3 is located (around 40%), regardless of whether the weather is cold or warm. For the average value, although the goods selected in the location where the sensor2 is located are the least, it can be seen that its proportion is increasing as the temperature increases (from 20% of extreme cold scenario to 40% of extreme warm scenario).

5.4.3 Comparison with the actual case

In reality, the sort criterion of the actual case is the time of goods in the warehouse. According to the results of the experiment in Section 5.4, our recommended sort criteria is the average quality loss in different storage locations. The simulation results based on the recommended scheme and the actual case are compared as follows.

The information about the shipments in this simulation experiment is only generated once. The average indoor temperature starts to decrease from 22 °C. The results according to the two different sort criteria are as follows.

As is shown in Figure 5.17, when the temperature is less than 16 °C, the quality loss is relatively small (less than 130 €) and difference is not obvious between two cases. But as temperatures rise, the quality loss of the actual case grows rapidly. Its quality loss value is around 1370 € at 22 °C, which is ten times as much as that at 16 °C. In contrast, the results based on the recommended control strategy can effectively reduce quality loss when compared with the real case. At 22 °C, the quality loss is only about 200 euros.

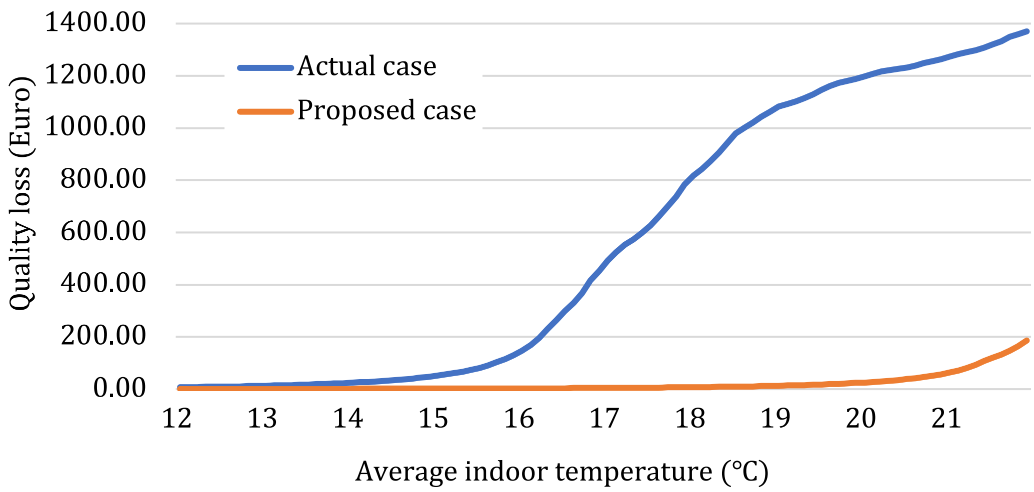


Figure 5. 17 Comparison between the actual case and the proposed case.

5.5 Discussion

In this chapter, we used the flowchart approach discussed in Section 3.3 as decision support method and simulated the quality loss distribution at different ambient temperatures based on the Monte Carlo approach. We set up four scenarios according to different weather conditions: extreme cold weather, cold weather, warm weather and extreme warm weather. In addition, based on strategic level and tactical level, we set the optional storage locations to the three currently in use and all the locations of the warehouse.

All the results were not normally distributed. Therefore, the bootstrapping method was introduced to calculate the confidence interval of the median of the quality loss for each distribution.

By comparing the experimental results generated by the three sort criteria, it can be concluded that sort criteria 2 (based on the average loss of shipments at all locations) can better reduce the total quality loss. In contrast, sorting by the time of arrival at the warehouse (sort criteria 1) results in significantly more quality loss, especially when the weather became warmer. Therefore, we recommend the use of sort criteria 2 as the basis for control strategy.

When comparing the results at tactical level and strategic level, it can be found that under the same scenario, the quality loss caused by tactical planning is greater, which indicates that the current temporary storage locations of the warehouse are not optimal.

Finally, we compared simulation results based on actual operations with those based on the proposed control strategy. The results showed that when the temperature is low, the quality loss caused by the two is similar, but with the increase of indoor temperature, the proposed control decision can better reduce the quality loss of products.

Chapter 6 Conclusions and Recommendations

In this final chapter, the conclusion on a temperature-based temporarily storage location choice model is presented. First, the main research questions and three sub-questions raised in Section 1.3 were answered. Additionally, the contribution of this study was discussed. Finally, the recommendations for future research and KLM Cargo were presented.

6.1 Answers to Research Questions

The problem researched in this thesis is that KLM Cargo currently does not have any control strategy considering temperate as the deciding factor to reduce the quality loss of perishable goods while the temperature may have a great impact on these shipments. Therefore, the main objective of this thesis is to develop a control strategy to reduce the quality loss when perishable goods (pharmaceuticals and flowers) are outside the storage temperature range in the stage of warehouse storage location choice.

To realize the objective, three sub-questions were organized. The first sub-question was interpreted as:

What is the relationship between the outside weather (temperature, humidity, wind, rainfall, etc.) and the indoor temperature at div storage locations?

In the data preparation phase, we collected data from1st June 2018 to 31st May 2019, contains hourly temperature data from 12 detectors inside the warehouse and hourly weather data outside the warehouse. After the five-step data processing in Section 4.1, we classified the data set according to whether the heating is turned or not. Each data set contains the following information: one dependent variable (temperature of each detector) and independent variables (corresponding outdoor temperature, wind speed, dew point value, precipitation amount, air pressure, cloud cover, humidity, rainfall or not, thunder or not).

To predict indoor temperature, we established the multiple linear regression model and the neural network model by R language. The RMSE value was used to evaluate model performance and the results for all models are shown in Appendix F. It was found that when the heating system was turned off, both models have good predictive power and the fitting degree of the neural network model is better than that of the multiple regression model. When the heating system was turned on, the RMSE value of both models of more than half of the detectors increased significantly. The main reason that under the condition of the heating, the indoor temperature of some detector areas was relatively stable and was not affected by the outside weather obviously.

In the process of model optimization, we found that the RMSE value for all models was very large at some time of the day. To reduce model errors as much as possible, we decided to do the model refinement. The prediction results showed that the refined model was closer to the real value. In addition, K-fold cross-validation approach was introduced to figure out the optimal parameters and to avoid the overfitting.

In general, the neural network model has a higher fitting degree than the multiple regression model for all sensors. Therefore, it was applied in the final simulation.

After choosing a suitable indoor temperature prediction model, we need to establish a control strategy to reduce the total quality loss of perishable goods in the location choice procedure. Thus, it is reasonable to ask the second sub-question:

How to design a control strategy to reduce total quality loss?

The control strategy consists of three models: indoor temperature prediction model, quality loss model and location choice model. The input and output parameters of each model are shown in Table 6.1. The approach to collect them are discussed in Section 3.1.

Table 6. 1 Inputs and outputs of the control strategy

Model	Inputs	Outputs
Indoor temperature prediction model	Outdoor temperature, Wind speed, Precipitation amount, Air pressure, Cloud cover, Humidity, Rainfall or not, Thunder or not.	Hourly indoor temperature at different locations of the warehouse.
Quality loss model	Product category (flower or pharmaceutical products), Arrival and departure time, Storage temperature range, Product value and quantity, Hourly storage temperature.	Estimated quality loss of product per hour at different locations.
Location choice model	Estimated quality loss per hour for each product at different locations, Simulation time, The capacity of each location.	Storage location allocation results, Estimated total quality loss of all goods.

To answer the second sub-question, we first created a quality loss model to measure the quality loss of medicines and flowers due to temperature. The final quality loss model is composed of the temperature-time quality loss function and the time-money quality loss model. The figure below shows the calculation process of quality loss. For the temperature-time quality loss functions, according to the relationship between temperature and flower vase life/ medicine shelf life, the hourly vase life/shelf life loss at a given temperature can be estimated. Then we can calculate the total vase life/shelf life loss of the product over a certain period. After that, we end up converting the total vase life/shelf life loss to money, based on the time-money quality loss model.

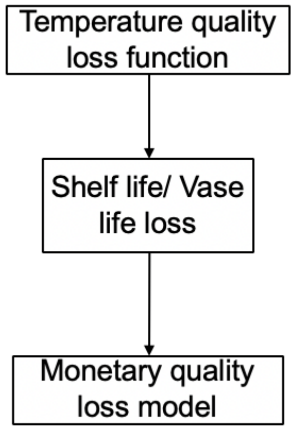


Figure 6. 1 The process of quality loss calculation.

For location choice model, we introduced two different approaches. The first one is the operations research model, which describes the functional relationship between each variable and the total quality loss. It requires a lot of modelling and solving efforts, which is out of the scope of my contribution. Therefore, I proposed a more straightforward method with a flowchart in Section 3.4.2.

The flowchart for storage locations choice is shown below. For the m_{th} shipment, after checking the storage capacity of each location, we will filter out all available storage locations. The quality loss of the shipment at each location is then calculated and the location with the least quality loss value will be selected. After that, the model updates the capacity information for each location and the $(m+1)_{th}$ shipment begins to choose the storage location.

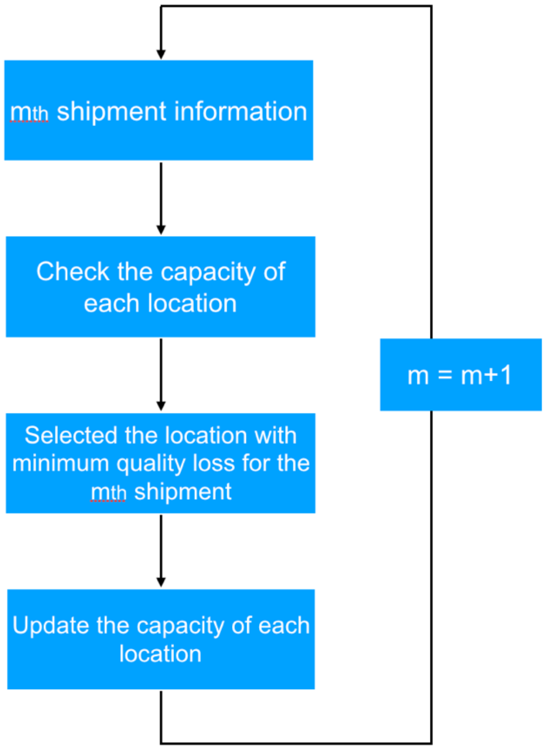


Figure 6. 2 The flowchart approach for storage location choice.

For the flowchart method, we need to decide in advance the sequence in which the goods are allocated. Since the sort criteria in the flowchart have a great influence on the result, we will try different sort criteria in the simulation experiment to find the best one in the control strategy. Therefore, it is inevitable to ask the third sub-question:

Given the information of perishable goods and weather forecasts, how much improvement can the control strategy provide according to the simulation results?

To answer this question, we first introduced three sort criteria of the flowchart approach: the estimated arrival time of the shipment at Schiphol; the average quality loss of the shipment at all storage locations; the standard deviation of the quality loss of the shipment at all storage locations.

In order to discover the effect of external weather on location choice results, we established four different scenarios using three sort criteria: the scenario of extreme cold weather; the scenario of cold weather; the scenario of warm weather; the scenario of extreme warm weather. In addition, we divided the simulation experiment into two parts, tactical planning and at strategic level, to explore how to optimize the selection of temporary storage locations from tactical and strategic perspectives.

The simulation relies on supercomputers at the Leibniz Supercomputer Centre. The simulation ran 10,000 times, with a total running time of about 24 hours. By comparing the experimental results, it can be concluded that sort criteria 2 (based on the average quality loss of shipments at all locations) can better reduce the total quality loss. Therefore, we recommend the use of sort criteria 2 as the basis for control strategy.

At the same time, we provided the percentage of each location selected in the simulation of strategic level and tactical level. This result helps us understand which locations are more suitable for temporary storage and provides the basis for KLM cargo to make adjustments in the future.

Last but not least, we compared the simulation results based on the recommended control strategy and the actual case. The results show that when the temperature is low, the quality loss caused by the two is similar, but with the increase of indoor temperature, the proposed control decision can better reduce the quality loss of products.

6.2 Contributions

Updated Literature Review

First, previous literature reviews have limited descriptions of KPIs for flowers. This thesis provides an updated version of this review. We summarized KPIs used in past studies to measure quality loss of flowers, including appearance indicators: flowerhead diameter; number of ray florets of flowerheads; height, number of buds; colour, appearance of leaves on the stem; physiological indicators: vase life, duration of flower development, flowering, inflorescences, respiration, fresh weight, electrical conductivity, malondialdehyde. The summary of this information will be helpful for the study of new flower quality loss models in the future.

Hourly temperature prediction

The previous temperature prediction models are generally accurate to the day. However, we designed an hourly temperature prediction model in this project. Our model needs to predict the temperature of every hour in a certain period, which needs to be more detailed. We also found that outdoor weather factors may have different effects on the indoor temperature at a different time of day. In order to make the model more accurate, we divided the model of each sensor into four small models according to the four time periods of each day. All these attempts are helpful for the development of real-time temperature prediction models in the future.

Validation of different prediction models

In this study, the k-fold cross validation method was used for the first time to validate both the multiple linear regression model and the neural network analysis model. In the past few studies have mentioned the validation of temperature prediction models. In addition, although some referred to multiple linear regression validation or neural network validation, no study has been conducted to apply the unified verification method to both models.

Quality loss model for multiple products

In previous studies, no generic model was applied to measure the quality loss of multiple products caused by temperature. In this study, a temperature-based model was designed to measure the quality loss of different types perishable products. Using this model, we can predict the quality loss (€) according to the storage temperature. At present, this model is mainly applicable to flowers and pharmaceutical products, but it provides a design idea for the quality loss model of other kinds of goods in the future.

Temperature based storage location choice model

For the previous location selection model, they mainly considered factors including transportation distance, transportation time and transportation cost. For KLM Cargo, the KPIs used in previous operations were time-dependent and they did not take temperature into account. The storage location choice model in this study takes temperature as a reference factor and aims to reduce product quality loss, which has rarely been shown in previous studies. In addition, the control strategy in this study combined temperature prediction, quality loss assessment, and location choice, which has not been shown in previous studies.

6.3 Recommendations for future study

Temperature prediction model

For the temperature prediction model, we mainly consider the influence of outdoor weather on indoor temperature, while some other factors are not considered. For example, the distance between the temporary storage location and the entrance, the distance from the cool room may have a hug influence on the indoor temperature. Therefore, the temperature prediction model can be improved based on the actual situation of the warehouse for KLM Cargo in the future.

Quality loss model for perishable goods

Since there have been few previous studies on the quality loss of perishable products, this has created many challenges for this study. For our quality loss model, we hypothesized some parameters of flowers and pharmaceutical products. If more studies are completed in the future, we can set more reasonable parameters in the model, and also set different parameters for different kinds of flowers and medicines. We expect to see more quality loss models for pharmaceutical products or flowers from other studies in the future.

KPIs for perishable goods

In order to quantify product quality loss, we used vase life and shelf life as KPIs for the quality loss model. This study is the first attempt to use the same KPIS to measure different types of perishable products. In the future, we hope to have more temperature based KPIs that can be used to measure the quality loss of products from different perspectives.

Storage location choice model

The storage location choice model in this study is based on the flowchart approach. This method can only reduce the quality loss of shipments to a certain extent, but it may not be the optimal solution. The operations research approach described in 3.4.2 can be used to find the optimal solution and is worthy of further study in the future.

6.4 Recommendations for KLM Cargo

Error of temperature detector

In this study, we ignored the measurement error caused by the temperature detector itself, while the measurement error range of the current common temperature detector is 1.5°C. Therefore, it is necessary to screen out the temperature outliers caused by the temperature detector itself in the data processing stage.

Kalman filter is a widely used tool for filtering temperature outliers. This filtering model provides an intuitive prediction-correction method for state estimation, in which it balances model predictions and data according to the degree of certainty of both. However, this method can only be used if the initial measurement error and estimated error are known. Therefore, if KLM cargo can provide relevant data in the future, it will make our prediction more accurate.

Historical data record

In this study, because the temperature detection equipment in some locations has only been used since the spring of 2019, a considerable part of the data volume is missing. Therefore, the model is divided according to whether the heating system is turned on or not. But previous studies have shown that the relationship between indoor and outdoor temperatures varies from season to season. Therefore, if there is new data available in the future, we can refine the temperature prediction model according to seasons, which is conducive to improving the prediction ability. Furthermore, KLM cargo did not record when the heating was turned on, so our data set for with/out heating were obtained based on statistical analysis. In future studies, it will be more conducive to the establishment of the prediction model if the specific time information of heating was turned on can be obtained.

KPIs for perishable goods

Currently, the KPIs used by KLM cargo are mainly measured by time, such as time out of refrigeration, storage time in the warehouse. These KPIs ignore the effect of temperature on product quality loss. In the future, we hope KLM Cargo can gradually try to use some temperature-based KPIs in practical decisions.

Temporary storage location choice

Through the simulation experiment in Chapter 5, we provided the percentage of each location selected in the simulation at strategic level and tactical level. We found that for the three storage locations currently in use, location 3 was the best for both hot and cold weather. In the long run, locations 18, 13, 20 and 19 are more suitable for temporary storage than others. It is hoped that these simulation results will serve as a basis for KLM cargo to adjust the layout and planning of temporary storage locations in the future.

Appendix A: Scientific paper

Temperature-based control strategy for temporary storage location choice of perishable goods.

***Abstract*— Ambient temperature has been identified as the most common reason for quality loss of most high value perishable goods. In actual operation, it is not cost-effective to cool perishable products in active and/or passive containers when their storage time in the warehouse is short. In order to o reduce the quality loss caused by the temperature, this research proposed a control strategy to find suitable temporary storage locations of the warehouse in advance for different types of perishable goods. The control strategy consists of indoor temperature prediction model, quality loss model and location choice model. By comparing with the actual case data, it is found that when the indoor average temperature increases, the proposed control strategy can significantly reduce the quality loss.**

***Keywords*– perishable goods, storage, logistics, transportation**

I. Introduction

1.1 Perishable goods supply chain

With the consideration of health, environment and convenience, today's consumer preferences are shifting toward fresh, and healthy consumable products, thus the demands of perishable goods are gradually growing(Diop & Jaffee, 2005). Research reveals that perishable products accounted for 38% of total-store sales while they made up about 65% of total store shrink for 43 countries in 2011 (Cognizant, 2015). This lifestyle change creates new business opportunities for logistics service providers, who need to provide more convenience to consumers through perishable goods that can be prepared, transported and supplied quickly and reliably.

Ambient temperature has the greatest impact on affecting the shelf-life of perishable goods among all of the particular factors (Mc Carthy et al., 2018). It is estimated that due to the poor processing technology after harvest, including poor temperature management, the total waste of agricultural products before the consumption stage accounts for about 60% in the supply chain (Gustavsson, Cederberg, Sonesson, Van Otterdijk, & Meybeck, 2011). For pharmaceutical products, it is estimated that as about 15 to 20 percent of the products are wasted during transportation due to storage temperature and have to be discarded upon delivery (Sales, 2016). Thus, temperature management is important to reduce the loss of perishable products.

Perishable supply chains are more complex than other product supply chains due to the short shelf-life of goods, which requires the cooperation of product growers, retailers, carriers and logistics companies. Ensuring the quality of these perishable products and providing the freshest products to customers is also one of the challenges that researchers are struggling with. The earliest research in this field can be traced back to (Straughn & Church, 1909), who studied the quality decline of agricultural products from the first harvest to the warehouse (Paterson, Tanksley, & Sorrells, 1991). Thanks to the development of technology, some new technologies are constantly emerging to control the quality of perishable products. For instance, the Wireless Sensor Network (WSN) has been proposed to provide real-time environmental information about perishable goods to decision-makers (Wang et al., 2015). Radio Frequency Identification (RFID) can reliably gather environmental information through sensors and help reduce waste (Lorite et al., 2017). These technologies bring new insights to the temperature-based quality control of perishable products.

1.2 Quality loss evaluation

Key performance indicators (KPIs) are quantifiable metrics used by organizations to measure the performance of products over time. For the floral industry, much previous research has used vase life as an indicator of changes in the quality of flowers. Hettiarachchi and Balas compared the changes of vase life under different storage environments and revealed that floral preservatives and stem-end treatments could better guarantee the quality of flowers (Hettiarachchi & Balas, 2003). Zheng and Guo studied the influence of cerium on the quality of carnation and found that cerium helps to regenerate ascorbate and glutathione, thus extending the vase life (Zheng & Guo, 2018). Shelf life is an important KPI of pharmaceutical products. This index has been widely used in previous studies for various types of pharmaceutical products, such as N-acetylcysteine, vitamin C, ascorbic acid, metronidazole infusion and isotretinoin products (Jaworska, Szulińska, Wilk, & Tautt, 1999; Moussa & Haushey, 2015; Shaheen, 2005; Taylor & Keenan, 2006).

The product quality loss model can be traced back to the quality control of television products by Sony-USA (Ealey, 1988). Products are good when their performance is within the limits of the upper and lower specification, and bad when they are outside the range of specifications. To make the quality loss model consistent with the facts, Japanese engineer Genichi Taguchi introduced a new concept of a loss function, which is an effective method for quality engineering.

1.3 Research background

It is important that perishable products are not unnecessarily exposed to too high or too low temperatures for too long. Here, we take an airport warehouse of an airline as the case background. The goods carried by the company mainly include pharmaceutical products and flowers. There are three main types of pharmaceutical products: COL, CRT and PIL, and their corresponding storage temperature range is 2 °C to 8 °C, 15 °C to 25 °C and 2 °C to 25 °C respectively. For fresh products such as flowers, they only have the two product categories of COL and PIL. If the storage time in the warehouse for COL or CRT products is more than eight hours, they will be stored in active and/or passive containers with active temperature control. The temperature inside the container can be adjusted to maintain the specified temperature range.

However, it is not cost-effective to cool perishable products in active and/or passive containers when they are broken down into cases and stored in temporary storage locations. For these locations, there is no temperature control equipment to ensure the temperature is within the storage temperature range of goods. When the ambient temperature exceeds the storage temperature range, the quality of fresh products in these locations may be affected.

1.4 Paper scope

In actual operation, it is not guaranteed that perishable products will always be stored in the temperature-controlled equipment, especially in the case of transhipment. For some perishable shipments (mainly pharmaceuticals and flowers), when the ambient temperature exceeds the storage temperature range, the quality of products in these locations may be affected. This paper aims to propose a control strategy to find suitable temporary storage locations of the warehouse in advance for different types of goods, so as to reduce the quality loss caused when the temperature is higher than the recommended storage temperature. The remainder of this paper is organized as follows. In Section 2, we introduce the control strategy for temporary location choice of perishable goods. Section 3, we apply the control strategy in simulation and compare it with the actual case. In Section 4, we discuss the results of this paper and presents recommendations for future research.

II. METHODOLOGY

The control strategy consists of three models: indoor temperature prediction model, quality loss model and location choice model. The input and output parameters of each model are shown in Table 1.

Table 1. Inputs and outputs of the control strategy

Model	Inputs	Outputs
Indoor temperature prediction model	Outdoor temperature, Wind speed, Precipitation amount, Air pressure, Cloud cover, Humidity, Rainfall or not, Thunder or not.	Hourly indoor temperature at different locations of the warehouse.
Quality loss model	Product category, Arrival and departure time, Storage temperature range, Product value and quantity, Hourly storage temperature.	Estimated quality loss of product per hour at different locations.
Location choice model	Quality loss per hour for each product at each location, Simulation time, The capacity of each location.	Storage location allocation results, Estimated total quality loss of all goods.

2.1 Temperature prediction model

2.1.1 Model structure

In this section, the neural network model for indoor temperature prediction will be introduced. As a subset of machine learning, Artificial Neural Network (ANN) model can simulate complex and non-linear relationship between input and output streams based on mathematical regression (Yang & Chen, 2015). A typical ANN architecture consists of three layers: input layer, hidden layer and output layer. Input parameters are stored in the input layer and fed into the system for processing. The role of hidden neurons is to intervene in a useful way between external input and network output. Result parameters are included in the output layer depending on the purpose of the system.

According to Table 1, the neural network model is composed of 9 variables for models without/with the heating system (8 inputs and 1 output). The warehouse internal temperature is the target variable. 80% of the training data set is used for learning, and 20% of the training data set is used for model validation. For both data sets, the hidden layer number of the ANN was set to 4, 8, 16 and 17.

Here, the "brnn package" of R can be used to build neural networks. This package was developed based on Nguyen algorithm to assign initial weights and was optimized by Gauss-Newton algorithm (Nguyen & Widrow, 1990). The optimization function of R will minimize:

$$F = \beta E_D + \alpha E_W \tag{1}$$

where E_D is the error sum of squares, E_W is the sum of squares of network parameters (weights and biases), $\beta = \frac{1}{2\sigma_e^2}$, $\alpha = \frac{1}{2\sigma_\theta^2}$, σ_θ^2 is a dispersion parameter for weights and biases.

After modelling, the overall significance of the model should be checked by the Root mean squared error (RMSE). Here, we use 1.5 °C as a threshold to measure the accuracy of the model. The value of 1.5°C is derived from the temperature accuracy of commonly used temperature sensors.

2.1.2 Model validation

In the process of machine learning, the final model usually overfits the training data. Therefore, the cross-validation method is introduced to prevent overfitting due to model complexity and estimate the skill of machine learning models. K-fold cross-validation method is to divide a given data set into K folds, where each fold is used as a testing set at a certain point (James, Witten, Hastie, & Tibshirani, 2013). Typically, k =10 is chosen in the field of applied machine learning (Refaeilzadeh, Tang, & Liu, 2009).

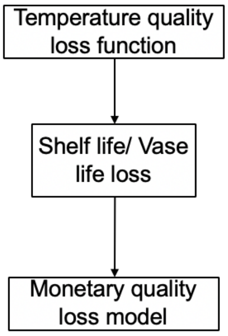
2.2 Quality loss model

2.2.1 Model structure

The final quality loss model is composed of the temperature-time quality loss function and the time-money quality loss model. The figure below shows the calculation process of quality loss.

For the temperature-time quality loss functions, according to the relationship between temperature and flower vase life/ medicine shelf life, the hourly vase life/shelf life loss at a given temperature can be estimated. Then we can calculate the total vase life/shelf life loss of the product over a certain period. After that, we end up converting the total vase life/shelf life loss to money, based on the time-money quality loss model.

Fig. 1. The process of quality loss calculation.



2.2.2 Time-money quality loss model

Based on the smaller-is-better loss function developed by Taguchi (Kumar, Chandrakar, Kumar, & Chandrakar, 2012), a time-based quality loss model has been built.

$$L(y) = k.(y)^2, k = A/\Delta^2 \tag{2}$$

where L(y) (€/unit) is the monetary loss caused by 100% product quality loss per unit of products; y (hours) is the loss of vase life for flowers and the loss of shelf life for pharma; Δ (hours) is the maximum vase life for flowers or maximum shelf life for pharma and A (€) is the price per unit of products.

2.2.3 Vase life/shelf life loss function

Take gypsophila as an example. Based on the experiments from Çelikel's, when the storage temperature deviates from the optimal storage temperature ΔT, the corresponding vase life is (Çelikel & Reid, 2004):

$$\text{Vase life}_{(\Delta T)} = 370.13 - 65.91 \cdot e^{0.1092 \cdot \Delta T} \text{ (hours)} \tag{3}$$

The corresponding vase life loss is:

$$\text{Vase life loss}_{(\Delta T)} = \text{Maximum vase life} - \text{vase life}_{(\Delta T)}$$
$$= -65.91 + 65.91 \cdot e^{0.1092 \cdot \Delta T} \text{ (hours)} \tag{4}$$

The hourly vase life loss is:

$$\text{Hourly vase life loss}_{(\Delta T)} = \text{Vase life loss}_{(\Delta T)} / \text{Vase life}_{(\Delta T)} \tag{5}$$

The pharma shelf life loss function was established according to Q10 model. Here, a drug is assumed to be moderately sensitive to temperature. Its shelf life is S(T1) (unit: hour) when the storage temperature is T1. When the storage temperature is T2 (ΔT = T2 − T1), the corresponding shelf life is:

$$S_{(T2)} = \frac{S^{(T1)}}{4^{\left(\frac{T2-T1}{10}\right)}} \text{ (hours)} \tag{6}$$

The corresponding shelf life loss is:

$$\text{Shelf life loss}_{(\Delta T)} = S(T1) - S(T2) \text{ (hours)} \tag{7}$$

The shelf life loss per hour at T2 is:

$$\text{Hourly vase life loss}_{(\Delta T)} = \text{Shelf life loss}_{(\Delta T)} / S_{(T2)} \tag{8}$$

2.3 Location choice model

In this section, the location choice approach with a flowchart is proposed. The flowchart for selecting storage locations of shipments is shown in Fig. 2. For the m_{th} shipment, after checking the storage capacity of each location, we will screen out all available storage locations. The quality loss at each location for this shipment is then calculated and the one with the least value is selected. After this, the model updates the storage capacity information for each location and begins to select storage locations for the $(m+1)_{th}$ shipment. For this model, we need to determine in advance the order in which shipments are selected for storage. Sorting rules and simulation results of this model are discussed in Section 3.

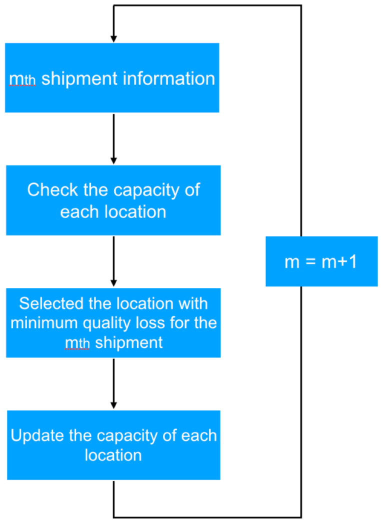


Fig. 2. The flowchart approach for storage location choice.

III. Simulation

In this section, we will try different sort criteria in the simulation experiment to find the best one in the control strategy. In addition, we will simulate the quality loss based on the actual case and compare it with the proposed control strategy of Section 2.

3.1 Indoor temperature prediction

3.1.1 Data preparation

This initial data set covering from 1st June 2018 to 31st May 2019, contains hourly temperature data from 12 detectors inside the warehouse and hourly weather data outside the warehouse.

The inside temperature exceeding the threshold that greater than or less than the daily average value five times the standard deviation will be considered outliers and removed from the data set. Since the warehouse is equipped with heating facilities. The data set is classified according to whether the heating is turned on or not. After data processing, each data set contains the following information: a dependent variable (temperature of each detector) and independent variables (corresponding outdoor temperature, wind speed, dew point value, precipitation amount, air pressure, cloud cover, humidity, rainfall or not, thunder or not).

3.1.2 Model performance

Models with the smallest RMSE are selected as final neural networks models for detectors, which are shown below. Models without the heating system fit better overall than models with the heating system. For all models, all of them have RMSE values less than 1.5. For models without the heating system, RMSE values of detectors #1, #2 are close to 1.5, which indicates that the model cannot well describe the relationship between independent variables and dependent variables.

Table 2. Selected ANN models.

Location	Without heating		With heating	
	RMSE	Model structure	RMSE	Model structure
1	1.41		0.54	
2	0.82		0.56	
3	1.19		0.56	
5	1.47		0.56	
15	0.87		1.38	
16	1.13		1.37	
18	1.14	8-4-1	1.28	8-8-1
19	0.96		1.36	
20	0.88		1.17	
21	0.83		1.08	
22	0.71		1.19	
23	0.75		1.32	

3.2 Control strategy implementation

The simulation relies on the supercomputer at the Leibniz Supercomputer Centre. Simulations will be run 10,000 times. The computation time for each model is around three hours and the total computation time is around 24 hours. During the simulation, goods information will be generated randomly. The Settings of simulation parameters are described in Table 3.

Table 3. Parameters setting of the simulation

[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

3.2.1 Sort criteria

The simulation in this chapter is based on the flowchart method described in Section 2 for decision making. We need to decide in advance the sequence in which the goods are allocated. To test the method, we assume three possible cases.

- Case 1: The information about the shipments arriving in the next 72 hours is constantly updated.
- Case 2: All information about the shipments arriving in the next 72 hours is available.

For the first case, we will take the estimated arrival time of shipments as the basis for determining the sequence. Goods arriving early have the right to choose storage locations first. For the second case, we will determine the sequence based on the mean and standard deviation of the quality loss of each good at all storage locations.

3.2.2 Scenarios Building

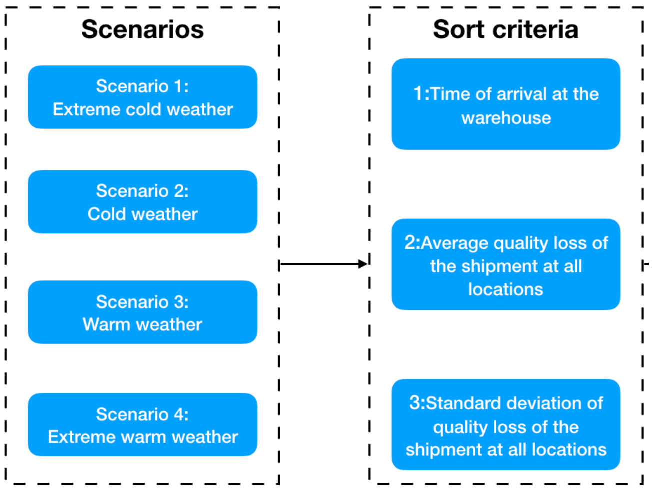


Fig. 3. Comparison between the actual case and the proposed case.

As shown in Fig. 3, in order to understand the impact of different weather conditions on the results of storage location selection and figure out the best sort criteria, we will simulate four different scenarios using three sort criteria. By analysing meteorological data from June 1, 2018 to May 31, 2019, we selected four representative periods.

3.2.3 Simulation results

After we have the final distributions of results, we first assessed whether the distribution is normal. Based on the goodness of fit test, it was found that these three distributions are not normal. Thus, we applied the bootstrapping method to calculate confidence intervals of the median of the quality loss value, which should be the median of the distribution. The results are shown in Table 4.

By comparing the experimental results, it can be concluded that sort criteria 2 (based on the average quality loss of shipments at all locations) can better reduce the total quality loss. Therefore, we recommend the use of sort criteria 2 as the basis for control strategy.

Table 4. 95% Confidence interval of quality loss (Euro) based on the Bootstrap approach.

		Sort criteria 1	Sort criteria 2	Sort criteria 3
Extreme cold	Lower bound	32.38	27.13	27.23
	Upper bound	32.71	27.37	27.51
Cold	Lower bound	42.37	37.44	38.09
	Upper bound	42.76	37.78	38.44
Hot	Lower bound	152.28	105.15	109.12
	Upper bound	154.08	106.27	110.33
Extreme hot	Lower bound	221.07	130.91	131.16
	Upper bound	223.67	132.32	132.55

Then, we compared the simulation results based on the proposed control strategy with sort criteria 2 and the actual case. At present, this company takes time into account when selecting the location of its shipments. The selection of temporary storage locations is mainly based on the storage time in the warehouse. To make the simulation more realistic, it is assumed that this company currently uses a sort criterion based on the length of time the shipment is stored in the warehouse. Products that stay in the warehouse for a long time can have the right to choose the storage location first.

The average indoor temperature starts to decrease from 22 °C. As is shown in Fig. 4, when the temperature is less than 16 °C, the quality loss is relatively small (less than 130 €) and difference is not obvious between two cases. But as temperatures rise, the quality loss of the actual case grows rapidly. Its quality loss value is around 1370 € at 22 °C, which is ten times as much as that at 16 °C. In contrast, the results based on the recommended control strategy can effectively reduce quality loss when compared with the real case. At 22 °C, the quality loss is only about 200 euros.

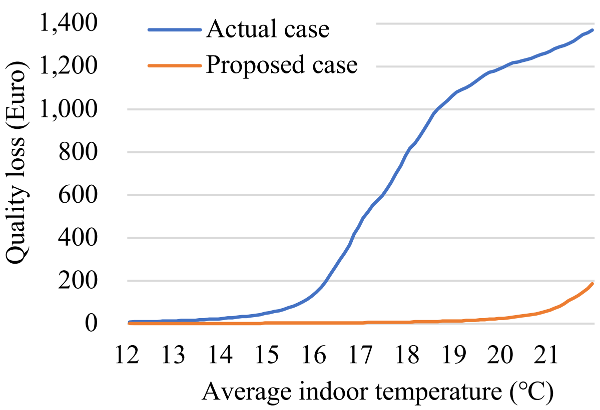


Fig. 4. Comparison between the actual case and the proposed case.

IV. Conclusion

This paper presents a control strategy to reduce the quality loss of perishable goods (pharmaceuticals and flowers) when are in the stage of warehouse storage location choice. The control strategy consists of three models: indoor temperature prediction model, quality loss model and location choice model. To predict indoor temperature, we propose the neural network model. The quality loss model is composed of the temperature-time quality loss function and the time-money quality loss model. The flowchart for storage locations choice is also introduced.

For this control strategy, we need to decide in advance the sequence in which the goods are allocated. Therefore, the simulation experiment is provided to find out the best sort criteria as well as to demonstrate the advantages of the proposed control strategy. For indoor temperature prediction, it is found that when the heating system is turned off, the neural network model has good predictive power. When the heating system is turned on, the RMSE value of both models of more than half of the detectors increase significantly. The main reason that under the condition of the heating, the indoor temperature of some detector areas is relatively stable and is not affected by the outside weather obviously. Then, we introduce three sort criteria of the flowchart approach and find that the average quality loss of the shipment at all storage locations is the best sort criteria based on simulation results. We also compare the simulation results based on the recommended control strategy with the best sort criteria and the actual case. The results show that when the temperature is low, the quality loss caused by the two is similar, but with the increase of indoor temperature, the proposed control decision can better reduce the quality loss of products.

In this study, a temperature-based model is designed to measure the quality loss of different types perishable products. At present, this model is mainly applicable to flowers and pharmaceutical products, but it provides a design idea for the quality loss model of other kinds of goods in the future. In addition, we design a temperature-based control strategy to select storage locations, which combines temperature prediction, quality loss assessment, and location choice. It has not been shown in previous studies.

Since there have been few previous studies on the quality loss of perishable products, this has created many challenges for this study. The quality loss model in this research only for flowers and pharmaceutical products. We expect to see more quality loss models for other types of perishable products in the future. Besides, the storage location choice model in this study is based on the flowchart approach. This method can only reduce the quality loss of shipments to a certain extent. The operations research approach can be used to find the optimal solution and is worthy of further study in the future.

V. Bibliography

Çelikel, F. G., & Reid, M. S. (2004). Temperature and postharvest performance of rose (Rosa hybrida L. First Red) and gypsophila (Gypsophila paniculata L. Bristol Fairy) flowers. V International Postharvest Symposium 682, 1789–1794.

Cognizant. (2015).https://www.cognizant.com/InsightsWhitepapers/high-velocity- supply-chains-for-perishables-a-fresher-proposition-codex1325.pdf, (Accessed 13-8-2019).

Diop, N., & Jaffee, S. (2005). Fruits and vegetables: Global trade and competition in fresh and processed product markets. Global Agricultural Trade and Developing Countries, 237–257.

Ealey, L. (1988). Quality by design. Taguchi Methods and US industry. Dearborn. MI: American Supplier Institute.

Foresee, F. D., & Hagan, M. T. (1997). Gauss-Newton approximation to Bayesian learning. Proceedings of International Conference on Neural Networks (ICNN'97), 3, 1930–1935. IEEE.

Gustavsson, J., Cederberg, C., Sonesson, U., Van Otterdijk, R., & Meybeck, A. (2011). Global food losses and food waste. FAO Rome.

Haykin, S. (1999). Self-organizing maps. Neural Networks-A Comprehensive Foundation, 2nd Edition, Prentice-Hall.

Hettiarachchi, M. P., & Balas, J. (2003). Postharvest quality of cut anthurium flowers (Anthurium andraeanum L.) after long-distance shipment. VIII International Symposium on Postharvest Physiology of Ornamental Plants 669, 329–336.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112). Springer.

Jaworska, M., Szulińska, G., Wilk, M., & Tautt, J. (1999). Capillary electrophoretic separation of N-acetylcysteine and its impurities as a method for quality control of pharmaceuticals. Journal of Chromatography A, 853(1–2), 479–485.

Kim, S. Y., Lee, S. M., Park, K. S., & Ryu, K. H. (2018). Prediction Model of Internal Temperature using Backpropagation Algorithm for Climate Control in Greenhouse. Horticultural Science and Technology, to Be Submitted.

Kumar, R., Chandrakar, R., Kumar, A., & Chandrakar, H. R. (2012). Taguchi loss function as optimized model for supplier selection and evaluation. International Journal of Advanced Engineering Technology, 3(1), 268–270.

Lorite, G. S., Selkälä, T., Sipola, T., Palenzuela, J., Jubete, E., Viñuales, A., … Uusitalo, S. (2017). Novel, smart and RFID assisted critical temperature indicator for supply chain monitoring. Journal of Food Engineering, 193, 20–28.

MacKay, D. J. (1992). Bayesian interpolation. Neural Computation, 4(3), 415–447.

Mc Carthy, U., Uysal, I., Badia-Melis, R., Mercier, S., O'Donnell, C., & Ktenioudaki, A. (2018). Global food security–Issues, challenges and technological solutions. Trends in Food Science & Technology, 77, 11–20.

Moussa, N., & Haushey, L. A. (2015). The shelf life of Vitamin C in aw/o emulsion according to the Q10 Method. International Journal of Pharmaceutical Sciences Review and Research, 30, 33–39.

Nguyen, D., & Widrow, B. (1990). Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. 1990 IJCNN International Joint Conference on Neural Networks, 21–26. IEEE.

Paterson, A. H., Tanksley, S. D., & Sorrells, M. E. (1991). DNA markers in plant improvement. In Advances in agronomy (Vol. 46, pp. 39–90). Elsevier.

Refaeilzadeh, P., Tang, L., & Liu, H. (2009). Cross-validation. Encyclopedia of Database Systems, 532–538.

Sales, M. (2016). Aviation logistics: The dynamic partnership of air freight and supply chain. Kogan Page Publishers.

Shaheen, S. M. (2005). Accelerated stability study of metronidazole infusion 100 ml. TAJ: Journal of Teachers Association, 18(2), 118–121.

Straughn, M. N., & Church, C. G. (1909). The Influence of Environment on the Composition of Sweet Corn, 1905-1908. US Government Printing Office.

Taylor, P. W., & Keenan, M. H. (2006). Pharmaceutical quality of generic isotretinoin products, compared with Roaccutane. Current Medical Research and Opinion, 22(3), 603–615.

Wang, J., Wang, H., He, J., Li, L., Shen, M., Tan, X., … Zheng, L. (2015). Wireless sensor network for real-time perishable food supply chain management. Computers and Electronics in Agriculture, 110, 196–207.

Yang, H., & Chen, H. (2015). Biomass gasification for synthetic liquid fuel production. In Gasification for Synthetic Fuel Production (pp. 241–275). Elsevier.

Zheng, M., & Guo, Y. (2018). Cerium improves the vase life of Dianthus caryophyllus cut flower by regulating the ascorbate and glutathione metabolism. Scientia Horticulturae, 240, 492–495.

Appendix B: Classification criteria of heating and non-heating data

The heating system is turned on from September to April of the next year. When the system detects the indoor temperature below 16 degrees Celsius, the warehouse will be warmed automatically. According to this control strategy, we cannot directly distinguish whether the indoor temperature data is in the heating situation or not. Moreover, this heating system does not record the specific turn-on time. Thus, it is necessary to distinguish the data with and without the heating system based on data analysis.

Assumptions:

1) If the heating system is not activated, the relationship between indoor average temperature and outdoor temperature from May to August was the same as that between September and April.

2) If the heating system is not activated, there is a linear relationship between indoor average temperature and outdoor temperature.

Based on the setting of the heating system, it can be confirmed that the heating system is closed from May to August every year. Here, the average indoor temperature per hour and the corresponding outdoor temperature are calculated during these four months. By using SPSS 25.0 software, the function between indoor average temperature without heating system and outdoor temperature is obtained. The parameters and performance indicators of the function are as follows:

Table 1 Model coefficients

Model	Unstandardized Coefficients	
	B	Std. Error
Constant	11.858	0.095
Outdoor temperature	0.484	0.005

Table 2 Model summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
0.86	0.74	0.74	1.52

According to Table 2, the linear fitting results are relatively ideal, and indoor and outdoor air temperature conforms to the linear relationship. The residuals of the linear model follow a normal distribution with a 95% confidence interval of +/-0.056 degrees Celsius. It is assumed that when the residual is within the confidence interval, the corresponding indoor average temperature follows to the above linear relationship, that is, the corresponding indoor average temperature is in the state of without heating. By using this linear model, the predicted value of indoor average temperature from April to September can be calculated. We can also obtain the residuals between the predicted and true values of the average indoor temperature.

Here, the actual indoor average temperature data below 16 degrees Celsius from September to April will be deleted. This is because when the actual indoor temperature is lower than 16 degrees Celsius, it is difficult to judge whether the heating is turned on based on purely data analysis: for example, when the outdoor temperature is 0 degrees Celsius, the corresponding predicted indoor temperature is 12.05 degrees Celsius, while the real temperature is 12 Celsius degrees. According to the data analysis, it fits the linear model and is classified as non-heating data. But when the average indoor temperature is 12 degrees, it is impossible that the heating system is turned off.

To sum up, the classification criteria for heating and non-heating data are as follows:

1) Heating data: Actual indoor average temperature data is higher than 16 degrees Celsius; The residual between the predicted and true values is higher than 0.056 degrees Celsius.

2) Non-heating data: Actual indoor average temperature data is higher than 16 degrees Celsius; The residual between the predicted and true values is between +/-0.056 degrees Celsius.

Appendix C: The evaluation results of multiple regression models including all effective meteorological parameters

The change of R-squared value and standard deviation can reflect the importance of deleted variables to the model. It can be concluded that although sunshine duration, radiation, precipitation duration and fog amount cannot be obtained in the weather forecast, they do have an impact on the prediction of indoor temperature.

Table 3 Comparison of RMSE values for model with all effective meteorological parameters and model based on the step 4 of data processing (Without heating sub-data sets)

Detector	Model with all effective meteorological parameters	Model based on the step 4 of data processing
1	1.27	1.57
2	0.82	0.86
3	1.01	1.27
5	1.36	1.62
15	0.83	0.94
16	1.03	1.27
18	1.04	1.22
19	1.00	1.10
20	0.85	1.04
21	0.82	0.95
22	0.73	0.78
23	0.75	0.80

Table 4 Comparison of RMSE values for model with all effective meteorological parameters and model based on the step 4 of data processing (With heating sub-data sets)

Detector	Model with all effective meteorological parameters	Model based on the step 4 of data processing
1	0.75	0.79
2	0.59	0.63
3	0.80	0.83
5	1.43	1.45
15	1.98	2.00
16	2.11	2.15
18	3.17	3.21
19	1.97	2.00
20	1.96	2.04
21	1.53	1.50
22	1.46	1.47
23	1.70	1.70

Appendix D: Code for building the neural network

```
#install.packages("brnn")
#document location
setwd("/Users/kangzhao/Downloads/")
#setwd("~/Desktop/NN")
library('plyr')
library('Metrics')
library('brnn')
library("caret")

#train and test
raw_data <- read.csv("without heating.csv", header=T)
#raw 1-15: temperature data of sensor 1-15; raw 16-22: outdoor weather data
data <- raw_data[,c(1,16,17,18,19,20,21,22)]
data <-data[complete.cases(data[,1]),]
split <-createDataPartition(y=data$SENSOR_1, p=0.8, list=FALSE)
train <- data[split, ]
test <- data[-split, ]
#neurons=number of hidden layers
nn <- brnn(SENSOR_1 ~
TEMP+WINS+DEWP+PREA+AIRP+CLOC+HUM,neurons=7,train)
p<-predict(nn,test)
t<-test$SENSOR_25
mse(t,p)
rmse(t,p)
relation<-lm(p~t)
print(summary(relation))
nn

#k-cross validatation
cv.error <- NULL
k <- 10
for(i in 1:k){
  index <- sample(1:nrow(data),round(0.8*nrow(data)))
  train.cv <- data[index,]
  test.cv <- data[-index,]
```

```
   nn<-brnn(SENSOR_1 ~
TEMP+WINS+DEWP+PREA+AIRP+CLOC+HUM,neurons=7,data=train.cv)
   pr.nn<-predict(nn,test.cv)
   test.cv.r <- test.cv$SENSOR_1
   cv.error[i] <- mse(pr.nn,test.cv.r)
}
mean(cv.error)
cv.error
boxplot(cv.error,xlab='MSE CV',col='cyan',
            border='blue',names='CV error (MSE)',horizontal=TRUE)
```

Appendix E: Parameters of each regression model

Table 5 Selected parameters of each regression model

Sensor	Model	Without heating	With heating
1	sum	TEM+WIN+AIR+CLO+HUM+RAI+THU	TEM+WIN+AIR+CLO+RAI
	1-6	TEM+WIN+AIR+CLO+HUM+RAI	TEM+WIN+AIR+CLO+HUM+RAI
	7-12	TEM+WIN+AIR+CLO+HUM+RAI+THU	TEM+WIN+AIR+CLO+HUM+THU
	13-18	TEM+AIR+CLO+HUM+RAI	TEM+WIN+AIR+CLO+HUM+RAI
	19-0	TEM+WIN+PRE+AIR+CLO+HUM+RAI	TEM+WIN+PRE+AIR+CLO+RAI
2	sum	TEM+HUM+RAI	TEM+WIN+PRE+AIR+CLO+HUM
	1-6	TEM+RAI	TEM+WIN+PRE+AIR+CLO+HUM
	7-12	TEM+HUM+RAI	TEM+RAI
	13-18	TEM+WIN+RAI	TEM+WIN
	19-0	TEM+WIN+RAI	TEM+WIN+PRE+CLO
3	sum	TEM+WIN+AIR+CLO+HUM+RAI+THU	TEM+WIN+PRE+CLO+HUM
	1-6	TEM+WIN+AIR+CLO+RAI	TEM+WIN+CLO+HUM
	7-12	TEM+WIN+CLO+HUM+RAI+THU	TEM+WIN+THU
	13-18	TEM+HUM+RAI	TEM+WIN+RAI
	19-0	TEM+WIN+PRE+AIR+CLO+HUM+RAI+THU	TEM+WIN+PRE+CLO+HUM
5	sum	TEM+WIN+AIR+CLO+HUM+THU	TEM+WIN+PRE+AIR+CLO+HUM
	1-6	TEM+WIN+AIR+CLO+HUM	TEM+WIN+AIR+HUM
	7-12	TEM+AIR+HUM+THU	TEM+CLO+RAI+THU
	13-18	TEM+WIN+AIR+HUM	TEM+CLO
	19-0	TEM+WIN+PRE+AIR+CLO+HUM	TEM+WIN+PRE+AIR+HUM
7	sum	TEM+WIN+CLO+HUM	TEM+AIR+RAI
	1-6	TEM+CLO+HUM	RAI
	7-12	TEM+HUM+THU	TEM+WIN+PRE

	13-18	TEM+WIN	TEM
	19-0	TEM+WIN+CLO+HUM+RAI	TEM+CLO
9	sum	TEM+PRE+CLO+HUM	WIN+AIR+HUM+RAI
	1-6	TEM+WIN+CLO+HUM	HUM+RAI
	7-12	TEM+PRE+HUM+RAI	HUM+RAI
	13-18	TEM+WIN+AIR	TEM+WIN+PRE+AIR
	19-0	TEM+PRE+CLO	TEM+WIN+AIR+CLO+HUM
11	sum	TEM	TEM+WIN+PRE+AIR+CLO+RAI+THU
	1-6	TEM	TEM+AIR+CLO+RAI+PRE
	7-12	TEM+WIN+CLO	WIN+CLO+HUM+RAI
	13-18	TEM+THU	WIN+PRE+HUM
	19-0	TEM+AIR+HUM	TEM+AIR
13	sum	TEM+CLO+HUM	TEM+PRE+AIR
	1-6	TEM+CLO+HUM	PRE+CLO
	7-12	TEM+CLO+HUM	TEM+PRE+AIR+RAI
	13-18	TEM+WIN+HUM+THU	TEM+WIN+PRE+AIR
	19-0	TEM+WIN+HUM+RAI+CLO	AIR+CLO
15	sum	TEM+CLO+HUM	TEM+WIN+AIR+RAI
	1-6	TEM+CLO+HUM	TEM+WIN
	7-12	TEM+PRE+AIR+RAI	WIN+AIR+HUM
	13-18	TEM+AIR+HUM+RAI	WIN+RAI
	19-0	TEM+WIN+HUM	TEM+WIN
17	sum	TEM+WIN+CLO+HUM	TEM+CLO+HUM+RAI
	1-6	TEM+CLO+HUM	TEM+WIN
	7-12	TEM+HUM+THU	WIN+AIR+HUM
	13-18	TEM+WIN+AIR	RAI+THU
	19-0	TEM+WIN+CLO+HUM+RAI	TEM+CLO+HUM+RAI
19	sum	TEM+HUM+RAI	TEM+WIN+CLO+HUM+RAI
	1-6	TEM+HUM+RAI	TEM+CLO+RAI
	7-12	TEM+WIN+PRE+CLO+THU	TEM+WIN+HUM+RAI
	13-18	TEM+RAI	TEM+RAI
	19-0	TEM+WIN+RAI	TEM+CLO+HUM
21	sum	TEM+AIR+HUM+RAI	TEM+WIN+AIR+CLO+HUM+RAI
	1-6	TEM+CLO+HUM+RAI	TEM+PRE+CLO
	7-12	TEM+CLO	PRE+AIR

	13-18	TEM+RAI+THU	TEM+AIR+HUM+RAI
	19-0	TEM+WIN+AIR+HUM+RAI	TEM+WIN+CLO+RAI

Appendix F: RMSE results for all temperature prediction models

Table 6 RMSE results for all temperature prediction models

Sensor		Without heating		With heating	
		MLR Model	ANN model	MLR Model	ANN model
1	sum	1.57	1.41	0.71	0.54
	1-6	1.31	1.17	0.61	0.49
	7-12	1.08	0.82	0.55	0.49
	13-18	1.28	0.85	0.66	0.50
	19-0	1.19	0.97	0.68	0.53
2	sum	0.86	0.82	0.63	0.56
	1-6	0.80	0.75	0.53	0.48
	7-12	0.74	0.69	0.50	0.47
	13-18	0.70	0.62	0.60	0.53
	19-0	0.73	0.68	0.56	0.47
3	sum	1.27	1.19	0.67	0.56
	1-6	1.03	0.94	0.61	0.47
	7-12	0.90	0.81	0.54	0.50
	13-18	1.11	1.07	0.60	0.53
	19-0	0.94	0.86	0.62	0.53
5	sum	1.62	1.47	0.75	0.56
	1-6	1.47	1.25	0.64	0.50
	7-12	1.19	0.87	0.42	0.42
	13-18	1.26	1.16	0.66	0.53
	19-0	1.31	1.07	0.81	0.55
15	sum	0.94	0.87	1.98	1.38
	1-6	0.77	0.66	1.65	1.21
	7-12	0.81	0.73	2.21	1.58
	13-18	0.80	0.82	2.16	1.64
	19-0	0.77	0.66	1.77	1.18
16	sum	1.27	1.13	1.85	1.37
	1-6	1.04	0.97	1.47	1.36
	7-12	1.05	0.91	2.15	1.48
	13-18	1.01	0.97	2.10	1.44
	19-0	0.96	0.89	1.67	1.14

18	sum	1.22	1.14	1.76	1.28
	1-6	1.18	0.95	1.15	0.81
	7-12	0.87	0.85	1.40	1.23
	13-18	0.91	0.85	1.99	1.28
	19-0	0.88	0.81	1.78	0.99
19	sum	1.09	0.96	1.98	1.36
	1-6	1.04	0.72	1.65	1.13
	7-12	1.04	0.85	2.23	1.35
	13-18	0.95	1.09	2.13	1.37
	19-0	0.91	0.85	1.70	1.11
20	sum	1.03	0.88	2.08	1.17
	1-6	0.76	0.59	1.80	1.20
	7-12	0.72	0.73	2.06	1.28
	13-18	0.78	0.76	2.27	1.06
	19-0	0.81	0.68	1.79	1.09
21	sum	0.94	0.83	1.56	1.08
	1-6	0.75	0.62	1.27	0.95
	7-12	0.82	0.86	1.82	1.15
	13-18	0.79	0.76	1.74	1.10
	19-0	0.74	0.74	1.31	0.95
22	sum	0.77	0.71	1.46	1.19
	1-6	0.58	0.56	1.13	0.85
	7-12	0.64	0.64	1.86	1.90
	13-18	0.73	0.68	1.79	1.77
	19-0	0.64	0.56	1.11	1.07
23	sum	0.79	0.75	1.68	1.32
	1-6	0.68	0.62	1.28	0.90
	7-12	0.70	0.75	2.28	2.20
	13-18	0.76	0.70	1.99	1.87
	19-0	0.69	0.60	1.26	1.00

Appendix G: The Settings of simulation parameters

Table 7 Parameters setting of quality loss model

Pharmaceutical products	Q10=4 Maximum shelf life=4320 hours
Flowers	Maximum vase life=304.22 hours

Table 8 Parameters setting of control strategy

[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

Table 9 Parameters setting of simulation

Simulation duration	72 hours
Number of runs	10,000 times
[redacted]	[redacted]

Appendix H: Code for simulation

```
setwd("~/Desktop/")
#install.packages('brnn')
library('brnn')
raw_data <- read.csv("samplew.csv", header=T)
raw_data <- read.csv("samplec.csv", header=T)
raw_data <- read.csv("samplewx.csv", header=T)
raw_data <- read.csv("samplecx.csv", header=T)
###temperature prediction for all sensors in future 72 hours
##without heating peridod (May to August)
{#sensor1
SEN1.2 <- predict(nn1.4.2,raw_data[raw_data$X.1%in%
c('1:00:00','2:00:00','3:00:00','4:00:00','5:00:00','6:00:00'), ])
Y2 <- c(1:6,25:30,49:54)
aa <- cbind(Y2,SEN1.2)
SEN1.3 <- predict(nn1.4.3,raw_data[raw_data$X.1%in%
c('7:00:00','8:00:00','9:00:00','10:00:00','11:00:00','12:00:00'), ])
Y3 <- c(7:12,31:36,55:60)
bb <- cbind(Y3,SEN1.3)
SEN1.4 <- predict(nn1.4.4,raw_data[raw_data$X.1%in%
c('13:00:00','14:00:00','15:00:00','16:00:00','17:00:00','18:00:00'), ])
Y4 <- c(13:18,37:42,61:66)
cc <- cbind(Y4,SEN1.4)
SEN1.5 <- predict(nn1.4.4,raw_data[raw_data$X.1%in%
c('19:00:00','20:00:00','21:00:00','22:00:00','23:00:00','0:00:00'), ])
Y5 <- c(19:24,43:48,67:72)
dd <- cbind(Y5,SEN1.5)
Total1 <- rbind(aa,bb,cc,dd)

#sensor2
SEN2.2 <- predict(nn2.4.2,raw_data[raw_data$X.1%in%
c('1:00:00','2:00:00','3:00:00','4:00:00','5:00:00','6:00:00'), ])
SEN2.3 <- predict(nn2.4.3,raw_data[raw_data$X.1%in%
c('7:00:00','8:00:00','9:00:00','10:00:00','11:00:00','12:00:00'), ])
SEN2.4 <- predict(nn2.4.4,raw_data[raw_data$X.1%in%
c('13:00:00','14:00:00','15:00:00','16:00:00','17:00:00','18:00:00'), ])
```

```
SEN2.5 <- predict(nn2.4.4,raw_data[raw_data$X.1%in%
c('19:00:00','20:00:00','21:00:00','22:00:00','23:00:00','0:00:00'), ])
Total2 <- t(cbind(t(SEN2.2),t(SEN2.3),t(SEN2.4),t(SEN2.5)))

#sensor3
SEN3.2 <- predict(nn3.4.2,raw_data[raw_data$X.1%in%
c('1:00:00','2:00:00','3:00:00','4:00:00','5:00:00','6:00:00'), ])
SEN3.3 <- predict(nn3.4.3,raw_data[raw_data$X.1%in%
c('7:00:00','8:00:00','9:00:00','10:00:00','11:00:00','12:00:00'), ])
SEN3.4 <- predict(nn3.4.4,raw_data[raw_data$X.1%in%
c('13:00:00','14:00:00','15:00:00','16:00:00','17:00:00','18:00:00'), ])
SEN3.5 <- predict(nn3.4.4,raw_data[raw_data$X.1%in%
c('19:00:00','20:00:00','21:00:00','22:00:00','23:00:00','0:00:00'), ])
Total3 <- t(cbind(t(SEN3.2),t(SEN3.3),t(SEN3.4),t(SEN3.5)))

#sensor5
SEN5.2 <- predict(nn5.4.2,raw_data[raw_data$X.1%in%
c('1:00:00','2:00:00','3:00:00','4:00:00','5:00:00','6:00:00'), ])
SEN5.3 <- predict(nn5.4.3,raw_data[raw_data$X.1%in%
c('7:00:00','8:00:00','9:00:00','10:00:00','11:00:00','12:00:00'), ])
SEN5.4 <- predict(nn5.4.4,raw_data[raw_data$X.1%in%
c('13:00:00','14:00:00','15:00:00','16:00:00','17:00:00','18:00:00'), ])
SEN5.5 <- predict(nn5.4.4,raw_data[raw_data$X.1%in%
c('19:00:00','20:00:00','21:00:00','22:00:00','23:00:00','0:00:00'), ])
Total5 <- t(cbind(t(SEN5.2),t(SEN5.3),t(SEN5.4),t(SEN5.5)))

#sensor15
SEN15.2 <- predict(nn15.4.2,raw_data[raw_data$X.1%in%
c('1:00:00','2:00:00','3:00:00','4:00:00','5:00:00','6:00:00'), ])
SEN15.3 <- predict(nn15.4.3,raw_data[raw_data$X.1%in%
c('7:00:00','8:00:00','9:00:00','10:00:00','11:00:00','12:00:00'), ])
SEN15.4 <- predict(nn15.4.4,raw_data[raw_data$X.1%in%
c('13:00:00','14:00:00','15:00:00','16:00:00','17:00:00','18:00:00'), ])
SEN15.5 <- predict(nn15.4.4,raw_data[raw_data$X.1%in%
c('19:00:00','20:00:00','21:00:00','22:00:00','23:00:00','0:00:00'), ])
Total15 <- t(cbind(t(SEN15.2),t(SEN15.3),t(SEN15.4),t(SEN15.5)))

#sensor16
```

```
SEN16.2 <- predict(nn16.4.2,raw_data[raw_data$X.1%in%
c('1:00:00','2:00:00','3:00:00','4:00:00','5:00:00','6:00:00'), ])
SEN16.3 <- predict(nn16.4.3,raw_data[raw_data$X.1%in%
c('7:00:00','8:00:00','9:00:00','10:00:00','11:00:00','12:00:00'), ])
SEN16.4 <- predict(nn16.4.4,raw_data[raw_data$X.1%in%
c('13:00:00','14:00:00','15:00:00','16:00:00','17:00:00','18:00:00'), ])
SEN16.5 <- predict(nn16.4.4,raw_data[raw_data$X.1%in%
c('19:00:00','20:00:00','21:00:00','22:00:00','23:00:00','0:00:00'), ])
Total16 <- t(cbind(t(SEN16.2),t(SEN16.3),t(SEN16.4),t(SEN16.5)))

#sensor18
SEN18.2 <- predict(nn18.4.2,raw_data[raw_data$X.1%in%
c('1:00:00','2:00:00','3:00:00','4:00:00','5:00:00','6:00:00'), ])
SEN18.3 <- predict(nn18.4.3,raw_data[raw_data$X.1%in%
c('7:00:00','8:00:00','9:00:00','10:00:00','11:00:00','12:00:00'), ])
SEN18.4 <- predict(nn18.4.4,raw_data[raw_data$X.1%in%
c('13:00:00','14:00:00','15:00:00','16:00:00','17:00:00','18:00:00'), ])
SEN18.5 <- predict(nn18.4.4,raw_data[raw_data$X.1%in%
c('19:00:00','20:00:00','21:00:00','22:00:00','23:00:00','0:00:00'), ])
Total18 <- t(cbind(t(SEN18.2),t(SEN18.3),t(SEN18.4),t(SEN18.5)))

#sensor19
SEN19.2 <- predict(nn19.4.2,raw_data[raw_data$X.1%in%
c('1:00:00','2:00:00','3:00:00','4:00:00','5:00:00','6:00:00'), ])
SEN19.3 <- predict(nn19.4.3,raw_data[raw_data$X.1%in%
c('7:00:00','8:00:00','9:00:00','10:00:00','11:00:00','12:00:00'), ])
SEN19.4 <- predict(nn19.4.4,raw_data[raw_data$X.1%in%
c('13:00:00','14:00:00','15:00:00','16:00:00','17:00:00','18:00:00'), ])
SEN19.5 <- predict(nn19.4.4,raw_data[raw_data$X.1%in%
c('19:00:00','20:00:00','21:00:00','22:00:00','23:00:00','0:00:00'), ])
Total19 <- t(cbind(t(SEN19.2),t(SEN19.3),t(SEN19.4),t(SEN19.5)))

#sensor20
SEN20.2 <- predict(nn20.4.2,raw_data[raw_data$X.1%in%
c('1:00:00','2:00:00','3:00:00','4:00:00','5:00:00','6:00:00'), ])
SEN20.3 <- predict(nn20.4.3,raw_data[raw_data$X.1%in%
c('7:00:00','8:00:00','9:00:00','10:00:00','11:00:00','12:00:00'), ])
SEN20.4 <- predict(nn20.4.4,raw_data[raw_data$X.1%in%
c('13:00:00','14:00:00','15:00:00','16:00:00','17:00:00','18:00:00'), ])
```

```
SEN20.5 <- predict(nn20.4.4,raw_data[raw_data$X.1%in%
c('19:00:00','20:00:00','21:00:00','22:00:00','23:00:00','0:00:00'), ])
Total20 <- t(cbind(t(SEN20.2),t(SEN20.3),t(SEN20.4),t(SEN20.5)))

#sensor21
SEN21.2 <- predict(nn21.4.2,raw_data[raw_data$X.1%in%
c('1:00:00','2:00:00','3:00:00','4:00:00','5:00:00','6:00:00'), ])
SEN21.3 <- predict(nn21.4.3,raw_data[raw_data$X.1%in%
c('7:00:00','8:00:00','9:00:00','10:00:00','11:00:00','12:00:00'), ])
SEN21.4 <- predict(nn21.4.4,raw_data[raw_data$X.1%in%
c('13:00:00','14:00:00','15:00:00','16:00:00','17:00:00','18:00:00'), ])
SEN21.5 <- predict(nn21.4.4,raw_data[raw_data$X.1%in%
c('19:00:00','20:00:00','21:00:00','22:00:00','23:00:00','0:00:00'), ])
Total21 <- t(cbind(t(SEN21.2),t(SEN21.3),t(SEN21.4),t(SEN21.5)))

#sensor22
SEN22.2 <- predict(nn22.4.2,raw_data[raw_data$X.1%in%
c('1:00:00','2:00:00','3:00:00','4:00:00','5:00:00','6:00:00'), ])
SEN22.3 <- predict(nn22.4.3,raw_data[raw_data$X.1%in%
c('7:00:00','8:00:00','9:00:00','10:00:00','11:00:00','12:00:00'), ])
SEN22.4 <- predict(nn22.4.4,raw_data[raw_data$X.1%in%
c('13:00:00','14:00:00','15:00:00','16:00:00','17:00:00','18:00:00'), ])
SEN22.5 <- predict(nn22.4.4,raw_data[raw_data$X.1%in%
c('19:00:00','20:00:00','21:00:00','22:00:00','23:00:00','0:00:00'), ])
Total22 <- t(cbind(t(SEN22.2),t(SEN22.3),t(SEN22.4),t(SEN22.5)))

#sensor23
SEN23.2 <- predict(nn23.4.2,raw_data[raw_data$X.1%in%
c('1:00:00','2:00:00','3:00:00','4:00:00','5:00:00','6:00:00'), ])
SEN23.3 <- predict(nn23.4.3,raw_data[raw_data$X.1%in%
c('7:00:00','8:00:00','9:00:00','10:00:00','11:00:00','12:00:00'), ])
SEN23.4 <- predict(nn23.4.4,raw_data[raw_data$X.1%in%
c('13:00:00','14:00:00','15:00:00','16:00:00','17:00:00','18:00:00'), ])
SEN23.5 <- predict(nn23.4.4,raw_data[raw_data$X.1%in%
c('19:00:00','20:00:00','21:00:00','22:00:00','23:00:00','0:00:00'), ])
Total23 <- t(cbind(t(SEN23.2),t(SEN23.3),t(SEN23.4),t(SEN23.5)))
}
```

```
FinalT <-
cbind(Total1,Total2,Total3,Total5,Total15,Total16,Total18,Total19,Total20,Total21,To
tal22,Total23)
FinalT <- FinalT[order(FinalT[,1]),]
colnames(FinalT) <-
c("TIME","SEN1","SEN2","SEN3","SEN5","SEN15","SEN16","SEN18","SEN19","SEN20","
SEN21","SEN22","SEN23")

#with heating period (September to April)
{
  average=cbind(FinalT[,1],rowMeans(FinalT[,2:13]))
#th: the critical temperature at which the heater is turned on
th <- 16
for (i in Y2){
  if (average[i,2]<th){
    FinalT[i,2]=predict(nw1.8.2,raw_data[i,3:10])
    FinalT[i,3]=predict(nw2.8.2,raw_data[i,3:10])
    FinalT[i,4]=predict(nw3.8.2,raw_data[i,3:10])
    FinalT[i,5]=predict(nw5.8.2,raw_data[i,3:10])
    FinalT[i,6]=predict(nw15.8.2,raw_data[i,3:10])
    FinalT[i,7]=predict(nw16.8.2,raw_data[i,3:10])
    FinalT[i,8]=predict(nw18.8.2,raw_data[i,3:10])
    FinalT[i,9]=predict(nw19.8.2,raw_data[i,3:10])
    FinalT[i,10]=predict(nw20.8.2,raw_data[i,3:10])
    FinalT[i,11]=predict(nw21.8.2,raw_data[i,3:10])
    FinalT[i,12]=predict(nw22.8.2,raw_data[i,3:10])
    FinalT[i,13]=predict(nw23.8.2,raw_data[i,3:10])
    }
}

for (i in Y3){
  if (average[i,2]<th){
    FinalT[i,2]=predict(nw1.8.3,raw_data[i,3:10])
    FinalT[i,3]=predict(nw2.8.3,raw_data[i,3:10])
    FinalT[i,4]=predict(nw3.8.3,raw_data[i,3:10])
    FinalT[i,5]=predict(nw5.8.3,raw_data[i,3:10])
    FinalT[i,6]=predict(nw15.8.3,raw_data[i,3:10])
    FinalT[i,7]=predict(nw16.8.3,raw_data[i,3:10])
    FinalT[i,8]=predict(nw18.8.3,raw_data[i,3:10])
```

```
        FinalT[i,9]=predict(nw19.8.3,raw_data[i,3:10])
        FinalT[i,10]=predict(nw20.8.3,raw_data[i,3:10])
        FinalT[i,11]=predict(nw21.8.3,raw_data[i,3:10])
        FinalT[i,12]=predict(nw22.8.2,raw_data[i,3:10])
        FinalT[i,13]=predict(nw23.8.2,raw_data[i,3:10])
    }
}

for (i in Y4){
    if (average[i,2]<th){
        FinalT[i,2]=predict(nw1.8.4,raw_data[i,3:10])
        FinalT[i,3]=predict(nw2.8.4,raw_data[i,3:10])
        FinalT[i,4]=predict(nw3.8.4,raw_data[i,3:10])
        FinalT[i,5]=predict(nw5.8.4,raw_data[i,3:10])
        FinalT[i,6]=predict(nw15.8.4,raw_data[i,3:10])
        FinalT[i,7]=predict(nw16.8.4,raw_data[i,3:10])
        FinalT[i,8]=predict(nw18.8.4,raw_data[i,3:10])
        FinalT[i,9]=predict(nw19.8.4,raw_data[i,3:10])
        FinalT[i,10]=predict(nw20.8.4,raw_data[i,3:10])
        FinalT[i,11]=predict(nw21.8.4,raw_data[i,3:10])
        FinalT[i,12]=predict(nw22.8.4,raw_data[i,3:10])
        FinalT[i,13]=predict(nw23.8.4,raw_data[i,3:10])
    }
}

for (i in Y5){
    if (average[i,2]<th){
        FinalT[i,2]=predict(nw1.8.5,raw_data[i,3:10])
        FinalT[i,3]=predict(nw2.8.5,raw_data[i,3:10])
        FinalT[i,4]=predict(nw3.8.5,raw_data[i,3:10])
        FinalT[i,5]=predict(nw5.8.5,raw_data[i,3:10])
        FinalT[i,6]=predict(nw15.8.5,raw_data[i,3:10])
        FinalT[i,7]=predict(nw16.8.5,raw_data[i,3:10])
        FinalT[i,8]=predict(nw18.8.5,raw_data[i,3:10])
        FinalT[i,9]=predict(nw19.8.5,raw_data[i,3:10])
        FinalT[i,10]=predict(nw20.8.5,raw_data[i,3:10])
        FinalT[i,11]=predict(nw21.8.5,raw_data[i,3:10])
        FinalT[i,12]=predict(nw22.8.5,raw_data[i,3:10])
        FinalT[i,13]=predict(nw23.8.5,raw_data[i,3:10])
```

```
  }
}
}

###Product information
##Pharma
pharma <- function(t1,t2,maxtem,value){
  temd<-FinalT[,2:13]-maxtem
  for (i in 1: nrow(temd)){
    for (j in 2:ncol(temd)){
      if (temd[i,j] < 0){
        temd[i,j]=0}
    }
  }
  #Q10=2:optimistic evaluation; Q10=3:stable evaluation; Q10=4:pessimistic
evaluation.
  Q10=4
  shelflife=4320
  shelflife1=shelflife/(Q10^(temd/10))
  shelflifeloss=shelflife-shelflife1
  houlysloss=shelflifeloss/shelflife
  totalloss=colSums(houlysloss[t1:t2,])
  costfloss=(value/(4320^2))*(totalloss^2)
  return(costfloss)
}

##Floral
floral <- function(t1,t2,maxtem,value){
  temd<-FinalT[,2:13]-maxtem
  for (i in 1: nrow(temd)){
    for (j in 2:ncol(temd)){
      if (temd[i,j] < 0){
        temd[i,j]=0
      }
      if(temd[i,j] >15){
        temd[i,j]=15
      }
    }
  }
```

```
  if (temd>15){
    vaselife=0}else{
      vaselife=370.13-65.91*exp(0.1092*temd)
    }
  vaselifeloss=304.22-vaselife
  hourlyfloss=vaselifeloss/vaselife
  totalloss=colSums(hourlyfloss[t1:t2,])
  costfloss=(value/(304.22^2))*(totalloss^2)
  return(costfloss)
}

##location choice based on time in
shipment<-read.csv("shipment.csv", header=T)
finallosst<-c()
rrab<-matrix(nrow=100,ncol=10000)
for (k in 1:10000){
  set.seed(k)
  range=1:70
  pro <- c("Flora","Pharma")
  opp <- c("COL","CRT","PIL")
  opf<-c("COL","PIL")
  for (p in 1:100){
    shipment[p,1]=sample(pro,1,replace = TRUE,prob = c(0.576,0.424))
    shipment[p,2]=sample(range,1,replace = TRUE)
    to<-as.numeric(shipment[p,2])+1
    too<-as.numeric(shipment[p,2])+9
    if (shipment[p,1] == "Pharma"){
      shipment[p,7]=sample(opp,1,replace = TRUE)
      if (shipment[p,7] == "COL"){
        shipment[p,4]=8
        if(shipment[p,2]<63){
          shipment[p,3]=sample(to:too,1,replace = TRUE)}
        else{shipment[p,3]=sample(to:72,1,replace = TRUE)}
      }
      if(shipment[p,7] == "CRT"){
        shipment[p,4]=25
        if(shipment[p,2]<63){
          shipment[p,3]=sample(to:too,1,replace = TRUE)}
        else{shipment[p,3]=sample(to:72,1,replace = TRUE)}
```

```
      }
      if(shipment[p,7] == "PIL"){
        shipment[p,4]=25
        shipment[p,3]=sample(to:72,1,replace = TRUE)}
    }
    if (shipment[p,1] == "Flora"){
      shipment[p,7]=sample(opf,1,replace = TRUE)
      if (shipment[p,7] == "COL"){
        shipment[p,4]=8
        if(shipment[p,2]<63){
          shipment[p,3]=sample(to:too,1,replace = TRUE)}
        else{shipment[p,3]=sample(to:72,1,replace = TRUE)}
      }
      if(shipment[p,7] == "PIL"){
        shipment[p,4]=25
        shipment[p,3]=sample(to:72,1,replace = TRUE)}
    }
    shipment[p,6]=sample(1:5,1,replace = TRUE)
    shipment[p,8]=shipment[p,3]-shipment[p,2]
    if(shipment[p,1] == 'Pharma'){
    shipment[p,5]=sample(40624.8759:136355.5350,1,replace = TRUE)
    }else{
      shipment[p,5]=sample(3118.9354:7188.9507,1,replace = TRUE)
    }
      }
colnames(shipment) <-
c("Type","Time.in","Time.out","Optimal","Value","Amount","Category")

shipment <- shipment[order(shipment$Time.in,-shipment$Time.out),]

a<-function(i){
  pharma(as.numeric(shipment[i,2]),as.numeric(shipment[i,3]),
         as.numeric(shipment[i,4]),as.numeric(shipment[i,5]))
}
b<-function(i){
  floral(as.numeric(shipment[i,2]),as.numeric(shipment[i,3]),
          as.numeric(shipment[i,4]),as.numeric(shipment[i,5]))
}
```

```
y<-ifelse(shipment$Type != 'Pharma',0,1)
y1<-ifelse(shipment$Type != 'Flora',0,1)
y1

c<-c()
for (i in 1:100){
  x<-a(i)*y[i]
  c<-rbind(c,x)
}

m<-c()
for (i in 1:100){
  x<-b(i)*y1[i]
  m<-rbind(m,x)
}

capacity<-c(0,120,120,120,0,0,0,0,0,0,0,0)
capacity
amount<-as.numeric(shipment$Amount)

n<-m+c
n<-t(n)
colnames <- c()
for (i in 1:100){
  name<- paste('product',i)
  colnames<-cbind(colnames,name)
}
colnames(n)<-colnames

n<-cbind(n,capacity)

ab<-c()
for(j in 1:100){
  n<-n[order(n[,j]),]
  n1<-n[order(n[,j]),]
  for (i in 1:12){
    if(n[i,ncol(n)] >= amount[j]){
      n[i,ncol(n)]<-n[i,ncol(n)]-amount[j]
      break;}
```

```
    }
    nnn<-n1-n
    x<-which(nnn==max(nnn),arr.ind=T)
    ab<-rbind(ab,x)
}
rab<-rownames(ab)
rab<-matrix(rab)
rrab[1:100,k]<-rab

locationchoice <- cbind(shipment,rab)
names(locationchoice)[7]<-c("location")
locationchoice
finallosst[k]<-0
for(i in 1:100){
    loss<-n[which(rownames(n)==rab[i]),i]
    finallosst[k]<-finallosst[k]+loss
    }
}
hist(finallosst)
finallosst<-matrix(finallosst)

choice<-t(rrab)
percentage<-matrix(nrow=12,ncol=1)
rr<-1:12
for (r in rr){
    name<-colnames(FinalT)
    name<-name[r+1]
    percentage[r,]=sum(choice==name)/(10000*100)
}
rrr<-colnames(FinalT)
rrr<-matrix(rrr)
rrr<-rrr[-1,]
percentage<-cbind(rrr,percentage)

write.table (finallosst, file ="finalloss-w1.csv",sep =",",row.names =FALSE)
write.table (percentage, file ="choice-w1.csv",sep =",",row.names =FALSE)

##location choice based on average loss
finallossa<-c()
```

```
rrab2<-matrix(nrow=100,ncol=10000)
for (k in 1:10000){
  set.seed(k)
  range=1:70
  pro <- c("Flora","Pharma")
  opp <- c("COL","CRT","PIL")
  opf<-c("COL","PIL")
  for (p in 1:100){
    shipment[p,1]=sample(pro,1,replace = TRUE,prob = c(0.576,0.424))
    shipment[p,2]=sample(range,1,replace = TRUE)
    to<-as.numeric(shipment[p,2])+1
    too<-as.numeric(shipment[p,2])+9
    if (shipment[p,1] == "Pharma"){
      shipment[p,7]=sample(opp,1,replace = TRUE)
      if (shipment[p,7] == "COL"){
        shipment[p,4]=8
        if(shipment[p,2]<63){
          shipment[p,3]=sample(to:too,1,replace = TRUE)}
        else{shipment[p,3]=sample(to:72,1,replace = TRUE)}
      }
      if(shipment[p,7] == "CRT"){
        shipment[p,4]=25
        if(shipment[p,2]<63){
          shipment[p,3]=sample(to:too,1,replace = TRUE)}
        else{shipment[p,3]=sample(to:72,1,replace = TRUE)}
      }
      if(shipment[p,7] == "PIL"){
        shipment[p,4]=25
        shipment[p,3]=sample(to:72,1,replace = TRUE)}
    }
    if (shipment[p,1] == "Flora"){
      shipment[p,7]=sample(opf,1,replace = TRUE)
      if (shipment[p,7] == "COL"){
        shipment[p,4]=8
        if(shipment[p,2]<63){
          shipment[p,3]=sample(to:too,1,replace = TRUE)}
        else{shipment[p,3]=sample(to:72,1,replace = TRUE)}
      }
      if(shipment[p,7] == "PIL"){
```

```
        shipment[p,4]=25
        shipment[p,3]=sample(to:72,1,replace = TRUE)}
    }
    shipment[p,6]=sample(1:5,1,replace = TRUE)
    if(shipment[p,1] == 'Pharma'){
      shipment[p,5]=sample(40624.8759:136355.5350,1,replace = TRUE)
    }else{
      shipment[p,5]=sample(3118.9354:7188.9507,1,replace = TRUE)
    }
  }
  colnames(shipment) <-
c("Type","Time.in","Time.out","Optimal","Value","Amount","Category")

  a<-function(i){
    pharma(as.numeric(shipment[i,2]),as.numeric(shipment[i,3]),
           as.numeric(shipment[i,4]),as.numeric(shipment[i,5]))
  }
  b<-function(i){
    floral(as.numeric(shipment[i,2]),as.numeric(shipment[i,3]),
           as.numeric(shipment[i,4]),as.numeric(shipment[i,5]))
  }

  y<-ifelse(shipment$Type != 'Pharma',0,1)
  y1<-ifelse(shipment$Type != 'Flora',0,1)
  y1

  c<-c()
  for (i in 1:100){
    x<-a(i)*y[i]
    c<-rbind(c,x)
  }

  m<-c()
  for (i in 1:100){
    x<-b(i)*y1[i]
    m<-rbind(m,x)
  }

  capacity<-c(0,120,120,120,0,0,0,0,0,0,0,0)
```

```
capacity
amount<-as.numeric(shipment$Amount)

n<-m+c
n<-t(n)
colnames <- c()
for (i in 1:100){
  name<- paste('product',i)
  colnames<-cbind(colnames,name)
}
colnames(n)<-colnames

mean<-colMeans(n)
meann<-rev(order(mean))
meannn<-rev(sort(mean))
n<-n[,meann]
n<-cbind(n,capacity)

ab<-c()
for(j in 1:100){
  n<-n[order(n[,j]),]
  n1<-n[order(n[,j]),]
  for (i in 1:12){
    if(n[i,ncol(n)] >= amount[j]){
      n[i,ncol(n)]<-n[i,ncol(n)]-amount[j]
      break;}
  }
  nnn<-n1-n
  x<-which(nnn==max(nnn),arr.ind=T)
  ab<-rbind(ab,x)
}
rab<-rownames(ab)
rab<-matrix(rab)
rrab2[1:100,k]=rab
locationchoice <- cbind(meannn,rownames(ab))
locationchoice <- locationchoice[,-1]
finallossa[k]<-0
for(i in 1:100){
  loss<-n[which(rownames(n)==rab[i]),i]
```

```
        finallossa[k]<-finallossa[k]+loss
    }
}
hist(finallossa)
finallossa<-matrix(finallossa)

choice2<-t(rrab2)
percentage2<-matrix(nrow=12,ncol=1)
rr<-1:12
for (r in rr){
    name<-colnames(FinalT)
    name<-name[r+1]
    percentage2[r,]=sum(choice2==name)/(10000*100)
}

write.table (finallossa, file ="212.csv",sep =",",row.names =FALSE)
write.table (percentage2, file ="schoice-w2.csv",sep =",",row.names =FALSE)


##location choice based on std
finallosss<-c()
rrab3<-matrix(nrow=100,ncol=10000)
for (k in 1:10000){
    set.seed(k)
    range=1:70
    pro <- c("Flora","Pharma")
    opp <- c("COL","CRT","PIL")
    opf<-c("COL","PIL")
    for (p in 1:100){
        shipment[p,1]=sample(pro,1,replace = TRUE,prob = c(0.576,0.424))
        shipment[p,2]=sample(range,1,replace = TRUE)
        to<-as.numeric(shipment[p,2])+1
        too<-as.numeric(shipment[p,2])+9
        if (shipment[p,1] == "Pharma"){
            shipment[p,7]=sample(opp,1,replace = TRUE)
            if (shipment[p,7] == "COL"){
                shipment[p,4]=8
                if(shipment[p,2]<63){
                    shipment[p,3]=sample(to:too,1,replace = TRUE)}
```

```
        else{shipment[p,3]=sample(to:72,1,replace = TRUE)}
      }
      if(shipment[p,7] == "CRT"){
        shipment[p,4]=25
        if(shipment[p,2]<63){
          shipment[p,3]=sample(to:too,1,replace = TRUE)}
        else{shipment[p,3]=sample(to:72,1,replace = TRUE)}
      }
      if(shipment[p,7] == "PIL"){
        shipment[p,4]=25
        shipment[p,3]=sample(to:72,1,replace = TRUE)}
    }
    if (shipment[p,1] == "Flora"){
      shipment[p,7]=sample(opf,1,replace = TRUE)
      if (shipment[p,7] == "COL"){
        shipment[p,4]=8
        if(shipment[p,2]<63){
          shipment[p,3]=sample(to:too,1,replace = TRUE)}
        else{shipment[p,3]=sample(to:72,1,replace = TRUE)}
      }
      if(shipment[p,7] == "PIL"){
        shipment[p,4]=25
        shipment[p,3]=sample(to:72,1,replace = TRUE)}
    }
    shipment[p,6]=sample(1:5,1,replace = TRUE)
    if(shipment[p,1] == 'Pharma'){
      shipment[p,5]=sample(40624.8759:136355.5350,1,replace = TRUE)
    }else{
      shipment[p,5]=sample(3118.9354:7188.9507,1,replace = TRUE)
    }
  }
  colnames(shipment) <-
c("Type","Time.in","Time.out","Optimal","Value","Amount","Category")

  a<-function(i){
    pharma(as.numeric(shipment[i,2]),as.numeric(shipment[i,3]),
               as.numeric(shipment[i,4]),as.numeric(shipment[i,5]))
  }
  b<-function(i){
```

```
    floral(as.numeric(shipment[i,2]),as.numeric(shipment[i,3]),
                as.numeric(shipment[i,4]),as.numeric(shipment[i,5]))
}

y<-ifelse(shipment$Type != 'Pharma',0,1)
y1<-ifelse(shipment$Type != 'Flora',0,1)
y1

c<-c()
for (i in 1:100){
    x<-a(i)*y[i]
    c<-rbind(c,x)
}

m<-c()
for (i in 1:100){
    x<-b(i)*y1[i]
    m<-rbind(m,x)
}

capacity<-c(0,120,120,120,0,0,0,0,0,0,0,0)
capacity
amount<-as.numeric(shipment$Amount)

n<-m+c
n<-t(n)
colnames <- c()
for (i in 1:100){
    name<- paste('product',i)
    colnames<-cbind(colnames,name)
}
colnames(n)<-colnames

std<-c()
for (i in 1:100){
    namee<- sd(n[1:12,i])
    std<-cbind(std,namee)
}
```

```
  stdd<-rev(order(std))
  stddd<-rev(sort(std))
  n<-n[,stdd]
  n<-cbind(n,capacity)
  ab<-c()
  for(j in 1:100){
    n<-n[order(n[,j]),]
    n1<-n[order(n[,j]),]
    for (i in 1:12){
      if(n[i,ncol(n)] >= amount[j]){
        n[i,ncol(n)]<-n[i,ncol(n)]-amount[j]
        break;}
    }
    nnn<-n1-n
    x<-which(nnn==max(nnn),arr.ind=T)
    ab<-rbind(ab,x)
  }
  rab<-rownames(ab)
  rab<-matrix(rab)
  rrab3[1:100,k]=rab
  locationchoice <- cbind(stddd,rownames(ab))
  locationchoice <- locationchoice[,-1]
  finallosss[k]<-0
  for(i in 1:100){
    loss<-n[which(rownames(n)==rab[i]),i]
    finallosss[k]<-finallosss[k]+loss
  }
}
hist(finallosss)
finallosss<-matrix(finallosss)

choice3<-t(rrab3)
percentage3<-matrix(nrow=12,ncol=1)
rr<-1:12
for (r in rr){
  name<-colnames(FinalT)
  name<-name[r+1]
  percentage3[r,]=sum(choice3==name)/(10000*100)
}
```

Bibliography

Accorsi, R., Gallo, A., & Manzini, R. (2017). A climate driven decision-support model for the distribution of perishable products. *Journal of Cleaner Production*, *165*, 917–929.
Afroz, Z., Urmee, T., Shafiullah, G. M., & Higgins, G. (2018). Real-time prediction model for indoor temperature in a commercial building. *Applied Energy*, *231*, 29–53.
Anderson, G., & Scott, M. (1991). Determination of product shelf life and activation energy for five drugs of abuse. *Clinical Chemistry*, *37*(3), 398–402.
Asumadu-Sakyi, A. B., Barnett, A. G., Thai, P., Jayaratne, E. R., Miller, W., Thompson, M. H., … Morawska, L. (2019). The relationship between indoor and outdoor temperature in warm and cool seasons in houses in Brisbane, Australia. *Energy and Buildings*, *191*, 127–142.
Aung, M. M., & Chang, Y. S. (2014). Temperature management for the quality assurance of a perishable food supply chain. *Food Control*, *40*, 198–207. https://doi.org/10.1016/j.foodcont.2013.11.016
Çelikel, F. G., & Reid, M. S. (2004). Temperature and postharvest performance of rose (Rosa hybrida L. First Red) and gypsophila (Gypsophila paniculata L. Bristol Fairy) flowers. *V International Postharvest Symposium 682*, 1789–1794.
Çelikel, Fisun G., Cevallos, J. C., & Reid, M. S. (2010). Temperature, ethylene and the postharvest performance of cut snapdragons (Antirrhinum majus). *Scientia Horticulturae*, *125*(3), 429–433.
Çelikel, Fisun G., & Reid, M. S. (2002). Storage temperature affects the quality of cut flowers from the Asteraceae. *HortScience*, *37*(1), 148–150.
Cevallos, J.-C., & Reid, M. S. (2001). Effect of dry and wet storage at different temperatures on the vase life of cut flowers. *HortTechnology*, *11*(2), 199–202.
CIOBANU, I., PUI, A., BAIA, M., HUSTI, A., & CANTOR, M. (2015). THE EFFECTS OF PRESERVATION SOLUTIONS ON MORPHODECORATIVE CHARACTERISTICS OF DAHLIA VARIABILIS CUT FLOWERS. *Contributii Botanice*, *50*.
Cognizant. (2015). https://www.cognizant.com/InsightsWhitepapers/high-velocity-supply-chains-for-perishables-a-fresher-proposition-codex1325.pdf, (Accessed 13-8-2019).
Cordis. Data filtering methods. https://cordis.europa.eu/docs/projects/cnect/5/215455/080/deliverables/ROADIDEA-D3-1-Data-filtering-methods-V1-1.pdf (Accessed 27-5-2019)
Ealey, L. (1988). Quality by design. Taguchi Methods and US industry. Dearborn. *MI: American Supplier Institute*.
Eric. (2018). The Floral Supply Chain: Cold, Competitive, Consolidating. https://onlinebusiness.syr.edu/blog/floral-supply-chain-valentines-day/, (Accessed 2-4-2019)

Foresee, F. D., & Hagan, M. T. (1997). Gauss-Newton approximation to Bayesian learning. *Proceedings of International Conference on Neural Networks (ICNN'97)*, *3*, 1930–1935. IEEE.
Greg Welch. The Kalman Filter. http://www.cs.unc.edu/~welch/kalm,an/index.html (Accessed 3-6-2019)
Gustavsson, J., Cederberg, C., Sonesson, U., Van Otterdijk, R., & Meybeck, A. (2011). *Global food losses and food waste*. FAO Rome.
Hansen, B. E. (1997). Approximate asymptotic p values for structuras-change tests. *Journal of Business & Economic Statistics*, *15*(1), 60–67.
Hawkins, D. M. (1980). *Identification of outliers* (Vol. 11). Springer.
Haykin, S. (1999). Self-organizing maps. *Neural Networks-A Comprehensive Foundation, 2nd Edition, Prentice-Hall*.
Heo, J. W., Chakrabarty, D., & Paek, K. Y. (2004). Longevity and quality of cut'Master'carnation and'Red Sandra'rose flowers as affected by red light. *Plant Growth Regulation*, *42*(2), 169–174.
Hettiarachchi, M. P., & Balas, J. (2003). Postharvest quality of cut anthurium flowers (Anthurium andraeanum L.) after long-distance shipment. *VIII International Symposium on Postharvest Physiology of Ornamental Plants 669*, 329–336.
Instruments, T. (2013). *„LM135/LM235/LM335*. LM135A/LM235A/LM335A Precision Temperature Sensors.
James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). Springer.
Jaworska, M., Szulińska, G., Wilk, M., & Tautt, J. (1999). Capillary electrophoretic separation of N-acetylcysteine and its impurities as a method for quality control of pharmaceuticals. *Journal of Chromatography A*, *853*(1–2), 479–485.
Ji, M., Fang, J., Zhang, W., Liao, L., Cheng, T. C. E., & Tan, Y. (2018). Logistics scheduling to minimize the sum of total weighted inventory cost and transport cost. *Computers & Industrial Engineering*, *120*, 206–215.
Kanlayanarat, S., Thawiang, N., & Buanong, M. (2007). Effect of Ethanol on Postharvest Quality of Cut Anthurium Flowers (Anthurium andraeanum L. cv. Marshall). *International Conference on Quality Management in Supply Chains of Ornamentals 755*, 419–424.
Katsaliaki, K., Mustafee, N., & Kumar, S. (2014). A game-based approach towards facilitating decision making for perishable products: An example of blood supply chain. *Expert Systems with Applications*, *41*(9), 4043–4059.
Kausar Shafaat, A. H., Kumar, B., Hasan, R., Prabhat, P., & Yadav, V. K. (2013). Storage of Pharmaceutical Products. *World Journal of Pharmacy and Pharmaceutical Sciences*, *2*, 2499–2515.
Kim, S., Kim, S., & Kiniry, J. R. (2018). Two-phase simulation-based location-allocation optimization of biomass storage distribution. *Simulation Modelling Practice and Theory*, *86*, 155–168.

Kim, S. Y., Lee, S. M., Park, K. S., & Ryu, K. H. (2018). Prediction Model of Internal Temperature using Backpropagation Algorithm for Climate Control in Greenhouse. *Horticultural Science and Technology, to Be Submitted*.

KLM. (2019). https://www.afklcargo.com/NL/en/local/homepage/homepage.jsp, (Accessed 13-8-2019)

KNMI. https://www.knmi.nl/nederland-nu/klimatologie/uurgegevens, (Accessed 29-6- 2019).

Kumar, R., Chandrakar, R., Kumar, A., & Chandrakar, H. R. (2012). Taguchi loss function as optimized model for supplier selection and evaluation. *International Journal of Advanced Engineering Technology*, *3*(1), 268–270.

London, E., & East, S. (2001). *STORAGE, STABILITY AND IN-USE SHELF-LIFE GUIDELINES FOR NON-STERILE MEDICINES*.

Lorite, G. S., Selkälä, T., Sipola, T., Palenzuela, J., Jubete, E., Viñuales, A., … Uusitalo, S. (2017). Novel, smart and RFID assisted critical temperature indicator for supply chain monitoring. *Journal of Food Engineering*, *193*, 20–28.

MacKay, D. J. (1992). Bayesian interpolation. *Neural Computation*, *4*(3), 415–447.

Marr, B. (2012). *Key Performance Indicators (KPI): The 75 measures every manager needs to know*. Pearson UK.

Mc Carthy, U., Uysal, I., Badia-Melis, R., Mercier, S., O'Donnell, C., & Ktenioudaki, A. (2018). Global food security–Issues, challenges and technological solutions. *Trends in Food Science & Technology*, *77*, 11–20.

McMurray, A., Pearson, T., & Casarim, F. (2017). *Guidance on applying the Monte Carlo approach to uncertainty analyses in forestry and greenhouse gas accounting*. Tech. rep., Winrock International, URL https://www. winrock. org/wp-content ….

Mehta, A. J., Mehta, H. A., Manjunath, T. C., & Ardil, C. (2008). A multi-layer artificial neural network architecture design for load forecasting in power systems. *Int. J. Appl. Math. Comput. Sci*, *5*, 207–220.

Moussa, N., & Haushey, L. A. (2015). The shelf life of Vitamin C in aw/o emulsion according to the Q10 Method. *International Journal of Pharmaceutical Sciences Review and Research*, *30*, 33–39.

Musavi, M., & Bozorgi-Amiri, A. (2017). A multi-objective sustainable hub location-scheduling problem for perishable food supply chain. *Computers & Industrial Engineering*, *113*, 766–778.

Nguyen, D., & Widrow, B. (1990). Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. *1990 IJCNN International Joint Conference on Neural Networks*, 21–26. IEEE.

Nguyen, J. L., & Dockery, D. W. (2016). Daily indoor-to-outdoor temperature and humidity relationships: A sample across seasons and diverse climatic regions. *International Journal of Biometeorology*, *60*(2), 221–229.

Nguyen, J. L., Schwartz, J., & Dockery, D. W. (2014). The relationship between indoor and outdoor temperature, apparent temperature, relative humidity, and absolute humidity. *Indoor Air*, *24*(1), 103–112.

Olsson, A. (2004). Temperature controlled supply chains call for improved knowledge and shared responsibility. *16th Annual NOFOMA Conference, Linköping, Sweden*.

Oren-Shamir, M., Shaked-Sachray, L., Nissim-Levi, A., & Weiss, D. (2000). Effect of growth temperature on Aster flower development. *HortScience*, *35*(1), 28–29.

Paterson, A. H., Tanksley, S. D., & Sorrells, M. E. (1991). DNA markers in plant improvement. In *Advances in agronomy* (Vol. 46, pp. 39–90). Elsevier.

Reppo, B., Mikson, E., & Vaarak, V. (2004). Relation between the indoor and outdoor climate in uninsulated cowsheds. *J. Agric. Sci*, *17*(2), 120–125.

Saeki, K., Obayashi, K., Iwamoto, J., Tone, N., Okamoto, N., Tomioka, K., & Kurumatani, N. (2014). The relationship between indoor, outdoor and ambient temperatures and morning BP surges from inter-seasonally repeated measurements. *Journal of Human Hypertension*, *28*(8), 482.

Sales, M. (2016). *Aviation logistics: The dynamic partnership of air freight and supply chain*. Kogan Page Publishers.

Shaheen, S. M. (2005). Accelerated stability study of metronidazole infusion 100 ml. *TAJ: Journal of Teachers Association*, *18*(2), 118–121.

Sharma, N. K., Cudney, E. A., Ragsdell, K. M., & Paryani, K. (2007). *Quality loss function-A common methodology for three cases*.

Smith, B. A., McClendon, R. W., & Hoogenboom, G. (2006). Improving air temperature prediction with artificial neural networks. *International Journal of Computational Intelligence*, *3*(3), 179–186.

Straughn, M. N., & Church, C. G. (1909). *The Influence of Environment on the Composition of Sweet Corn, 1905-1908*. US Government Printing Office.

Taylor, P. W., & Keenan, M. H. (2006). Pharmaceutical quality of generic isotretinoin products, compared with Roaccutane. *Current Medical Research and Opinion*, *22*(3), 603–615.

Tong, C., & Lock, A. (2015). A Computational Procedure for Mean Kinetic Temperature Using Unequally Spaced Data. *Proceedings of the Biopharmaceutical Section*, 2065–2070.

Tromp, S.-O., van der Sman, R. G., Vollebregt, H. M., & Woltering, E. J. (2012). On the prediction of the remaining vase life of cut roses. *Postharvest Biology and Technology*, *70*, 42–50.

van Doorn, W. G., & Tijskens, L. M. M. (1991). FLORES: A model on the keeping quality of cut flowers. *Agricultural Systems*, *35*(2), 111–127.

Van Meeteren, U. (2008). Causes of quality loss of cut flowers-a critical analysis of postharvest treatments. *IX International Symposium on Postharvest Quality of Ornamental Plants 847*, 27–36.
Van't Hoff, J. H., & Lehfeldt, R. A. (1899). *Lectures on theoretical and physical chemistry*.
Vega, H. (2008). *The Transportation Costs of Fresh Flowers: A Comparison between Ecuador and Major Exporting Countries*. Inter-American Development Bank.
Vršek, I., Karlović, K., & Židovec, V. (2002). Effect of day length, growth regulators and fertilization on growth and development of Michaelmas Daisy (Aster novi-belgii L.). *Agriculturae Conspectus Scientificus*, *67*(4), 207–213.
Wang, J., Wang, H., He, J., Li, L., Shen, M., Tan, X., … Zheng, L. (2015). Wireless sensor network for real-time perishable food supply chain management. *Computers and Electronics in Agriculture*, *110*, 196–207.
Wang, J.-X., Xia, T., Zhang, J.-M., & Zhou, S.-L. (2009). Isolation and characterization of fourteen microsatellites from a tree peony (Paeonia suffruticosa). *Conservation Genetics*, *10*(4), 1029–1031.
Xing, L., Li, L., Gong, J., Ren, C., Liu, J., & Chen, H. (2018). Daily soil temperatures predictions for various climates in United States using data-driven model. *Energy*, *160*, 430–440.
Yang, H., & Chen, H. (2015). Biomass gasification for synthetic liquid fuel production. In *Gasification for Synthetic Fuel Production* (pp. 241–275). Elsevier.
Yoo, Y.-K., & Kim, W.-S. (2003). Storage solution and temperature affect the vase life and quality of cut gerbera flowers. *Korean Journal of Horticultural Science and Technology*, *21*(4), 386–392.
Zheng, M., & Guo, Y. (2018). Cerium improves the vase life of Dianthus caryophyllus cut flower by regulating the ascorbate and glutathione metabolism. *Scientia Horticulturae*, *240*, 492–495.