Optimization of Strawberry Supply Chain from the Perspective of Producers

Author: Zhongbo Yao

Delft University of Technology

Optimization of Strawberry Supply Chain from the Perspective of Producers

by

Author: Zhongbo Yao

Student Name	Student Number
Zhongbo Yao	5473543

Supervisor I:	Neil Yorke-Smith
Supervisor II:	Junhan Wen
Project Duration:	12, 2022 - 08, 2023
Faculty:	Computer Engineering, TU Delft

Cover: Strawberries Image by iStock (modified)
Style: TU Delft Report Style, with modifications by Daan Zwaneveld

Abstract

This thesis aims at maximizing the profit of a strawberry producer while satisfying the retailer's demand and meeting other constraints. The amount of strawberries to be delivered to the retailer signed in the contract is the main decision variable to be optimized in the problem. Furthermore, the transportation scheduling is also optimized to help the producer reduce cost.

Firstly, three ensembles of LSTMs are applied to predict the yield, price and demand of strawberries in a target month. The structure and hyper-parameters of the LSTMs are chosen carefully using grid-search. Secondly, MILP model is applied for optimizing the contract demand and the transportation scheduling. Finally, robust optimization is applied to the model to mitigate the shortage cost that may be introduced by the uncertain yield amount. The feasibility of the problem and the sensitivity of the parameters are also discussed in the thesis.

The result shows that the contract optimized by the MILP model earns 7.96% more profit compared with the baseline contract. And the robust optimization increased the profit by 1.69% by saving the shortage cost (when Γ is set to be 0.1). Therefore, it can be concluded that the MILP model helps the producer to earn more profit and the robust optimization successfully increases the final profit by considering the bad cases when the real yield is less than the prediction.

The contribution of this thesis is the combination of the deep learning neural network with the traditional MILP optimization model, together with the improvement using robust optimization. A case study with more complex conditions is expected as the major further work of the thesis, in order to see whether the optimized contract can over-perform the strategies commonly adopted by companies in the real scenarios. The LSTMs in the thesis are also expected to be trained on more accurate datasets, which are difficult to obtain without access permission.

Contents

Constants & Variables

Constants

Symbol	Definition	Unit
M	A very large constant	-
ϵ	A very small constant	-
Z_h	Cost of keeping strawberries in the inventory	$[\$/Kg]$
Z_t	Transportation cost of strawberries	$[\$/Kg]$
Z_{truck}	Cost of each truck	[$]
$H_{i,j}$	The amount of strawberries harvested for batch i on day j	$[Kg]$
$D_w^{prediction}$	The predicted demand on week w	$[Kg]$
$P_w^{prediction}$	The predicted price on week w	$[\$/Kg]$
$D_j^{retailer}$	The minimum demand of the retailer on day j	$[Kg]$
T_t	The time consumption of the transportation	$[Days]$
w^{more}	Weight of the retailer's reaction if $ind_w^{Difference} = 1$	$[\$/Kg^2]$
w^{less}	Weight of the retailer's reaction if $ind_w^{Difference} = 0$	$[\$/Kg^2]$
$w_{discount}$	Discount in price	-
$limit$	The maximum amount of strawberries allowed to be sold in the discounted price	$[Kg]$
Cap_{inv}	The capacity of the producer's inventory	$[Kg]$
Cap_{truck}	The capacity of each truck	$[Kg]$
σ_i	The survival rate of strawberries in the producer's inventory	-
σ_t	The survival rate of strawberries during the transportation	-
σ_r	The survival rate of strawberries in the retailer's inventory	-

Decision Variables

Decision Variables

Symbol	Definition	Unit
$Profit$	The final profit, which is the objective	[$]
$Income_{contract}$	Income by executing the contract	[$]
$Income_{discount}$	Income by selling the strawberries in the discounted price	[$]
$C_{holding}$	Total cost of keeping the strawberries in the inventory	[$]
$C_{transportation}$	Total cost of transportation	[$]
$C_{shortage}$	Total cost of failing to deliver enough strawberries	[$]
$C_{dispose}$	Total cost of disposing strawberries	[$]
$C_{deterioration}$	Total cost because of deterioration	[$]
$D_w^{difference}$	The difference between $D_w^{contract}$ and $D_w^{prediction}$ on week w	$[Kg]$
$D_w^{contract}$	Demand of strawberries on week w stated in the contract	$[Kg]$
$P_w^{contract}$	Price of strawberries on week w in the contract	$[\$/Kg]$
w_w^{price}	Weight of the retailer's reaction towards the $D_w^{difference}$	$[\$/Kg^2]$
$left_{i,j}$	The amount of strawberries left in the producer's inventory before dispose from batch i on day j	$[Kg]$
$I_{i,j}$	The amount of strawberries left in the producer's inventory after dispose from batch i on day j	$[Kg]$
$T_{i,j}$	The amount of strawberries transported from batch i on day j	$[Kg]$
$F_{i,j}$	The amount of strawberries sold in the discounted price from batch i on day j	$[Kg]$
$dispose_{i,j}$	The amount of strawberries to be disposed from batch i on day j	$[Kg]$
$I_j^{retailer}$	The amount of strawberries in the retailer's inventory on day j	$[Kg]$
$left_j^{tot}$	The total amount of strawberries left in the producer's inventory before dispose on day j	$[Kg]$
T_j^{tot}	The total amount of strawberries transported on day j	$[Kg]$
F_j^{tot}	The total amount of strawberries sold in discounted price on day j	$[Kg]$
$dispose_j^{tot}$	The total amount of strawberries disposed on day j	
$ind_w^{Difference}$	Indicating whether $D_w^{difference}$ is larger than 0	-
ind_j^{cap}	Indicating whether the amount of strawberries exceeds the capacity on day j	-
$ind_{i,j}^{transportation}$	Indicating whether there are strawberries transported from batch i on day j	-

Symbol	Definition	Unit
$ind_{i,j}^{inventory}$	Indicating whether there are strawberries from batch i left before dispose in the producer's inventory on day j	-
$ind_{i,j}^{dispose}$	Indicating whether there are strawberries disposed from batch i on day j	-
$ind_{i,j}^{left}$	Indicating whether there are strawberries from batch i left after dispose in the producer's inventory on day j	-
$Truck_j$	The number of trucks scheduled on day j	-

1
Introduction

This chapter introduces the background and the meaning of the problem, followed by the introduction of the outline and the contributions of this thesis.

1.1. Background

Perishable products consumption is essential for humanity to keep healthy, like fruits and vegetables consumptions. Supply chains are playing an important role in distributing the products from factories or farms to customers or retailers. However, there are numerous decisions involved in the supply chain to determine a complete strategy to follow, making the supply chain fragile under the negative influence from like human errors, bad weather, transportation disruption, etc. These factors unavoidably introduce cost to each stage of the supply chain, hurting the interest of the producers, the customers and even the environment. Furthermore, the perishability of such products results in further loss of values.

The cost in such supply chains has already caught the researchers' attention. It is studied that 40-50% of all root crops, fruits, and vegetables are wasted [1]. Specifically, at the farm level, waste can be as high as 20% whereas post-harvest waste is estimated to be 3%; at the retail and consumer levels, waste increases to 12% and 28%, respectively [2, 3, 4]. In post-harvest, waste occurs during sorting, handling, storage, and distribution [5].

Therefore, it is profitable to find the best trade-offs and the best strategy in order to mitigate such costs. The decisions like fruits harbour scheduling, transportation scheduling and the determination of the price are usually considered as the optimization targets because there are always trade-offs among them. When trying to find the best solution, many constraints should be considered about, ranging from the maximum yield value of the products, the capacity of inventory, to the satisfaction of customers' demands.

Considering the requirements and needs for solving the problem, combinatorial optimization is often considered as an appropriate approach as it is able to optimize a objective function while satisfying the constraints. Various combinatorial optimization models are attempted and the mixed integer linear programming model (MILP) is the most frequently used one - the objective is usually to maximize the income and to minimize the cost introduced by transportation, keeping the products in the inventory, etc. In this way, the profit of all the parties involved in the supply chain can be optimised. Based on the MILP models, stochastic programming and robust optimization have also been explored to make the results robust under uncertainties, like the fluctuation in yield or the market price, or the weather.

When making decisions, there should be sufficient relevant data providing information as guidance, which is likely unknown at the time. For example, the local market price for the next month guides the manager to decide the price for their products. But the price information cannot be obtained without the help of prediction algorithms. Therefore, the accurate forecast of such data is significant. There are multiple algorithms can be used to forecast time-series data, like Long Short-term Memory (LSTM), AutoRegressive Integrated Moving Average (ARIMA) and Temporal Convolutional Networks (TCN), etc.

This thesis takes the strawberry supply chain as an example of the perishable products supply chains. The idea is to combine the deep learning model for data prediction with the MILP model for optimizing the objective, and then apply robust optimization making the result more robust. The LSTMs are used to predict the strawberry yield, market demand and market price information. Then, the information is passed to the MILP model for supporting the decision about the contract value and the transportation scheduling to maximize the producer's profit. The robust optimization also uses the predicted information to acknowledge the prediction distribution.

1.2. Related Work

Lots of efforts have been made to optimize problems about supply chains. There are many different models designed to optimize the problem, with different concerns, scenarios and objectives. The functional areas are considered by most of the models, others also focus on the decision-scenarios, the environmental effects and fruit characteristics [6].

Caixeta-Filho et al. [7] proposed an MIP model for the orange supply chain in Brazil. The model proposed is simple, including one farm, one inventory and one retailer. It mainly considers about the optimization of the harvest time and transportation cost. The model introduces a method of using total soluble solid and total acid as the measurement of the orange ripeness. It proves that the best harvesting time is earlier than the time when the yield peaks, but it ignores the effect of perishability.

Rocco et al. [8] proposed a conceptual model for the tomato supply chain in Brazil, which includes multiple farms, multiple processing plants and one retailer. The locations of the farms are carefully chosen by the plants according to their distance and production. And each farm can decide which tomato variations to plant for maximising the profit. The model is programmed in MLP and mainly considers about the cost in the processing plant. The model helps to optimize the decisions about planting, transportation and processing of Brazil tomato in the supply chain, ignoring the decay factor.

Ahumada et al. [9] came up with an LP model for the perishable product supply chain in Mexico. The model includes multiple farms, packaging centres, warehouses, distributors and customers, and all the three kinds of facilities have direct delivery methods to the customers. It mainly focuses on optimizing the arrangement of labour and transportation of the system. To consider about the effect of perishability, the model adds a linear decay factor in the objective function to stimulate the deterioration.

Widodo et al. [10] introduced an MIP model for the flower supply chain. It utilises the mature curves of the flowers, which has multiple generations at a time point, to optimisze the harvest scheduling. They also consider the decay process during the transportation by introducing a non-linear decay function into the model.

Dubinin et al. [11] introduced an MIP model integrated with decision tree method for optimizing the vegetable supply chain. Grillo et al.[12] proposed a multi-objective model, which is able to maximise the profit and to minimise shelf time simultaneously. Itoh et al.[13], Kakaz et al.[14] and Bezat-Jarzebowska et al.[15] also used stochastic algorithms to optimise the final profit in the supply chains by maximising the profit and minimising the product loss.

There are also papers using robust optimisation, making the model robust under various uncertainties. Bertsimas et al. [16] proposed a robust optimisation method that can be used to MLP/LP models. Different from the approach introduced by Soyster [17], which prevents the model from violating all the constraints by considering the worst-case scenarios, [16] reformulated the model so that the model is less conservative with slightly higher risk of violation, but raising the expectation of the profit. The paper also indicated a method of specifying the tolerance for violation of the constraints.

By using the approach proposed in [16], Bertsimas et al. themselves formulated a robust optimization model to solve the supply chain management problems [18] by considering about the uncertainties from the fluctuation in the retailers' order amount. It is proved that this method over-performs dynamic programming in terms of the computing tractability and the expectation of the final results. Yehuda Bassok et al., Yue Wu and Aharon Ben-Yal et al. also used robust optimisation to solve problems in supply logistic and to decide a best contract [19, 20, 21].

Different from the work mentioned above, the thesis only focuses on optimizing the profit of the producer, which means how the retailer processes with the strawberries is not considered. The contract signed between the producer and the retailer is the key in this optimization problem. The thesis also combines deep learning model with the MILP model, which has not been explored yet.

1.3. Thesis Outline and Contributions

After the introduction in the first chapter, the second chapter describes the scenario and the assumptions of the problem. The third chapter introduces the background knowledge of LSTM, and how the LSTM functions in this model. That chapter is mainly about the data processing and the structure of the LSTM. Then the MILP model's structure and mathematical formulations are covered in the fourth chapter. The fifth chapter illustrates the results of this model, followed by the further work and final conclusion in the sixth and seventh chapters respectively.

The contribution of this thesis is the combination of the deep learning neural network with the traditional MILP optimization model, together with the improvement using robust optimization. The influence of the loss and the confidence of the LSTMs on the optimized results is studied. Before solving the problem, the feasibility of the problem is proven in advance. Finally, the sensitivity analysis of the parameters is conducted to study the sensitivity of the model to the environment. By studying the results, reasons why the model makes a certain decision are also better understood.

2 Problem Description

This chapter introduces the structure of the supply chain and the assumptions of the problem.

2.1. Structure of the Supply Chain

The thesis picks a strawberry supply chain as the target to optimize, as there are supportive strawberry data and information about trading methods extracted from interviews and researches done by Junhan Wen with strawberry companies and markets in the Netherlands.

It is assumed that there are only one producer and one retailer in this supply chain, simplifying unnecessary details in the problem. The producer harvests the strawberries from his farm regularly and sells them to the retailer by signing the contract one month in advance. For example, on 1st June a contract for July is signed. The contract specifies the weekly amount of strawberries to be delivered in July, with the corresponding price and the transportation scheduling. When signing the contract, information about strawberry yield, market demand and market price is predicted by LSTMs. Figure 2.1 illustrates how the contracts are signed.

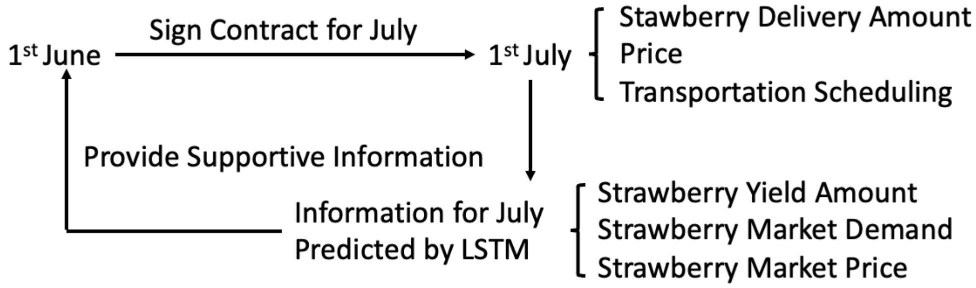

Figure 2.1: With information about strawberries provided by the LSTMs, a contract is signed one month earlier, including the amount of strawberries to be delivered, the corresponding price and the transportation scheduling.

After harvesting the strawberries, the producer stores them in the inventory immediately. The time and money costs for transporting the strawberries from the farm to the producer's inventory are ignored. Then the strawberries will be delivered to the retailer's inventory following the signed contract. Finally, the retailer sells the strawberries to meet the market demand. It is assumed that both producer and retailer have their own inventories, and strawberries stored in the inventories perish at a specific speed. The whole structure is shown in Figure 2.2.

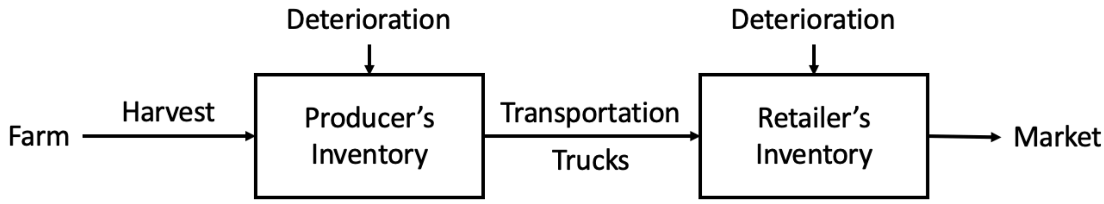

Figure 2.2: There are only a producer and a retailer considered in the structure. Both of them have inventories, in which the strawberries are deteriorating. Only the transportation cost between the producer and the retailer is considered.

2.2. Trade-offs

When signing a contract, it is assumed that the retailer has various reactions to the amount signed in the contract, affecting the final profit. Firstly, it is naturally recognized that if the strawberries offered by the producer are too many, the retailer would only receive them in a low price, because retailers do not want to buy excessive strawberries. Similarly, if the strawberries the producer decides to deliver are less than usual because of poor yield or other reasons, a slightly higher price is acceptable to the retailer.

Sometimes, the yield of strawberries is extremely good and there are still strawberries left in the inventory after full-filling the contract, causing waste. In this case, the excessive part can be sold to the retailer in a discounted price. But a limitation of the amount of strawberries sold in this way is applied, measured by the proportion of the amount agreed in the contract. Finally, the retailer gives feedback to the producer according to the comparison between the signed contract and the real data in that month. The feedback affects the prices slightly.

There are multiple trade-offs can be observed. For example, the final profit is mainly decided by the amount of strawberries signed in the contract and the resulted price, which changes reversely according to the signed amount. The producer should also consider about whether it is worthy to sign a contract in the risk of the shortage of the supply, which introduces shortage cost. All these and other optimization questions under this problem making the problem interesting to be studied.

2.3. Assumptions

The model is built based on the following assumptions:

- There is one producer and one retailer in the whole supply chain.
- Harvesting scheduling is fixed to harvest all the yield once a week on Sunday.
- The harvested strawberries are delivered immediately to the inventory of the producer without time cost and money cost.
- Strawberries begin to deteriorate immediately after being harvested.
- Each month equally has four weeks.
- The model has been used for a while rather than cold-started.
- When signing the contract, there is no backlog allowed. The amount of strawberries failed to be delivered will result in a shortage cost immediately.

Explanation for the cold-start problem: When the optimizer and the LSTMs are used for the first time, there will be a lack of information for the current month, because that the LSTMs are designed for predicting the situation a month later, as shown in Figure 2.3. When the optimizer and the LSTMs are initialized on the 'Current Date', the data about the past month is known as it has already happened, and the data about the next month (marked in orange circles) is also known because of the estimation by the LSTMs. However, there is no estimation nor knowledge about the information for the current month, which is marked by black dashed circles. Therefore, it is assumed the model is being used for several periods. Then the missing values are recorded one month ago by prediction.

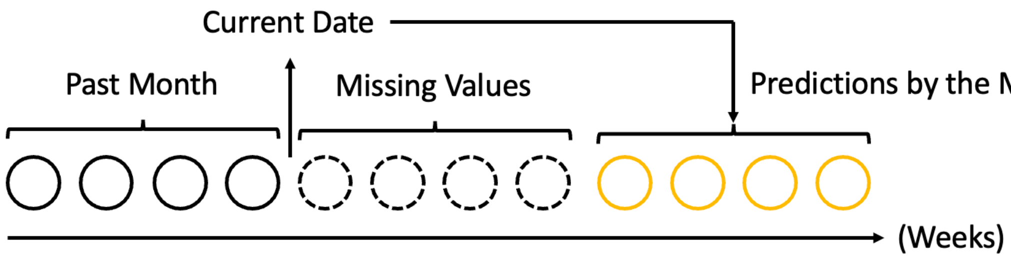

Figure 2.3: When initially started, the model can only be used to predict the data for one month later, which means the data for the current month is unknown, marked by black dashed circles.

3
Prediction

This chapter introduces how the ensembles of LSTMs predict the yield, the market demand and the market price of the strawberries to support the decisions made by the optimizer.

3.1. LSTM

The LSTM is short for long-short term memory neural network, which is able to process the entire sequence of data. The structure of an LSTM cell is shown in Figure 3.1.

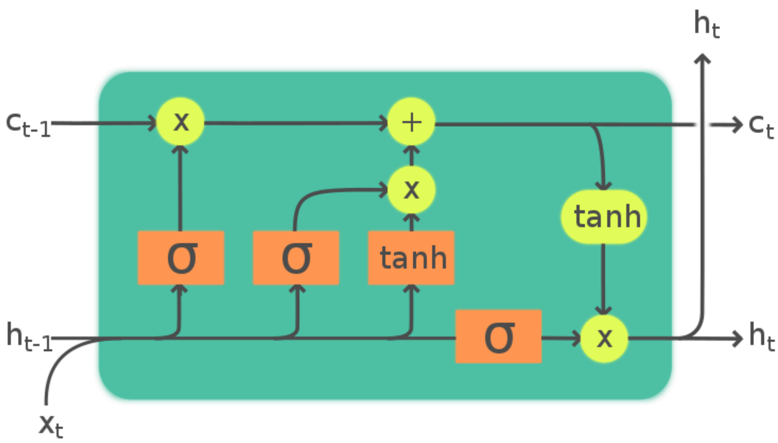

Figure 3.1: The structure of the LSTM. The various gates play important roles in it.

Importantly, an LSTM has a cell, an input gate, an output gate and a forget gate. Because of the function of both the cell and the input gate, an LSTM is able to memorize the sequence of the data and to learn the time-series pattern from it. The forget gate controls which history information to be forgotten, making sure that the memory too far away does not have a high weight on the loss calculation, preventing LSTM from gradient vanishment. Then the learnt pattern can be used to predict values that having the same distribution as the training data. Due to this characteristic, LSTM is suitable for predicting time series data, like the price in stock, precipitation, temperature, etc.

The functions of all gates and paramaters are expressed as following:

$$
\begin{aligned}
f_t &= \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \\
i_t &= \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \\
o_t &= \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \\
\hat{c}_t &= \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \\
c_t &= f_t \odot c_{t-1} + i_t \odot \hat{c}_t \\
h_t &= o_t \odot \sigma_h(c_t)
\end{aligned}
$$

Compared with other algorithms like ARIMA, SARIMA and VAR etc., the LSTM are more suitable for predicting data in short-term or mid-term range, delivering better results in higher accuracy [22, 23, 24]. Therefore, LSTM is chosen to predict the information in the project.

3.2. An Ensemble of LSTMs

In this thesis, three ensembles of LSTMs are used to study the influence of the prediction distribution and confidence over the final optimized profits. The distribution of the prediction results also provides information of the worst-case and best-case scenarios, which are exploited in the application of robust optimization.

The ensemble learning belongs to the integrated learning. In this project, there are 50 LSTMs in each ensemble to predict the wanted data. Firstly, the LSTMs are initialized randomly and then trained with bootstrapping technique. Both the random initialization and the bootstrapping are to introduce diversity to the prediction results. Afterwards, the LSTMs are validated and tested before the predicting.

The predictions of the LSTMs are plotted to visualize the distribution of the prediction results. The mean value of the predictions are considered as the prediction value of an ensemble. Figure 3.2 illustrates how does an ensemble of LSTMs function.

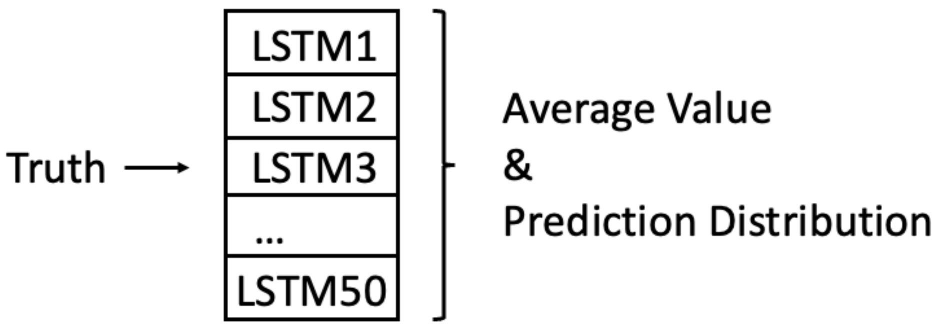

Figure 3.2: An ensemble of 50 LSTMs are trained together to get the predictions. The mean value is used as the final prediction and the prediction distribution can also be drawn, which is important in uncertainty analysis.

3.3. Data Processing

Because it is impossible to obtain monthly or weekly data of strawberry yield and market demand without the help from government or companies, and yearly data is far from being sufficient to train the model, the data for strawberry yield and market demand is replaced by other similar data. The strawberry price is accessed through [25]. The strawberry market demand is replaced by the data from [26], and strawberry yield is replaced by data from [27]. Specifically, the data for the strawberry demand since year 2022 is abandoned as there is a tremendous increasing trend observed, shown in Figure 3.3, caused by the COVID-19 probably. The LSTM cannot predict such a sudden event using the history data. The original data sets are the indices indicating the increase or decrease percent compared with a standard year, therefore they are scaled firstly before application.

Figure 3.3: There is a tremendous increasing trend observed in the data, which is probably caused by the pandemic shock and cannot be predicted. Therefore, the data after year 2022 is excluded.

Data is pre-processed before prediction. Taking the original dataset of history market price as an example, there are two columns of data in the dataset, date and price value respectively. Only the column of price is reserved as the LSTM should not directly link the price with the values of date. The nan values in the original dataset are replaced by the mean value of its nearest non-nan neighbours. Afterwards, the data is normalized using the Max-min scaler to improve the accuracy of the prediction.

After pre-process, data is scanned by the sliding-window algorithm and then is fed into LSTMs for prediction. The LSTMs are designed to predict the yield amount, market demand and market price for 1 month later, including four data points in each kind(4 weeks = 1 month). Each window is only responsible for the prediction for one week. Therefore to obtain four data points, the window are shifted four times. Considering that when making predictions, the real data for the current month is unknown and the available data for the current month is also made by LSTM predictions, there are 7 intervals between the training data and the truths, illustrated in Figure 3.4. It is the sum of the window shift times and the time interval of 4 weeks (because the data predicted is about one month later). A portion of 0.1 of all the windows are reversed for cross-validation for evaluating the performance and another 0.1 of all the windows are reserved as test data.

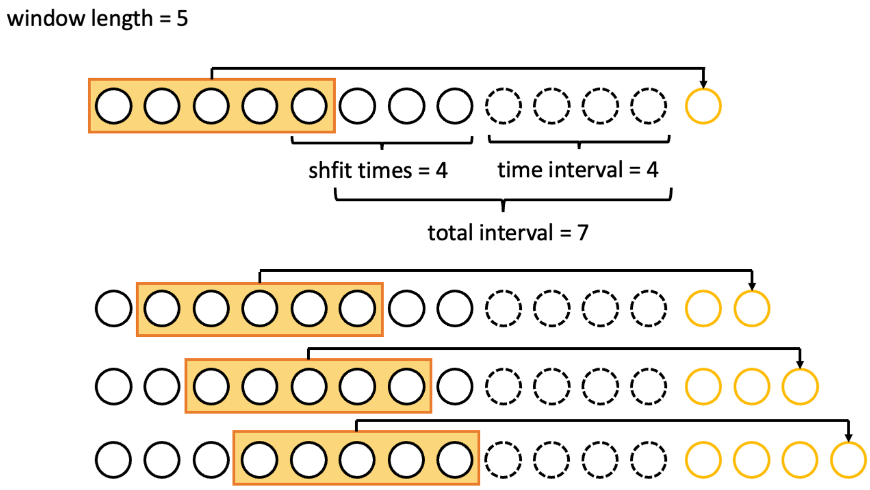

Figure 3.4: The dots represent the data points for any data, including the yield, market demand and maket price information. Before sending the data to the LSTMs, sliding window splits data into patches. Each window is only responsible for one prediction, and the window is shifted four times to obtain a complete prediction set.

3.4. Choice of Loss Function

When predicting the information using LSTMs, MSE is chosen as the loss function. The primary reason is that the prediction of price, demand and yield is a regression task, which makes it natural to adopt the MSE for measuring the distance between the truth and predictions. And the results given by the MSE can be better interpreted, exhibiting an intuitive impression about how the predictions mismatch with the truth. Futhermore, MSE emphasizes on the large errors. When faced with very noisy data, the LSTMs cannot predict completely accurately. Therefore, it is expected that at least the outliers are punished more harshly, offering an accurate range where the target values locate.

3.5. LSTM Structure

It is difficult to reasonably design the structure for LSTM to make its complexity appropriate. The model is faced with challenges that the training data is not sufficient and the patterns behind the strawberry yield and demand data are vague. A too complex model suffers from overfitting but a too simple model cannot capture the pattern. Therefore, the structure of the LSTM is carefully designed and the hyper-parameters are tuned using grid search.

The structure of the LSTM is illustrated in Figure 3.5. Following the LSTM cell, there are 3 fully connected layers. The Sigmoid activation function is used after the first and the second fully connected layers. The fully-connected layers and the activation functions are applied to increase the model's complexity as the training data is noisy and a simple LSTM layer fails to capture the pattern behind it. And at least one fully connected layer is necessary as the output of the LSTM layer should be reshaped by the fully connected layer. Early stopping with a warm-up is also implemented to all of the LSTMs to avoid ovefitting.

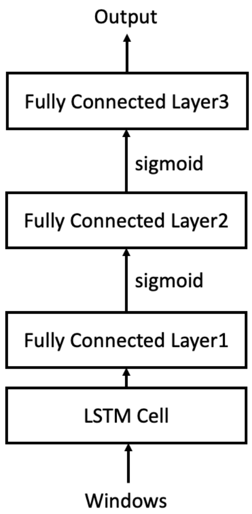

Figure 3.5: The structure of the LSTM model. Each LSTM contains a LSTM cell, 3 fully connected layers. The first and the second fully connected layers are activated by the Sigmoid function.

Figure 3.6 illustrates an example graph comparing the validation loss among various structures of the LSTM neural network for predicting market demand of strawberries. It can be seen that the configuration of 3 fully connected layers promises the best performance. Before 4 fc layers, both the complexity and accuracy increases with the number of fc layers, indicating the structure is more capable to capture the pattern. However, a tremendous increase of loss can be seen after 4 fc layers, because the data is not sufficient enough to support such a complex model.

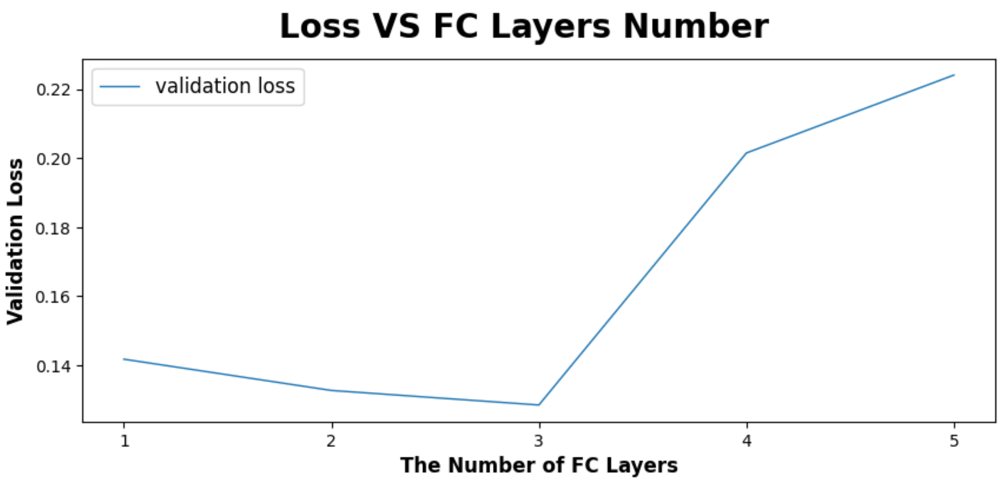

Figure 3.6: The result shows that 3 layers is the best choice for the LSTM structure. The validation loss decreases as the number of fc layers increases to 3, and then increases dramatically after the point.

All the hyper-parameters are also searched using grid search similarly. Figure 3.7 gives an example of the result of grid search for the hyper-parameter `Window Length` as an example. It can be seen that the best performance can be achieved by setting the `Window Length` as 17.

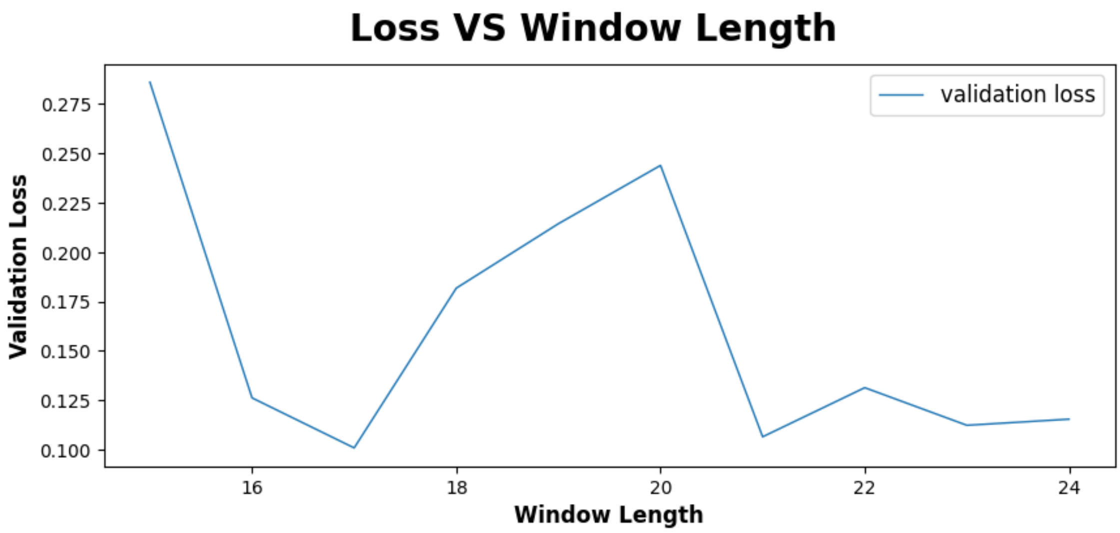

Figure 3.7: The comparison of loss among various configuration of the window length of the LSTMs for predicting demand. It can be seen that the best performance can be obtained by setting the window length as 17.

After the comprehensive grid search, the hyper-parameters for each LSTM are tuned as Table 3.1.

Table 3.1: Tuned Hyper-parameters for the Three Kinds of LSTM

	Window Length	Hidden Size	LSTM Layers Num	Learning Rate	Early Stopping Patience
Demand Prediction	17	24	1	0.00005	800
Price Prediction	12	10	2	0.0001	500
Yield Prediction	15	16	2	0.0001	1000

4 Mixed Integer Optimization Model

This chapter introduces the detailed structure and the formulation of the MILP model, which aims at finding the best contract demand and transportation scheduling in order to maximize the profit of the producer while satisfying the demand from the retailers and other constraints. In the end, robust optimization is also explored to make the optimized strategy robust under the uncertain yield prediction.

4.1. Formulation of the Model

The first part of the MILP is to calculate the contract price according to the difference between the contract demand and the predicted market demand, decided by the producer and the LSTM respectively. The calculation is shown in the following formulas:

$$D_j^{difference} = D_j^{contract} - D_j^{prediction} \tag{4.1.1}$$

$$D_j^{difference} \geq 0 \rightarrow w_j^{price} = w_j^{more} \tag{4.1.2}$$

$$D_j^{difference} \leq 0 \rightarrow w_j^{price} = w_j^{less} \tag{4.1.3}$$

$$P_j^{contract} = P_j^{prediction} - \frac{(D_j^{difference})^2}{25} * w_j^{price} \tag{4.1.4}$$

$$D_j^{contract}, P_j^{contract} \geq 0 \tag{4.1.5}$$

$$\forall j = \{0, 1, 2, 3\}$$

The price in contract $P_{\cdot}^{contract}$ changes non-linearly from the prediction price $P_{\cdot}^{prediction}$ according to the difference between the amount signed in the contract $D_{\cdot}^{contract}$ and the prediction demand $D_{\cdot}^{prediction}$, simulating the retailer's reaction towards the fluctuation of the contract. The weight $w_{\cdot}^{price}$ simulates the sensitivity of the retailer to the amount signed in the contract. The retailer has different feelings according to whether $D_{\cdot}^{contract}$ is larger than the $D_{\cdot}^{prediction}$ or not. When $D_{\cdot}^{contract} > D_{\cdot}^{prediction}$, the retailer has the sensitivity $w_{\cdot}^{more}$. Otherwise, the sensitivity is $w_{\cdot}^{less}$. The value of the two parameters are set to relatively low, representing that the producer is unwilling to decrease the price for extra strawberries, and the retailer is not willing to pay more if $D_{\cdot}^{contract}$ is less than the predictions. All these variables only represent the value for a week, therefore, j can be 0 to 3, indicating the week number within the target month.

Then the state of the inventory of the producer is described. The inventory capacity C_{inv}, transportation of strawberries to the retailer T_{ij} on day j from batch i, the excessive strawberries sold in the discounted price F_{ij} and the deterioration rate σ_{inv} are the major concerns.

$$left_{i,0} \leq H_{i,0} - T_{i,0} - F_{i,0} \tag{4.1.6}$$
$$left_{i,j} \leq I_{i,j-1} * \sigma_{inv} + H_{i,j} - T_{i,j} - F_{i,j} \tag{4.1.7}$$
$$I_{i,j} = left_{i,j} - dispose_{i,j} \tag{4.1.8}$$
$$T_j^{tot} = \sum_{i=0}^{Batches} T_{i,j} \tag{4.1.9}$$
$$F_j^{tot} = \sum_{i=0}^{Batches} F_{i,j} \tag{4.1.10}$$
$$\sum_{j=0}^{6} T_{w*7+j}^{tot} \leq D_w^{contract} \tag{4.1.11}$$
$$\sum_{j=0}^{6} F_{w*7+j}^{tot} \leq D_w^{contract} * limit \tag{4.1.12}$$
$$left_{i,j}, T_{i,j}, F_{i,j}, I_{i,j}, dispose_{i,j}, T_j^{tot}, F_j^{tot} \geq 0 \tag{4.1.13}$$
$$\forall i = \{0, 1, ..., batches\}$$
$$\forall j = \{0, 1, ..., 27\}$$
$$\forall w = \{0, 1, 2, 3\}$$

The formula (4.1.6) states the initial state of the inventory, and (4.1.7) describes the inventory state afterwards. $left_{i,j}$ indicates the amount of the strawberries left in the inventory after storing the yield H, the transportation T, discounted price sales H and deterioration. After dispose, the final left strawberries are kept in the inventory I. The dispose usually happens when $left_{\cdot,\cdot}$ is more than the inventory capacity, which will be introduced later. Then, the formula (4.1.11) regulates that the total delivery of strawberries in a week $T_{\cdot}^{tot}$ shall not exceed the amount agreed in the contract. Otherwise the retailer would take them for free, which is a kind of loss to the producer. (4.1.12) limits the maximum total amount of strawberry sold in the discounted price in a week $F_{\cdot}^{tot}$.

Specifically, the $batches$ is used to distinguish the strawberries arriving on different days. Strawberries arriving on the same day are grouped into the same batch. This is for tracking the different deterioration progress of the strawberries as the producer would like to deliver those closest to the expiration date to the retailer at first. As illustrated in Figure 4.1, the rows represent the batches, the columns represent the days, and the colour in each entry represents the amount of strawberries. It can be seen that the colour fades as the time goes by, indicating the strawberries are deteriorating.

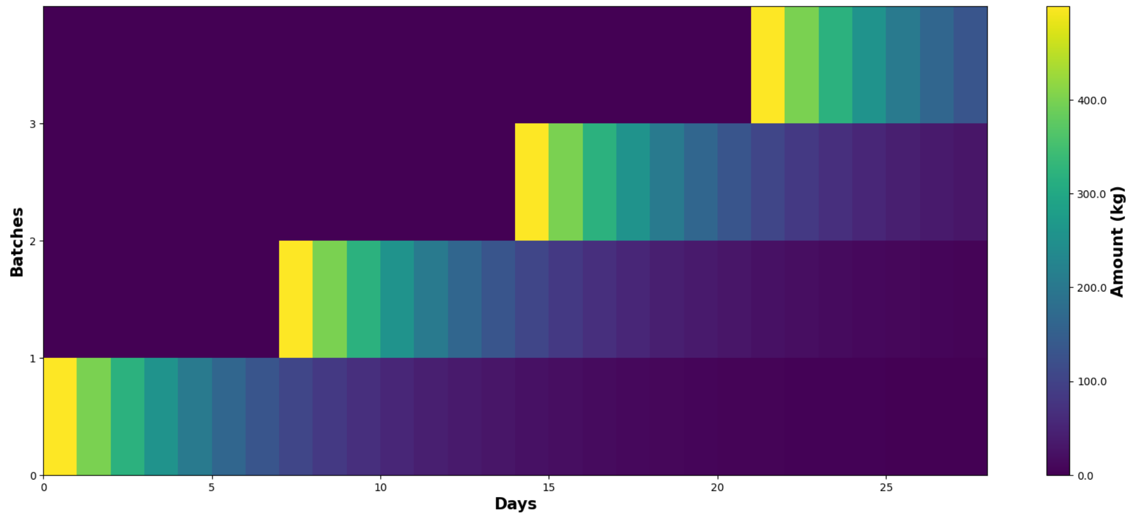

Figure 4.1: The data about the amount of strawberries kept in the inventory. The rows represent the batches of strawberries, the columns represent days, and the colour in each entry indicates the amount of strawberries. For example, the entry in the second row of the 14th column represents the amount of the second batch of strawberries left in the inventory on the 14th day is about 100kg in that month.

Then the model judges whether the strawberries should be disposed due to exceeding the capacity, and calculates the amount of the strawberries to dispose.

$$left_j^{tot} = \sum_{i=0}^{batches-1} left_{i,j} \tag{4.1.14}$$

$$left_j^{tot} \geq Cap_{inv} \rightarrow dispose_j^{tot} = left_j^{tot} - Cap_{inv} \tag{4.1.15}$$

$$\forall i = \{0, 1, ..., batches - 1\}$$

$$\forall j = \{0, 1, ..., 27\}$$

If the total amount of strawberries left in the inventory $left_{\cdot}^{tot}$ exceeds the inventory capacity, then the excessive amount will be disposed, represented by $dispose_{\cdot}^{tot}$.

The FIFO (first-in-first-out) is applied to transportation($T_{\cdot,\cdot}$), sales in discounted price($F_{\cdot,\cdot}$) and the dispose($dispose_{\cdot}^{tot}$). It is applied because that the producer would like to sell or dispose the strawberries arriving in the inventory at first to prevent strawberries from expiration in his own inventory. The FIFO for $T_{\cdot,\cdot}$ and $F_{\cdot,\cdot}$ is implemented as the following formulas:

$$(T_{i,j} + F_{i,j}) * \sum_{k=0}^{i-1} I_{k,j} = 0 \tag{4.1.16}$$

$$\forall i = \{0, 1, ..., batches - 1\}$$

$$\forall j = \{0, 1, ..., 27\}$$

Similarly, the FIFO for dispose is described as:

$$dispose_{i,j} * \sum_{k=0}^{i-1} left_{k,j} = 0 \quad (4.1.17)$$

$$\sum_{i=0}^{batches-1} dispose_{i,j} = dispose_j^{tot} \quad (4.1.18)$$

$$dispose_{i,j} \geq 0 \quad (4.1.19)$$

$$\forall i = \{0, 1, ..., batches - 1\}$$

$$\forall j = \{0, 1, ..., 27\}$$

The logic behind it is that if the model decides to transport or dispose strawberries from a specific batch, then there should be no strawberries left from all the previous batches according to the FIFO. If a specific batch still holds strawberries, then no transportation or dispose from later batches is allowed.

Another part of the model is that the strawberries left in the retailer's inventory $I_{\cdot,\cdot}^{retailer}$ should beyond a minimum level everyday. It is designed to be $\frac{1}{10}$ of the weekly prediction strawberries demand $D_{\cdot}^{prediction}$, for simulating an estimated daily demand from the market. In order to meet this requirement, multiple deliveries in a week can be observed. Otherwise the strawberries kept in the retailer's inventory will be less than the minimum level because of the deterioration if there are no supplements. It is formulated as the following:

$$D_j^{retailer} = D_{j//7}^{prediction}/10, \ \forall j = \{0, 1, 2, 3\} \quad (4.1.20)$$

$$I_0^{retailer} = 0 \quad (4.1.21)$$

$$I_j^{retailer} = T_{j-T_t}^{tot} * (\sigma_t^{T_t}) + I_{j-1}^{retailer} * \sigma_r, \ \forall j = \{1, 2, ..., 27\} \quad (4.1.22)$$

$$I_j^{retailer} \geq D_j^{retailer}, \ \forall j = \{1, 2, ..., 27\} \quad (4.1.23)$$

The formulas (4.1.21) and (4.1.22) describes the retailer's inventory state without distinguishing batches. It is also considered that there is a delay of T_t days and a deterioration rate $(1 - \sigma_t)$ during the transportation from the producer to the retailer. The strawberries kept in the retailer's inventory perish in rate $(1 - \sigma_r)$. It is often assumed that the strawberries deteriorate faster in the retailer's inventory compared with when being stored in the producer's inventory as the retailer's freezer often has worse performance.

Trucks are also needed for delivering the strawberries in $T_{\cdot,\cdot}$ and $F_{\cdot,\cdot}$:

$$Truck_j * Cap_{truck} \geq T_j^{tot} + F_j^{tot} \quad (4.1.24)$$

$$\forall j = \{0, 1, ..., 27\}$$

The constraint regulates that the number of trucks should be sufficient to carry all the strawberries to be delivered. The number of the trucks shall be integer.

The costs and the profit for the problem is defined as following:

$$Income_{contract} = \sum_{j=0}^{3} D_j^{contract} * P_j^{contract} \tag{4.1.25}$$

$$Income_{discount} = \sum_{j=0}^{3} D_j^{contract} * P_j^{contract} * w_{discount} \tag{4.1.26}$$

$$C_{holding} = \sum_{j=0}^{27} \sum_{i=0}^{batches-1} I_{i,j} * Z_h \tag{4.1.27}$$

$$C_{transporation} = \sum_{j=0}^{27} \sum_{i=0}^{batches-1} (T_{i,j} + F_{i,j}) * Z_t + \sum_{j=0}^{27} Truckj * Z_{truck} \tag{4.1.28}$$

$$C_{shortage} = \sum_{i=0}^{3} D_i^{contract} - \sum_{j=0}^{6} T_{i*7+j}^{tot} * P_i^{contract} * w_{shortage} \tag{4.1.29}$$

$$C_{dispose} = \sum_{j=0}^{27} dispose_j^{tot} \tag{4.1.30}$$

$$C_{deterioration} = \sum_{j=0}^{27} \sum_{i=0}^{batches-1} I_{i,j} * \sigma_i \tag{4.1.31}$$

And the target of the model is:

$$\begin{aligned} Objective = {} & Income_{contract} + Income_{discount} \\ & - C_{holding} - C_{transportation} - C_{shortage} - C_{dispose} - C_{deterioration} \end{aligned} \tag{4.1.32}$$

The objective is composed by the income from selling the strawberries through the contract $Income_{contract}$ and sales in discounted price $Incone_{discount}$, the cost introduced by keeping strawberries in the inventory $C_{holding}$, transportation cost $C_{transportation}$, shortage cost $C_{shortage}$ caused by failing to transport enough amount of strawberries to the retailer as signed in the contract, and the penalties for disposing the strawberries $C_{dispose}$ and deterioration $C_{deterioration}$. Though $C_{dispose}$ and $C_{deterioration}$ will not be reflected on the producer's profit in the reality, as it has already caused loss of strawberries without gaining income, the penalties are still added to encourage the optimizer to find other better solutions rather than disposing the strawberries or letting them deteriorate.

4.2. Linearize the Model

Then the MILP model is linearized for faster convergence and better performance. The linearized MILP model is described as following:

$$max\ Profit = Income_{contract} + Income_{discount} \qquad (4.2.1)$$
$$- C_{holding} - C_{transportation} - C_{shortage} - C_{dispose} - C_{deterioration}$$
$$s.t.\ M = 100000, \epsilon = 0.000001 \qquad (4.2.2)$$
$$D_w^{difference} = D_w^{contract} - D_w^{prediction} \qquad (4.2.3)$$
$$D_w^{difference} \geq M * (ind_w^{D^{difference}} - 1) \qquad (4.2.4)$$
$$D_w^{difference} \leq M * ind_w^{D^{difference}} \qquad (4.2.5)$$
$$w_w^{price} = w^{more} * ind_w^{D^{difference}} + w^{less} * (1 - ind_w^{D^{difference}}) \qquad (4.2.6)$$
$$P_w^{contract} = P_w^{prediction} - \frac{(D_w^{difference})^2}{25} * w_w^{price} \qquad (4.2.7)$$
$$left_{i,0} \leq H_{i,0} - T_{i,0} - F_{i,0} \qquad (4.2.8)$$
$$left_{i,j} \leq I_{i,j-1} * \sigma_{inv} + H_{i,j} - T_{i,j} - F_{i,j},\ j \neq 0 \qquad (4.2.9)$$
$$I_{i,j} = left_{i,j} - dispose_{i,j} \qquad (4.2.10)$$
$$T_j^{tot} = \sum_{i=0}^{Batches} T_{i,j} \qquad (4.2.11)$$
$$F_j^{tot} = \sum_{i=0}^{Batches} F_{i,j} \qquad (4.2.12)$$
$$\sum_{j=0}^{6} T_{w*7+j}^{tot} \leq D_w^{contract} \qquad (4.2.13)$$
$$\sum_{j=0}^{6} F_{w*7+j}^{tot} \leq D_w^{contract} * limit \qquad (4.2.14)$$
$$left_j^{tot} = \sum_{i=0}^{batches-1} left_{i,j} \qquad (4.2.15)$$
$$left_j^{tot} \geq Cap_{inv} + M * (ind_j^{cap} - 1) \qquad (4.2.16)$$
$$left_j^{tot} \leq Cap_{inv} + M * ind_j^{cap} \qquad (4.2.17)$$
$$dispose_j^{tot} \geq left_j^{tot} - Cap_{inv} - M * (1 - ind_j^{cap}) \qquad (4.2.18)$$
$$dispose_j^{tot} \leq left_j^{tot} \qquad (4.2.19)$$
$$T_{i,j} + F_{i,j} \geq -M * (1 - ind_{i,j}^{transportation}) + \epsilon \qquad (4.2.20)$$
$$T_{i,j} + F_{i,j} \leq M * ind_{i,j}^{transportation} + \epsilon \qquad (4.2.21)$$
$$\sum_{k=0}^{i-1} I_{k,j} \geq -M * (1 - ind_{i,j}^{inventory}) + \epsilon \qquad (4.2.22)$$
$$\sum_{k=0}^{i-1} I_{k,j} \leq M * ind_{i,j}^{inventory} + \epsilon \qquad (4.2.23)$$
$$ind_{i,j}^{transport} + ind_{i,j}^{inventory} \leq 1 \qquad (4.2.24)$$
$$dispose_{i,j} \geq -M * (1 - ind_{i,j}^{dispose}) + \epsilon \qquad (4.2.25)$$
$$dispose_{i,j} \leq M * ind_{i,j}^{dispose} + \epsilon \qquad (4.2.26)$$

$$\sum_{k=0}^{i-1} left_{k,j} \geq -M * (1 - ind_{i,j}^{left}) + \epsilon \tag{4.2.27}$$

$$\sum_{k=0}^{i-1} left_{k,j} \leq M * ind_{i,j}^{left} + \epsilon \tag{4.2.28}$$

$$ind_{i,j}^{dispose} + ind_{i,j}^{left} \leq 1 \tag{4.2.29}$$

$$dispose_{i,j} * \sum_{k=0}^{i-1} left_{k,j} = 0 \tag{4.2.30}$$

$$D_j^{retailer} = D_{j//7}^{prediction}/10 \tag{4.2.31}$$

$$I_0^{retailer} = 0 \tag{4.2.32}$$

$$I_j^{retailer} = T_{j-T_t}^{tot} * \sigma_t^{T_t} + I_{j-1}^{retailer} * \sigma_r,\ j \neq 0 \tag{4.2.33}$$

$$I_j^{retailer} \geq D_j^{retailer} \tag{4.2.34}$$

$$Truck_j * Cap_{truck} \geq T_j^{tot} + F_j^{tot} \tag{4.2.35}$$

$$Income_{contract} = \sum_{j=0}^{3} D_j^{contract} * P_j^{contract} \tag{4.2.36}$$

$$Income_{discount} = \sum_{j=0}^{3} D_j^{contract} * P_j^{contract} * w_{discount} \tag{4.2.37}$$

$$C_{holding} = \sum_{j=0}^{27} \sum_{i=0}^{batches-1} I_{i,j} * Z_h \tag{4.2.38}$$

$$C_{transporation} = \sum_{j=0}^{27} \sum_{i=0}^{batches-1} (T_{i,j} + F_{i,j}) * Z_t + \sum_{j=0}^{27} Truck_j * Z_{truck} \tag{4.2.39}$$

$$C_{shortage} = \sum_{i=0}^{3} D_i^{contract} - \sum_{j=0}^{6} T_{i*7+j}^{tot} * P_i^{contract} * w_{shortage} \tag{4.2.40}$$

$$C_{dispose} = \sum_{j=0}^{27} dispose_j^{tot} \tag{4.2.41}$$

$$C_{deterioration} = \sum_{j=0}^{27} \sum_{i=0}^{batches-1} I_{i,j} * (1 - \sigma_i) \tag{4.2.42}$$

$$D_w^{contract}, P_w^{contract} \geq 0 \tag{4.2.43}$$

$$left_{i,j}, I_{i,j}, T_{i,j}, F_{i,j}, dispose_{i,j} \geq 0 \tag{4.2.44}$$

$$ind_w^{Difference}, ind_j^{cap}, ind_{i,j}^{transportation}, ind_{i,j}^{inventory}, ind_{i,j}^{dispose}, ind_{i,j}^{left} = \{0, 1\} \tag{4.2.45}$$

$$\forall i = \{0, 1, ..., batches\}$$

$$\forall j = \{0, 1, ..., 27\}$$

$$\forall w = \{0, 1, 2, 3\}$$

The changes for the linearization mainly happen in the decision of $w_\cdot^{price}$, calculation of the $dispose_{\cdot,\cdot}$ and the implementations of the FIFO. However, the definition of the

$Income_{retailer}$ and $Income_{discount}$ relates to the multiplication of $D_{\cdot}^{contract}$ and $P_{\cdot}^{contract}$, which is a polynomial formula dominated by $D_{\cdot}^{contract^2}$ after simplification and cannot be linearized. Luckily, this component can still be solved by Gurobipy when the parameter `nonConvex` is set to 2.

4.3. Robust Optimization

The results given by the model mentioned above is determined, which means the best profit can only be observed if all the external values are impossibly exactly equal to the predictions. Sometimes even slight fluctuation results in great loss in final profit. To improve this, robust optimization is applied to mitigate such potential loss, by considering about the worst cases based on the prediction distribution.

A big disadvantage of the current model is that sometimes the uncertainty of the yield prediction may cause severe shortage cost. It happens when the yield is less than the prediction and cannot satisfy the amount signed in the contract. The robust optimization works by setting the yield amount to the worst case amount, forcing the constraints not to violate this amount. By tuning the hyper-parameter Γ, the threshold value of the worst case amount can be adjusted to be slightly higher than the worst-case value, controlling the conservatism of the model. The calculation of the worst-case value is visualized in Figure 4.2.

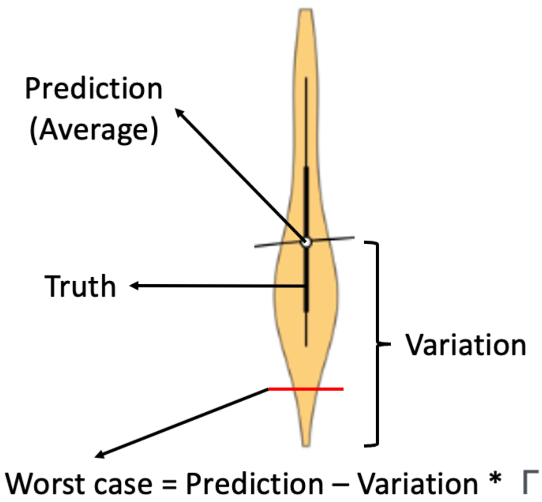

Figure 4.2: The worst case equals to the average value minus a specific portion of the variation. The portion controls the conservatism of the model.

The reformulation of the robust optimization model happens on the formulas describing the state of the producer's inventory. The formulas are changed from

$$left_{i,0} \leq H_{i,0} - T_{i,0} - F_{i,0} \tag{4.3.1}$$
$$left_{i,j} \leq I_{i,j-1} * \sigma_{inv} + H_{i,j} - T_{i,j} - F_{i,j} \tag{4.3.2}$$
$$\forall i = \{0, 1, ..., batches - 1\}$$
$$\forall j = \{1, 2, ..., 27\}$$

to the tight formulation of

$$left_{i,j} \leq \sum_{k=0}^{j}(H_{i,k} - T_{i,k} - F_{i,k} - qi, k * \Gamma_{i,k} - r_{i,k,j} * \sigma_i^{j-k} - \sum_{k=0}^{j-1} dispose_{i,k} \tag{4.3.3}$$
$$q_{i,j} + r_{i,k,j} \geq H_{i,k}^{variation} \tag{4.3.4}$$
$$\forall i = \{0, 1, ..., batches - 1\}$$
$$\forall j = \{0, 1, ..., 27\}$$

In this robust optimization, the box uncertainty is considered and the parameter Γ controls the worst cast-scenario by using:

$$H_{I,j}^{worst} = H_{i,j}^{average} - \Gamma * H_{i,j}^{variation} \tag{4.3.5}$$
$$\forall i = \{0, 1, ..., batches - 1\}$$
$$\forall j = \{0, 1, ..., 27\}$$

5

Results

In this chapter, the results of all experiments are illustrated and explained. The implementation of the entire project is done by programming in Python. Specifically, the building of the MILP model is done using the Gurobipy package. Pytorch is applied to build the LSTM neural networks. Other typical packages such as Numpy, Matplotlib are also used in the project.

5.1. Predictions Using LSTM

This section illustrates the prediction results and the loss of the LSTMs. The average values are used as the final predictions for all the ensembles of LSTMs.

The LSTMs are trained using the hyper-parameters mentioned in Table 3.1. Figure 5.1 to Figure 5.9 visualize the train loss, validation loss and the normalized test results of the three kinds of LSTMs compared with the truth.

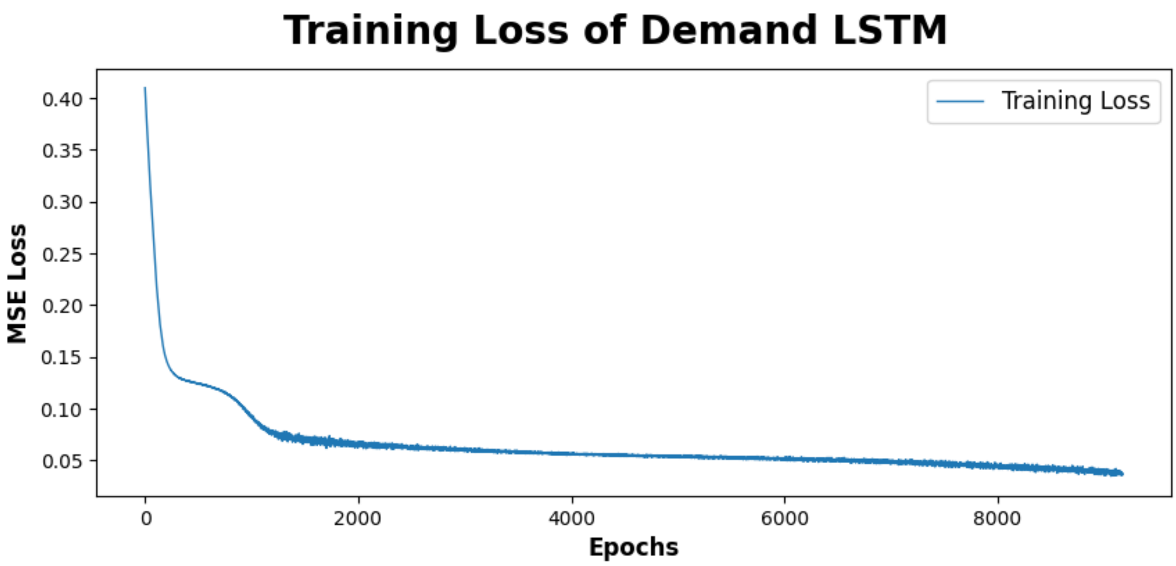

Figure 5.1: Demand LSTMs training loss.

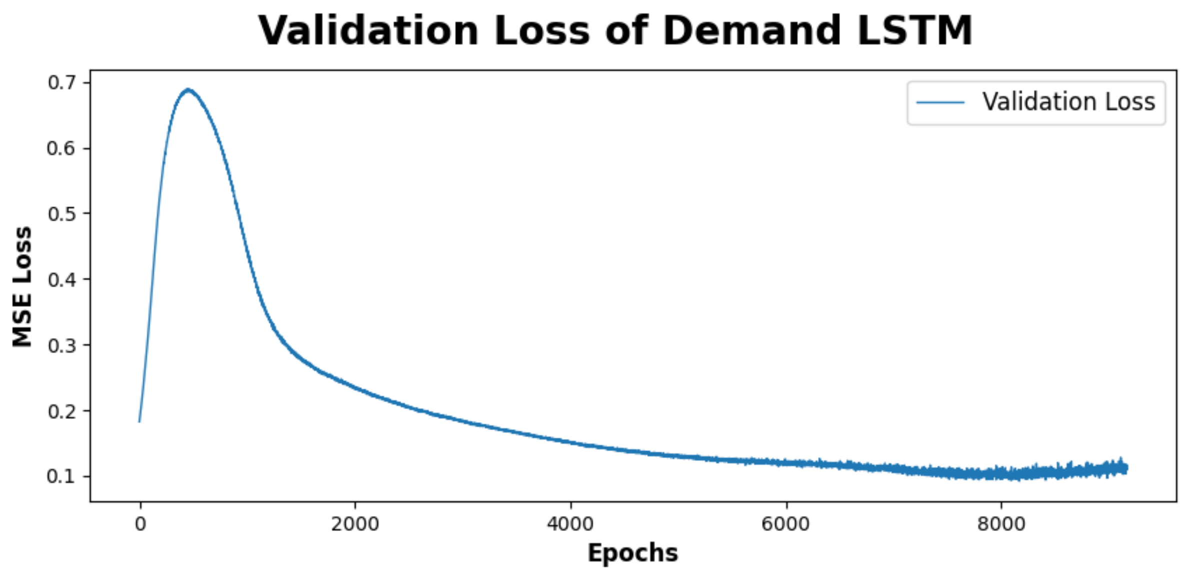

Figure 5.2: Demand LSTMs validation loss.

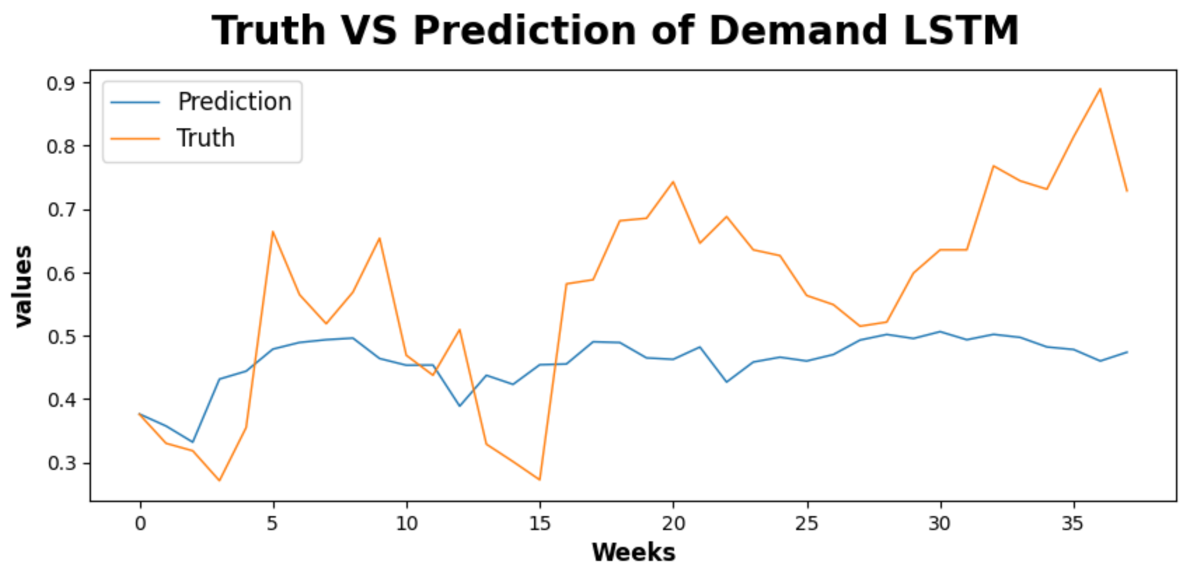

Figure 5.3: Demand LSTMs test.

It can be observed from the above figures that though the validation loss has reached the minimum value, the prediction does not match with the truth very perfectly for the demand LSTMs, but the general value level matches well. It is because that the pattern behind the data is much more noisy compared with other data, making it hard for LSTM to predict. However, a general value level is also meaningful to the producer for making decisions as it offers a possible range indicating where the demand will locate.

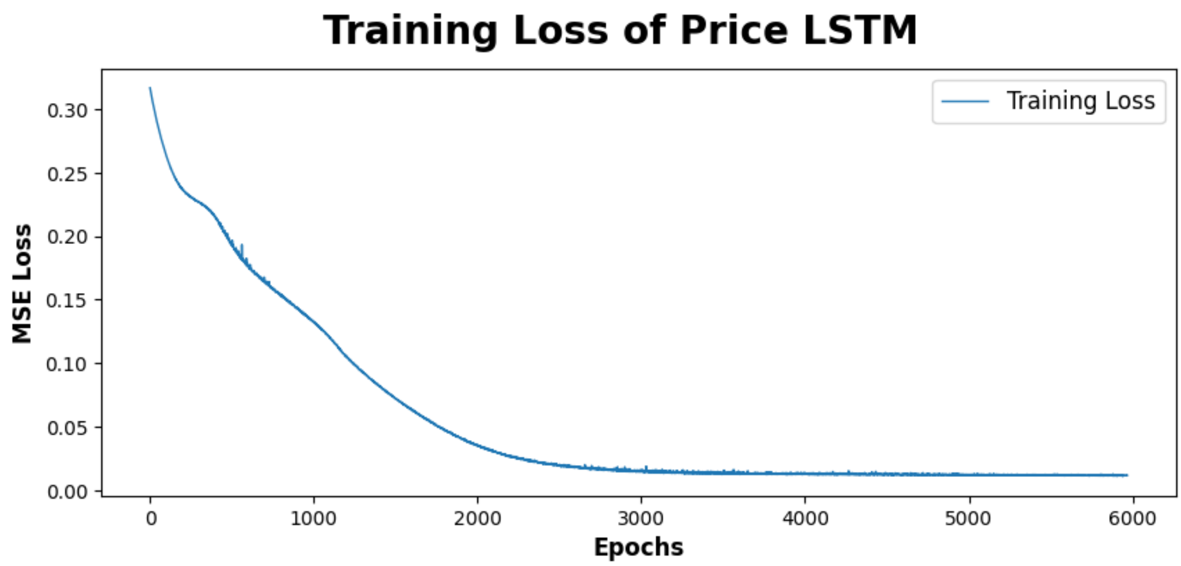

Figure 5.4: Price LSTMs training loss.

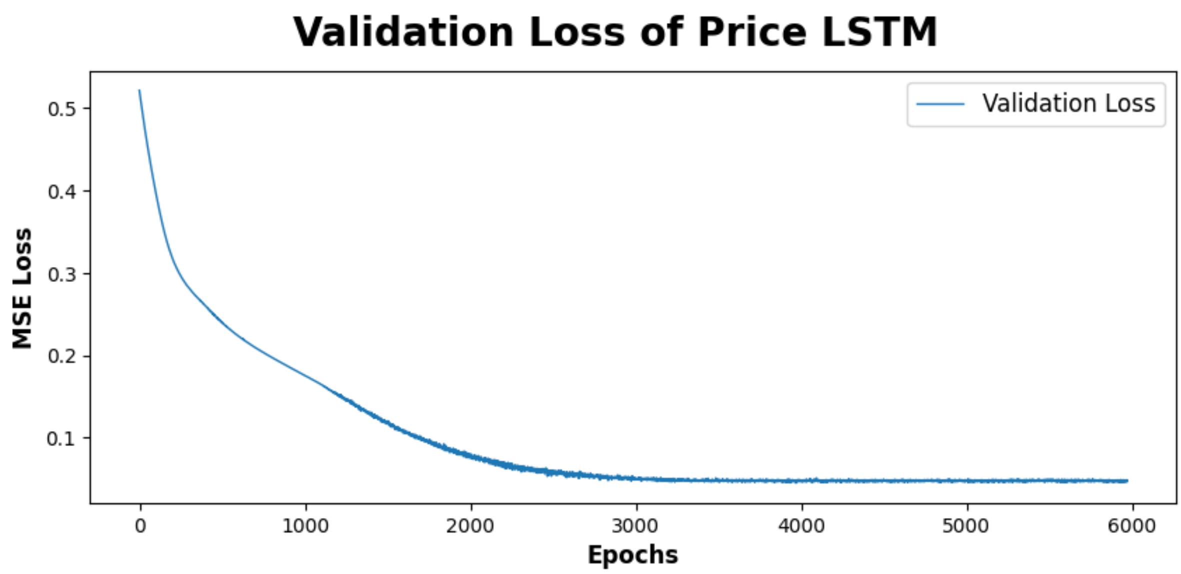

Figure 5.5: Price LSTMs validation loss.

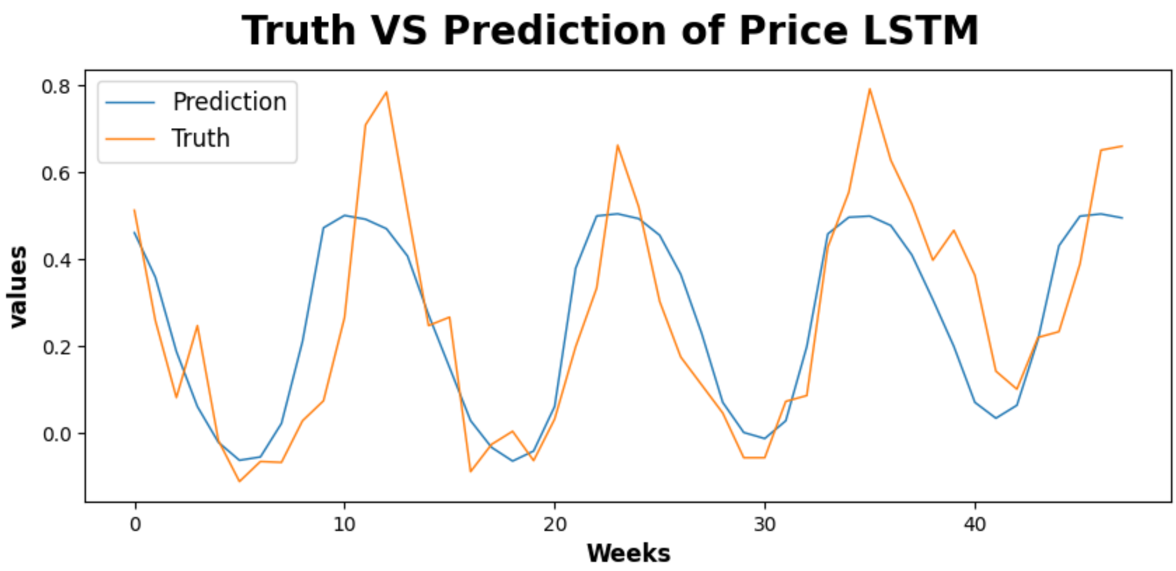

Figure 5.6: Price LSTMs test.

These figures illustrate that the price can be predicted better compared with predictions of the demand, because the price values follow a more regular sine-wave like distribution.

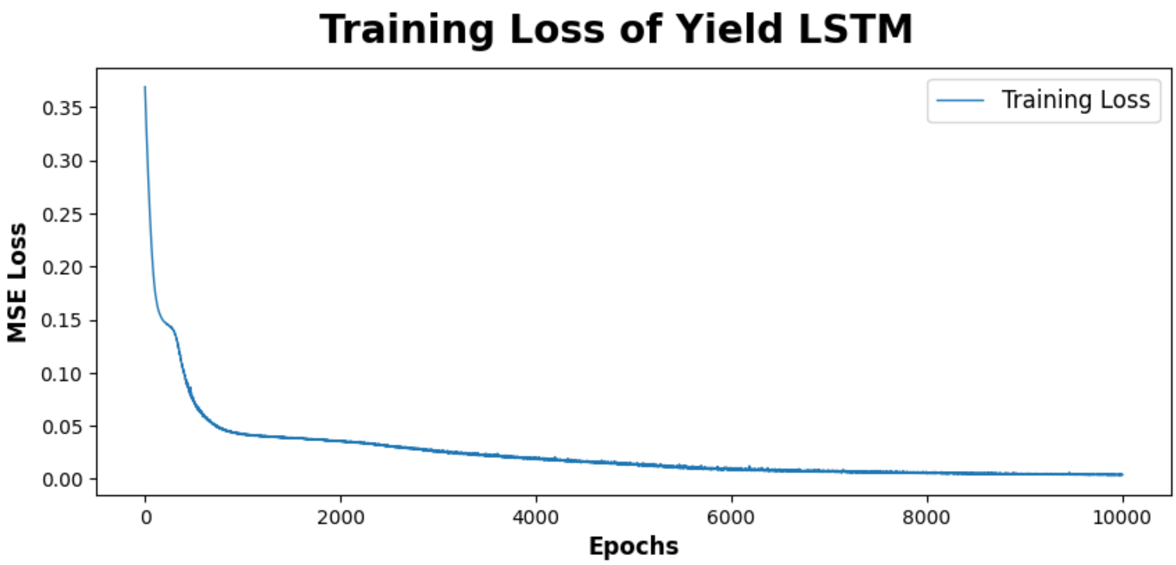

Figure 5.7: Yield LSTMs validation loss.

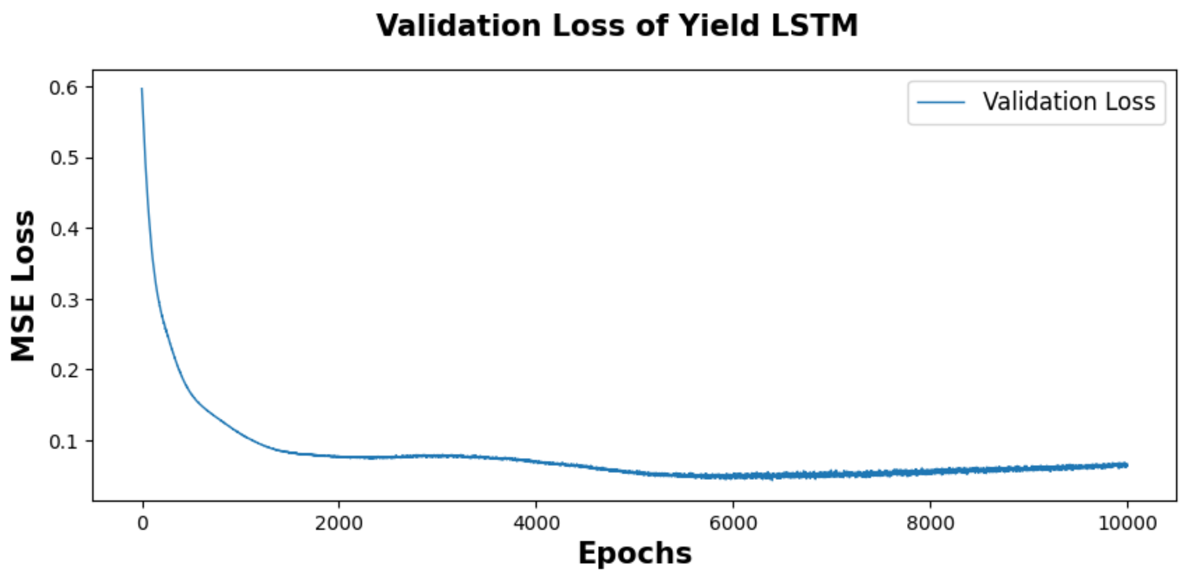

Figure 5.8: Yield LSTMs validation loss.

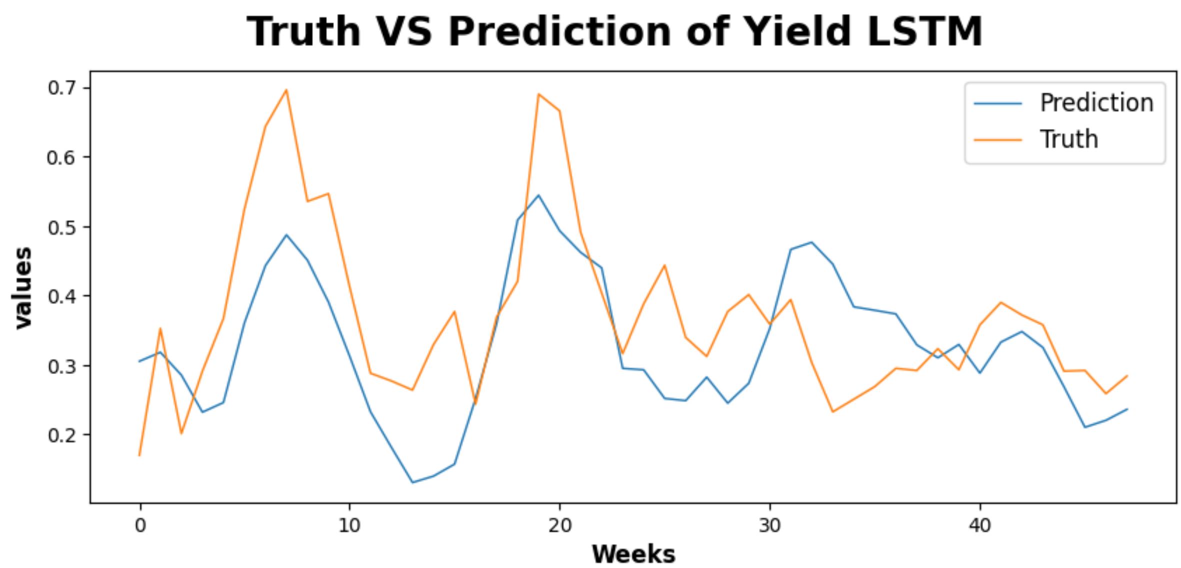

Figure 5.9: Yield LSTMs test.

It can be seen that the predictions of first-half yield also match relatively well with the truth, but there is a mis-match in the second-half. It is because that the regular pattern has disappeared since week 24.

The final test loss for the three kinds of the LSTMs is illustrated in Table 5.1, all measured by MSE.

Table 5.1: The training loss, validation loss and the test loss of the three kinds of LSTMs.

	Demand LSTM	Price LSTM	Yield LSTM
Train Loss	0.0359	0.0151	0.0057
Validation Loss	0.0874	0.0457	0.0642
Test Loss	0.0256	0.0552	0.0528

It can be observed that the LSTMs can generally capture the patterns of the data, and the price data can be predicted pretty well. In order to increase the performance, more complex LSTMs can be used but more data is required for the training.

5.2. Problem Validation

Before defining the MILP model to solve the problem, the convexity of how the objective profit changes with contract demands $D_{\cdot}^{contract}$ is studied. There is a need to build the MILP optimizer only if the problem is convex or partially convex. Otherwise if the objective changes linearly, the conclusion can be drawn like 'Always sell all the strawberries' without the assistance from the model. And if there are multiple optimal points on the curve, it is require to manually compare the locally optimal solutions given by the optimizer to find the global optimal solution.

When studying the convexity of problem, it is expected to see most of the trade-offs in the MILP model. Therefore, the parameters are set as Table 5.2, which make the problem interesting that most of the trade-offs can be observed and discussed.

Table 5.2: The setting of the parameters of the MILP model when verifying the convexity of the problem.

Parameter	Z_t	Z_h	Z_{truck}	Cap_{inv}	Cap_{truck}	$limit$	$w_{shortage}$
Value	0.1	0.1	100	5000	500	0.3	1
Parameter	σ_i	σ_t	σ_r	w_{more}	w_{less}	$w_{discount}$	$w_{difference}$
Value	0.9	0.9	0.6	0.0001	-0.0001	0.5	1

To study the convexity of the problem, a determined model executing the same strategy as the optimizer is built using Python. In this way, other decision variables will be fixed during modifying the values of the $D_{\cdot}^{contract}$.

5.2.1. The Influence of $D_0^{contract}$

Firstly, how the change of the first week contract demand $D_0^{contract}$ affects the profit is studied, because the contract demand $D_{\cdot}^{contract}$ are the primary decision variables affecting the profit. The result is shown in Figure 5.10

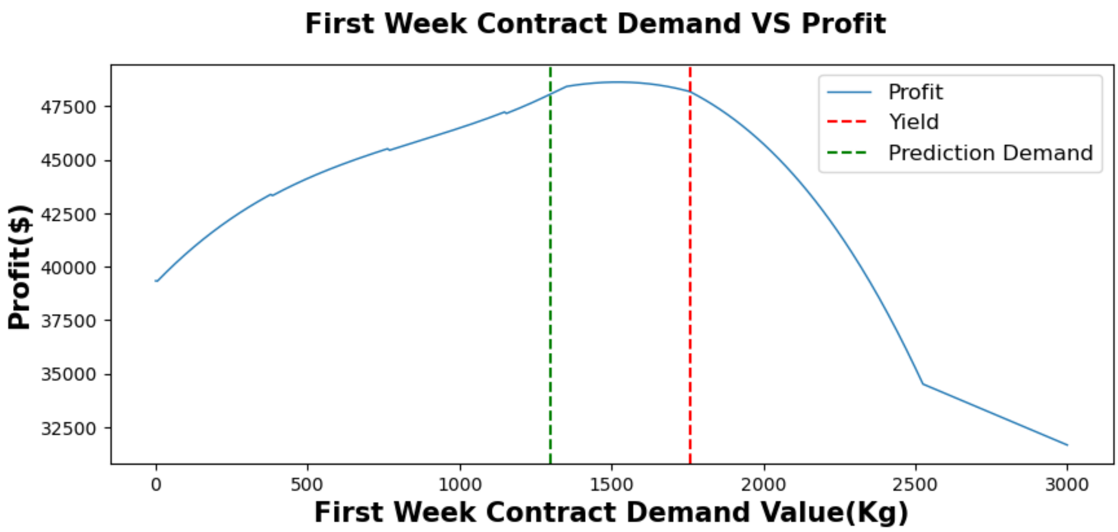

Figure 5.10: The curve indicates how the profit changes when the contract demand for the first week($D_0^{contract}$) is changing together. It can be observed the curve is convex, indicating the problem is meaningful. The optimal point can be observed between the prediction demand and the yield.

As the figure shows, the curve is convex, indicating the problem is solvable and meaningful. The best value for $D_0^{contract}$ is observed between the prediction demand $D_0^{prediction}$ and the prediction yield for that week, indicating that though the price is a bit lower, the extra amount of strawberries sold introduces more profit. However, the shortage is too huge if $D_0^{contract} > D_0^{prediction}$ and the producer is not advised to sign a contract demand value that high.

In order to understand why the profit changes like this, two components, the income and multiple costs, are studied together with the profit. The results are shown in Figure 5.11 and Figure 5.12. Several results in Figure 5.12 are multiplied by specific factors shown in the legend for making its shape obvious in the figure.

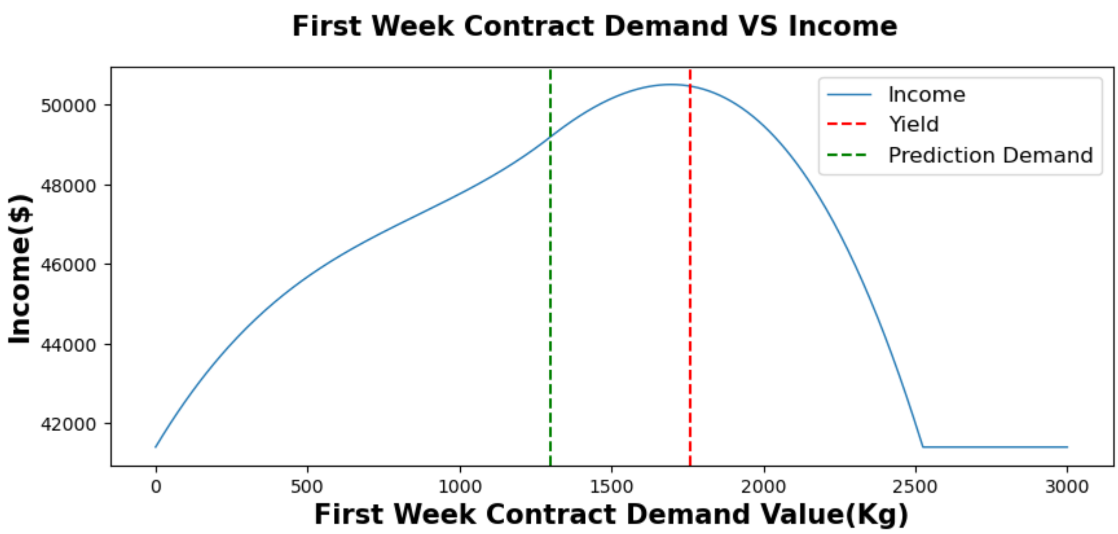

Figure 5.11: The curve shows how the income by contract changes as $D_0^{contract}$ changing. It can be seen that the shape is similar to the final profit.

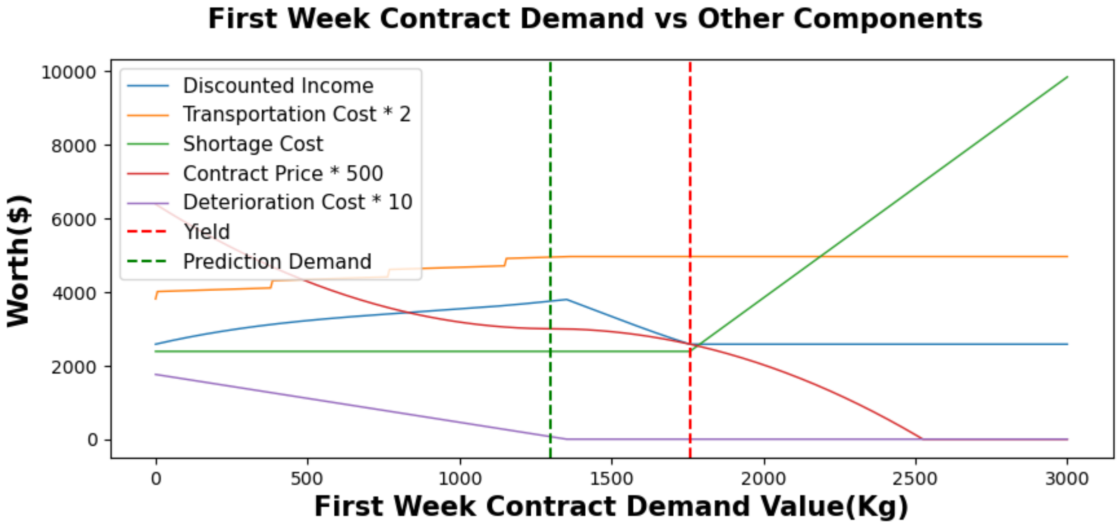

Figure 5.12: The curve shows how the income by contract changes as $D_0^{contract}$ changing. It can be seen that the shape is similar to the final profit.

It appears that the dominant component, $Income$ shown in Figure 5.11, is close to a quadratic function, resulted by the multiplication of $D_0^{contract}$ and $P_0^{contract}$. The income drops dramatically after the peak point between the yield and the prediction demand as from then on the $P_0^{contract}$ decreases too fast to be compensated by the extra sold amount of strawberries. The tail of the curve is flat as the $P_0^{contract}$ there is as low as 0, reaching its lower boundary.

The income gained by selling strawberries in discounted price F peaks at a point between the yield and prediction demand then drops, because all strawberries are delivered to coverage the $D_0^{contract}$ and no strawberries are left for F. Steps can be observed in the transportation cost because of trucks are required for transportation, and the number of trucks can only be integer. Finally, the deterioration cost continuously decreases as there are less strawberries left in the inventory as the amount transported increasing to cover the increasing $D_0^{contract}$.

5.2.2. The Influence of $D_0^{contract}$ and $D_1^{contract}$

Then, the change of profit and other values are studied when $D_0^{contract}$ and $D_1^{contract}$ are set to change independently. The results are shown in Figure 5.13 to Figure 5.17.

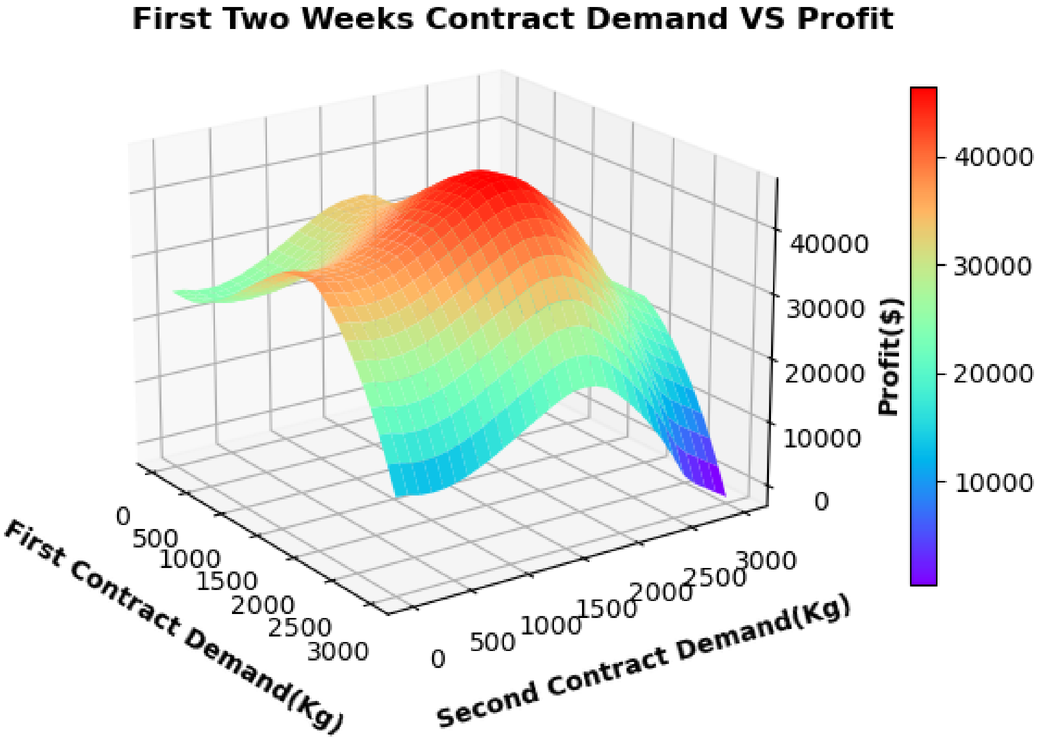

Figure 5.13: The change of profit.

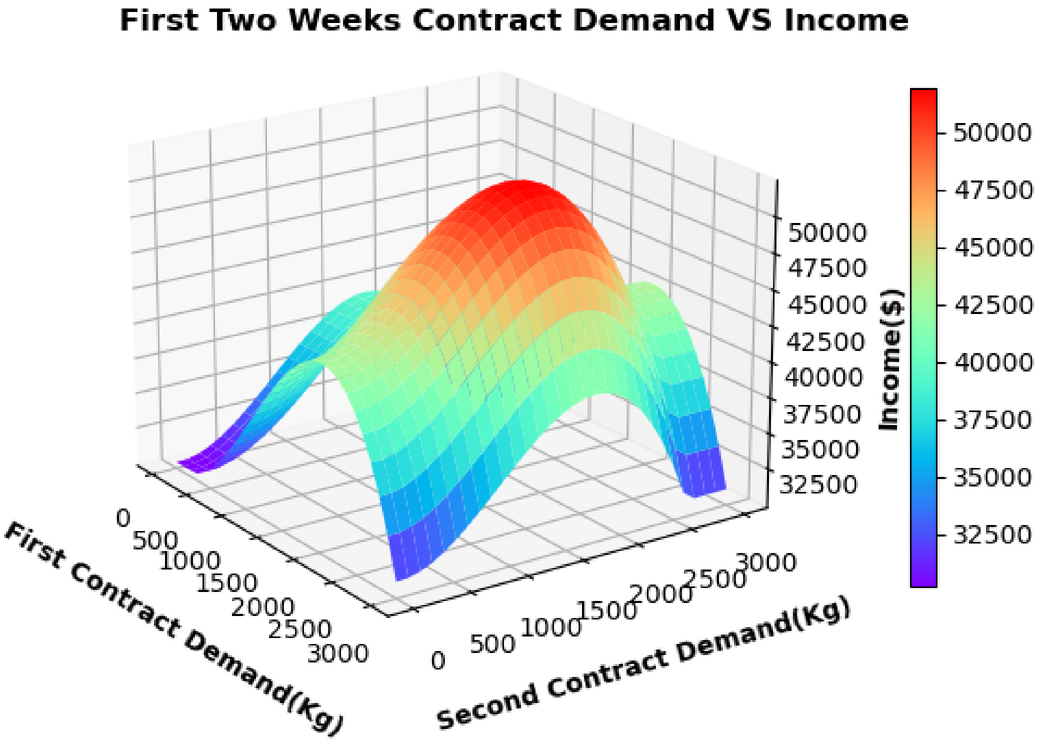

Figure 5.14: The change of income.

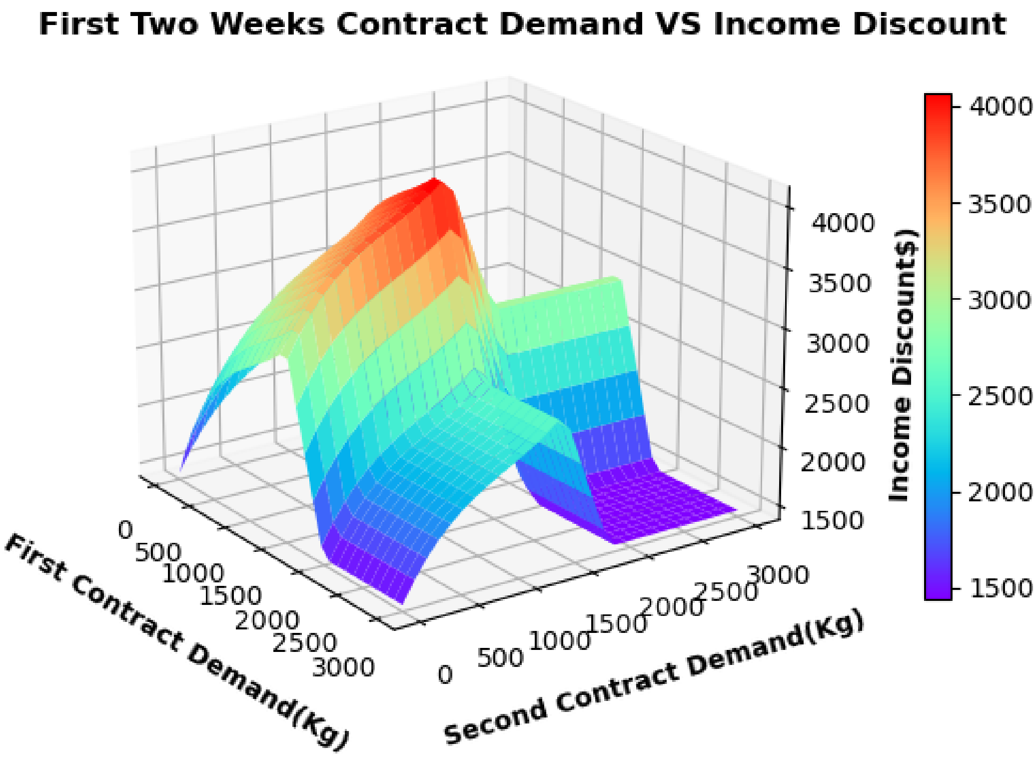

Figure 5.15: The change of discounted income.

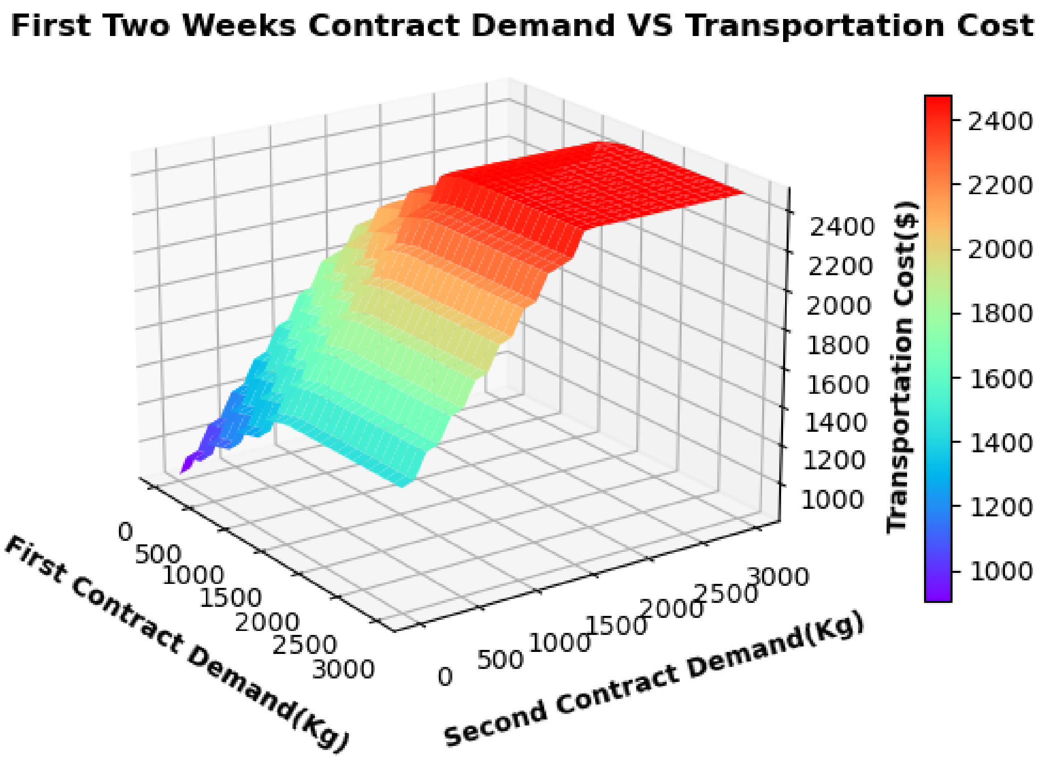

Figure 5.16: The change of transportation cost.

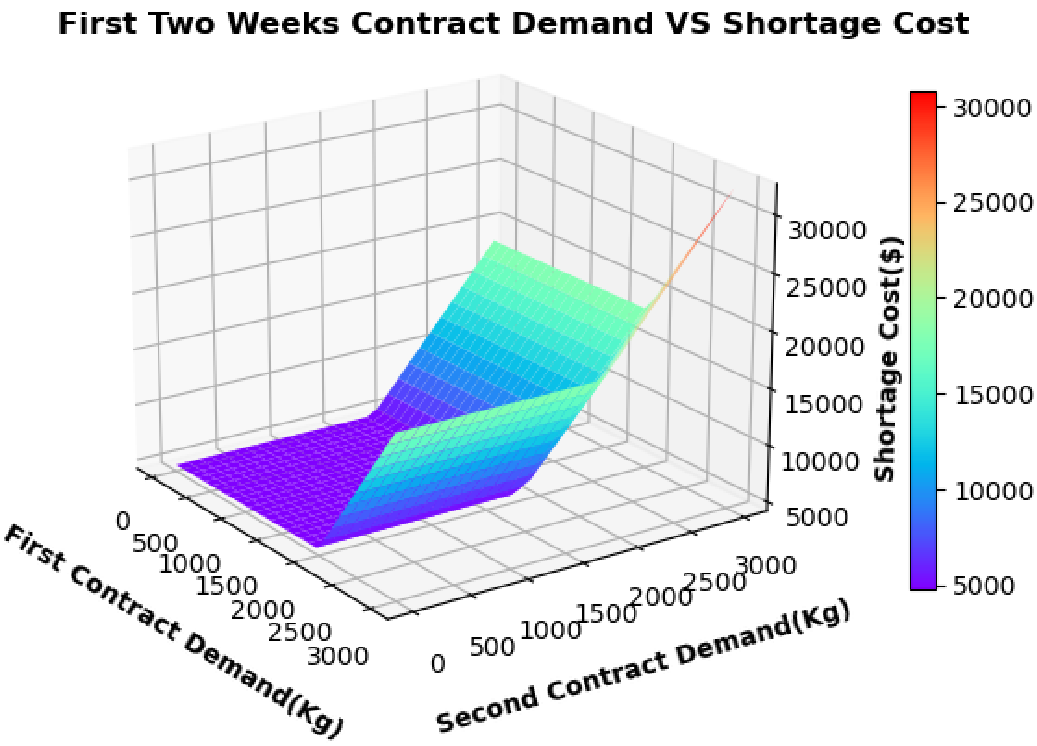

Figure 5.17: The change of shortage cost.

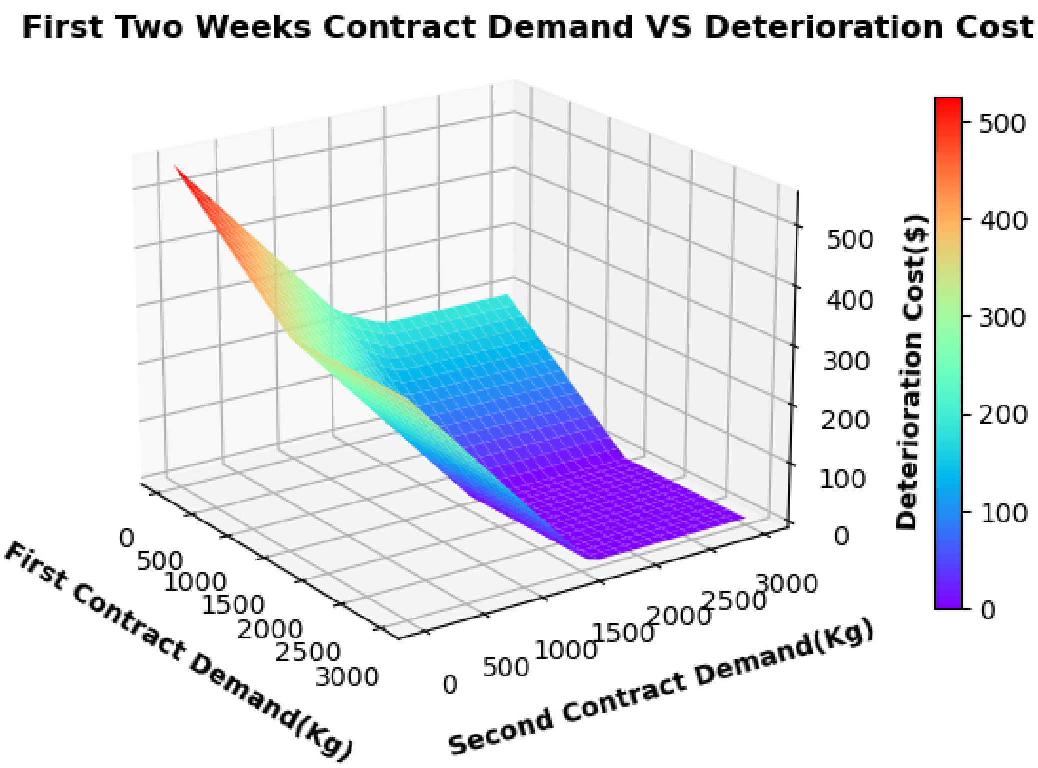

Figure 5.18: The change of deterioration cost.

5.2.3. The Influence of All Contract Demand

In the end, how the profit changes when all the contract demands $D_{\cdot}^{contract}$ change together in the same scale is studied. The standard value is set as the sum of the prediction demand $D_{\cdot}^{prediction}$. The result is shown in Figure 5.19.

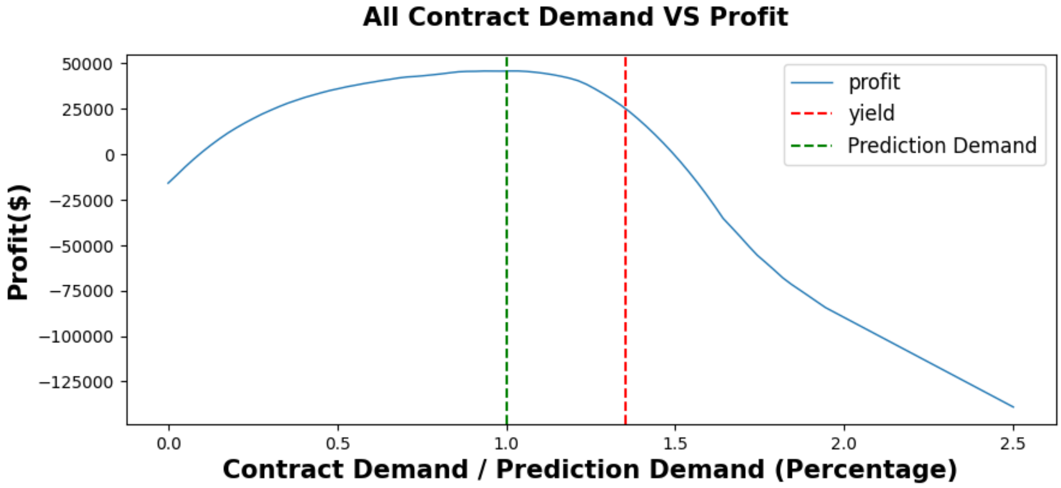

Figure 5.19: The change of profit with all the contract demand.

The curve is still convex and the best point can be observed near the prediction demand. The negative values of the beginning of the line is resulted by the holding cost and dispose cost. Therefore, it is worthy to study how to make the best choice considering all the trade-offs when all the contract demands $D_{\cdot}^{contract}$ can change independently.

5.3. Optimized Strategy from the MILP Model

This section illustrates the results optimized by the MILP model of one scenario picked within numerous possible settings, whose parameters are shown in Table 5.3. Compared with Table 5.2, $w_{shortage}$ is set to 2 as the retailer is not tolerant to the shortage. $limit$ is also decreased to so that the dispose of strawberries can be observed.

Table 5.3: The setting of the parameters of the MILP model when verifying the convexity of the problem.

Parameter	Z_t	Z_h	Z_{truck}	Cap_{inv}	Cap_{truck}	$limit$	$w_{shortage}$
Value	0.1	0.1	100	5000	500	0.1	2
Parameter	σ_i	σ_t	σ_r	w_{more}	w_{less}	$w_{discount}$	$w_{difference}$
Value	0.9	0.9	0.6	0.0001	-0.0001	0.5	1

The values for the $Y_{\cdot}$, $D_{\cdot}^{prediction}$ and $P_{\cdot}^{prediction}$ are given as the following table:

Table 5.4: Values for $Y_{\cdot}$, $D_{\cdot}^{prediction}$ and $P_{\cdot}^{prediction}$

	$j = 0$	$j = 1$	$j = 2$	$j = 3$
$D_j^{prediction}$	2685.43	2673.31	2670.38	2668.97
$P_j^{prediction}$	7.97	8.07	8.07	8.07
$Y_j^{prediction}$	4340.36	4183.75	4102.38	3996.90

After inputting the values from Table 5.4, the optimized strategy is shown from Figure 5.20 to Figure 5.27.

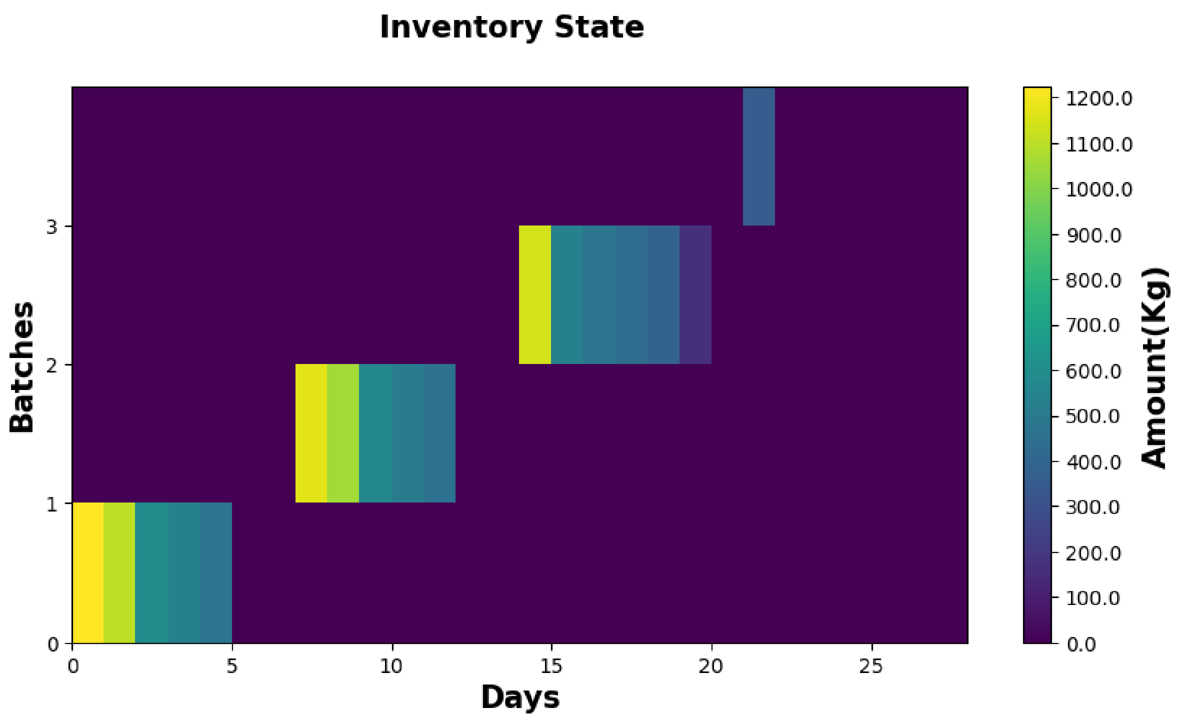

Figure 5.20: The amount of strawberries left in the producer's inventory.

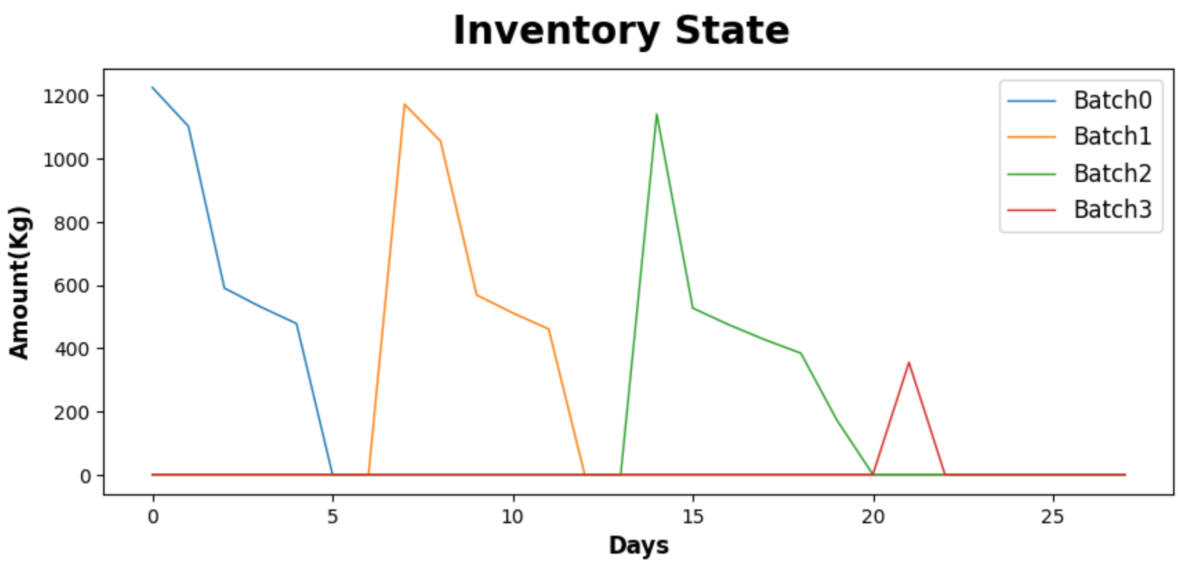

Figure 5.21: The amount of strawberries left in the producer's inventory illustrated using lines.

Figure 5.20 and Figure 5.21 illustrate the amount of strawberries left in the producer's inventory everyday after trading them out. Similar to Figure 4.1, the brightness of the colour of every entry in Figure 5.20 indicates the amount of the strawberries. The x-axis indicates the date and the y-axis represents various batches. As the colour is fading, it can be observed that the amount of the strawberries in the inventory is decreasing because of the deterioration. Figure 5.21 illustrates the amount level of the strawberries more intuitively.

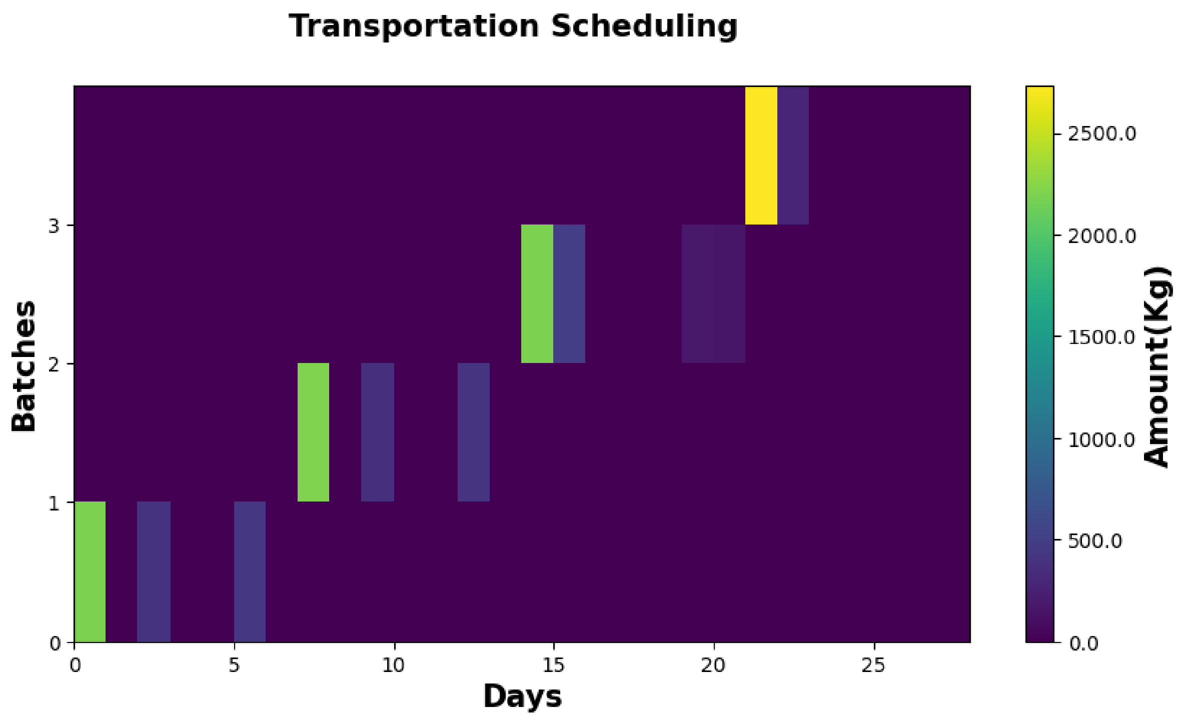

Figure 5.22: The transportation scheduling.

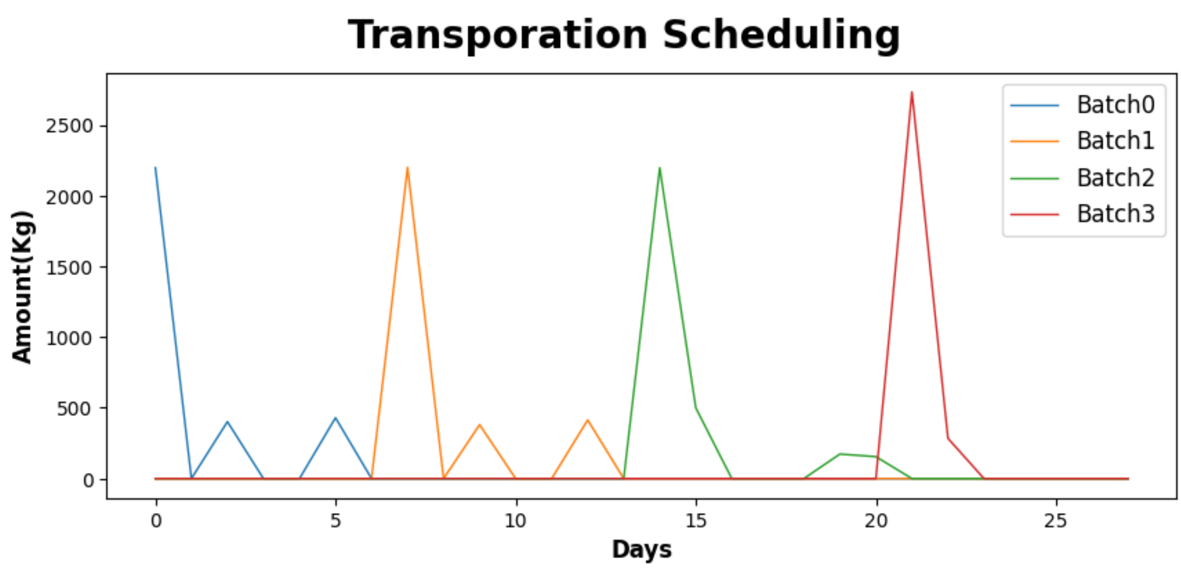

Figure 5.23: The transportation scheduling illustrated using lines.

Figure 5.22 illustrates the transportation scheduling for the producer. It can be observed that sometimes multiple deliveries are arranged from one batch in a week in order to ensure the retailer's inventory level is not less than the minimum requirement. Sometimes transportation delivering the same amount of strawberries is scheduled for several days continuously for just keeping the retailer's inventory level. Figure 5.23 illustrates the transportation scheduling using line plots.

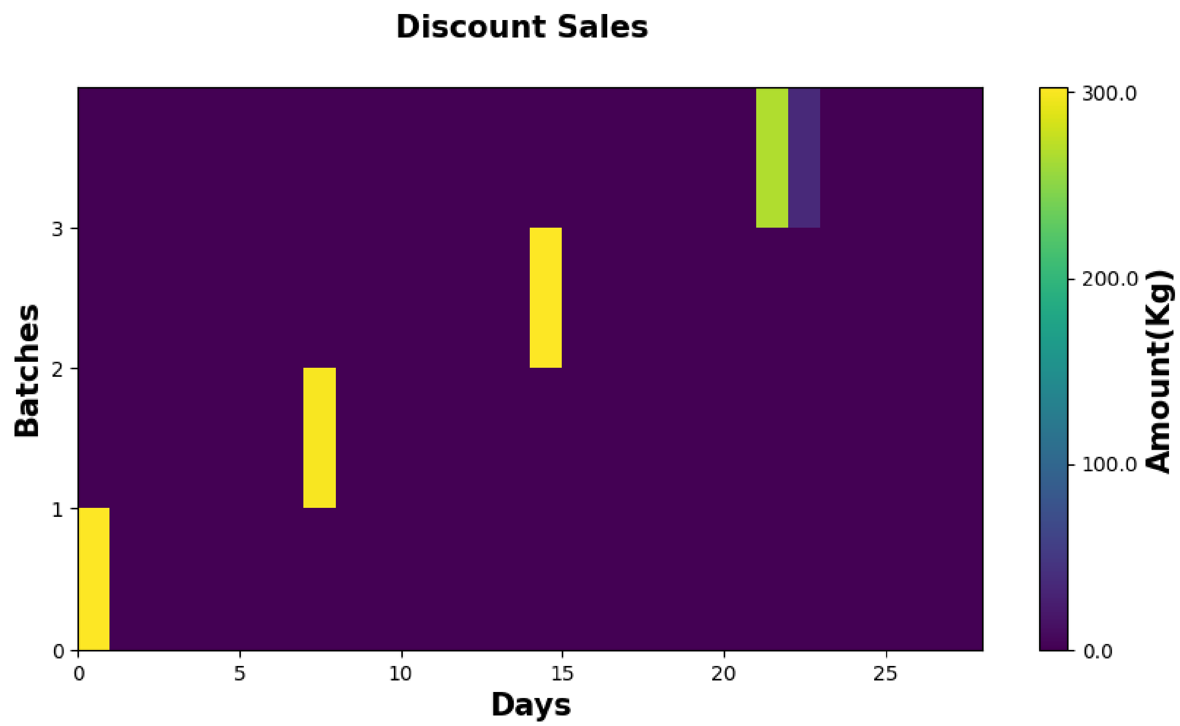

Figure 5.24: The discount sales.

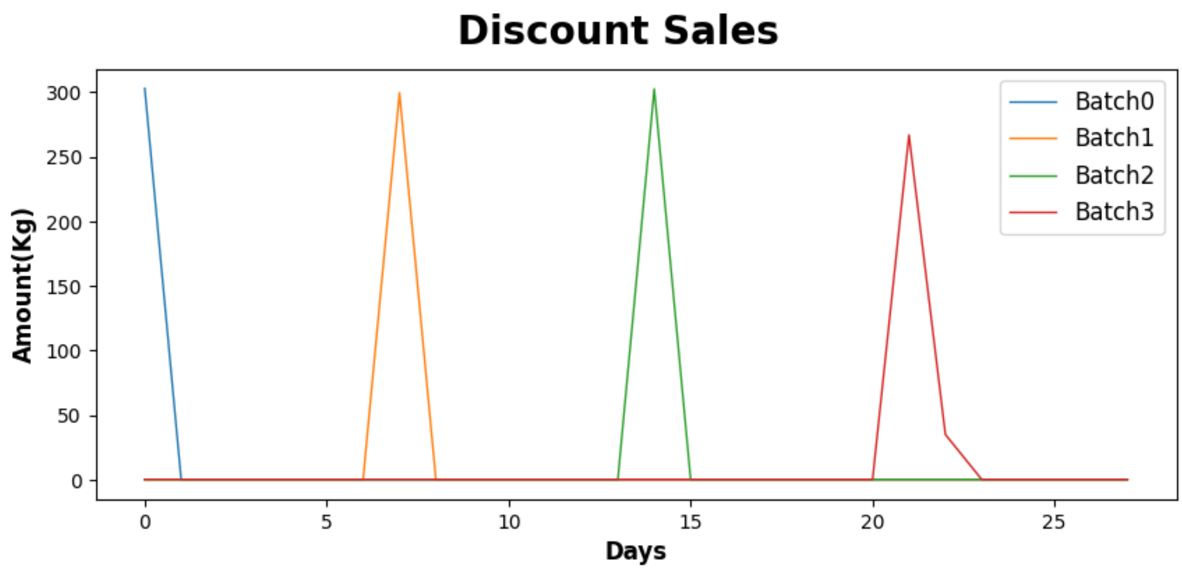

Figure 5.25: The discount sales illustrated lines.

The amount of strawberries sold in the discounted prices are illustrated in Figure 5.24 and Figure 5.25. There are strawberries sold in discounted price as there are excessive strawberries left after covering the amount signed in the contract.

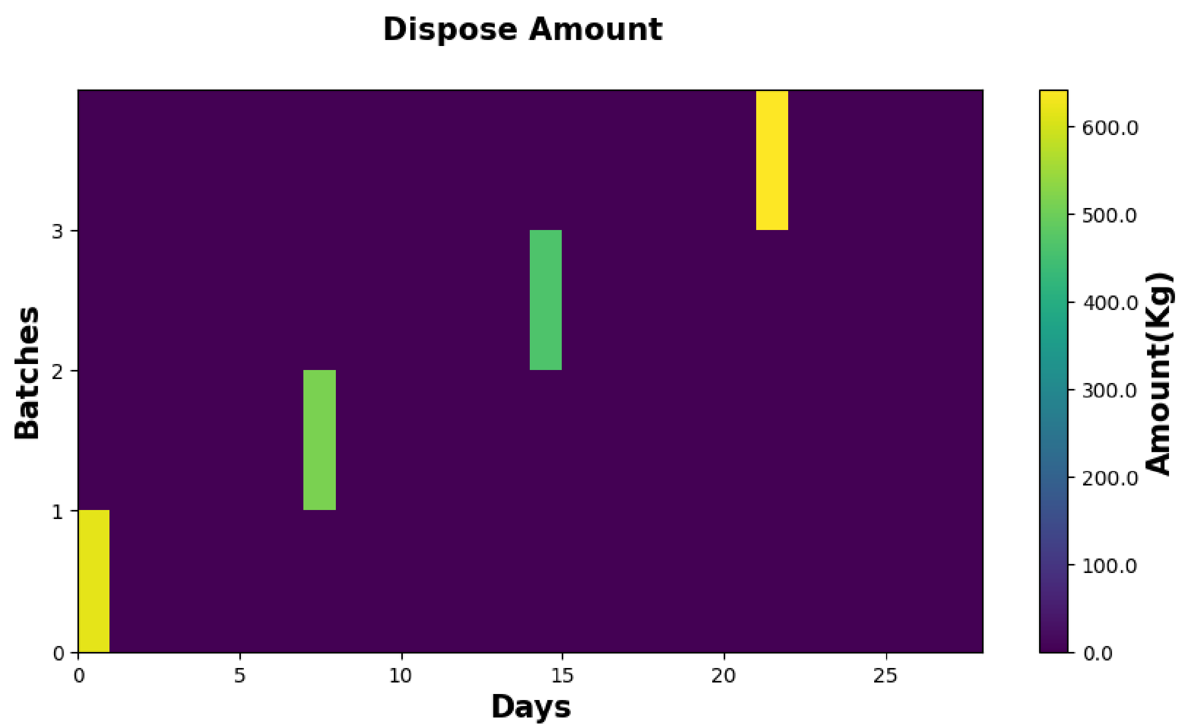

Figure 5.26: There are strawberries disposed in each week.

Figure 5.26 shows that on each Sunday there are strawberries disposed. It happens because that after covering all the demand stated in the contract and the amount to be sold in the discounted price, there are still strawberries left. Though the amount does not exceed the inventory capacity, the producer can do nothing with them. Compared with letting them deteriorate in the inventory introducing deterioration cost, disposing them in time seems to be a better choice.

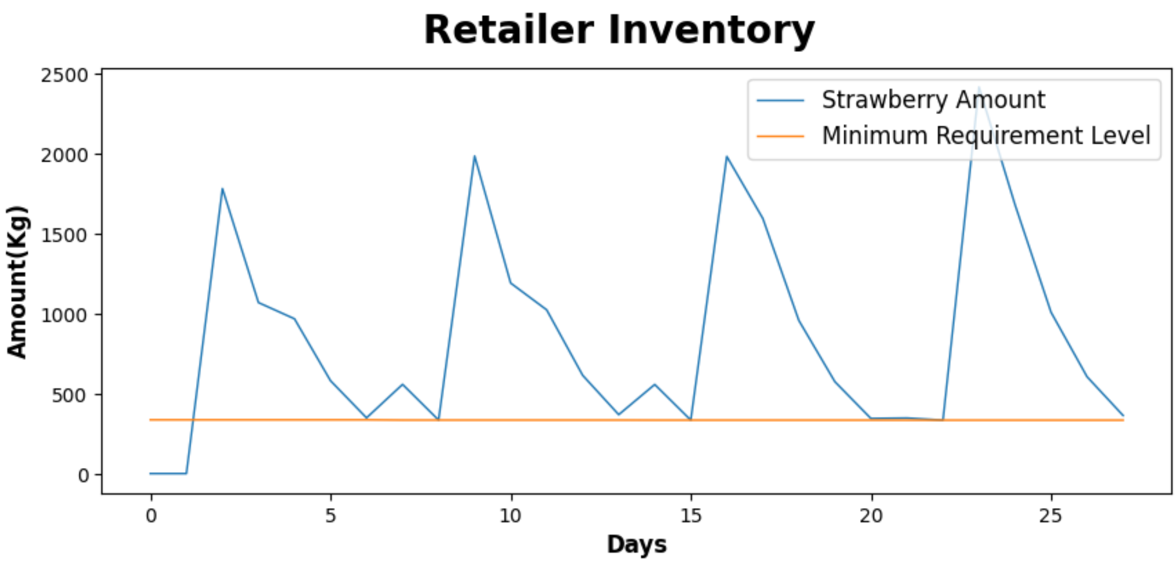

Figure 5.27: The amount of strawberries in the retailer's inventory. The orange line indicates the minimum requirement.

Figure 5.27 illustrates the amount of strawberries in the retailer's inventory. The information of transportation date and delay from the producer can be seen. The amount level is always kept not lower than the minimum requirement indicated using the orange line.

The optimized contract demands $D_{\cdot}^{contract}$, contract price $P_{\cdot}^{contract}$, profit and other income or costs are listed in Table 5.5 and Table 5.6. It can also be observed that the shortage cost is also avoided by the optimizer.

Table 5.5: Optimized $D_{\cdot}^{contract}$ and $P_{\cdot}^{contract}$ values.

	$j = 0$	$j = 1$	$j = 2$	$j = 3$
$D_j^{contract}$	3029.09	2995.46	3025.68	3017.45
$P_j^{contract}$	7.50	7.65	7.57	7.58

Table 5.6: Values for the profit, incomes and costs of the optimised strategy.

Name	Total Profit	Income	Discount Income	Transportation Cost
Worth($)	87347.73	91467.58	4573.37	4227.44
Name	Shortage Cost	Deterioration Cost	Dispose Cost	
Worth($)	0	2232.05	1116.86	

5.4. Evaluation

In this section, the optimized profit is compared with a baseline model. However, there is no access to obtain the data from the reality to evaluate the model. Therefore, the baseline model is set as the model executing a strategy which is chosen intuitively by most of people – setting the contract demands $D_{\cdot}^{contract}$ equal to the predictions $D_{\cdot}^{prediction}$.

The optimized results using the same information from Table 5.4 are listed in the table Table 5.7. It can be calculated that the profit is increased for 7.69% by the optimizer.

Table 5.7: Values for the profit, incomes and costs of the optimised strategy.

Name	Total Profit	Income	Discount Income	Transportation Cost
Worth($)	80629.89	86111.38	4305.56	3676.79
Name	Shortage Cost	Deterioration Cost	Dispose Cost	
Worth($)	0	3600.65	1254.80	

The optimizer is applied for 6 months for the evaluation of the model performance. The comparison between the profit by the optimized strategy and the standard strategy is shown in Figure 5.28 and Figure 5.29.

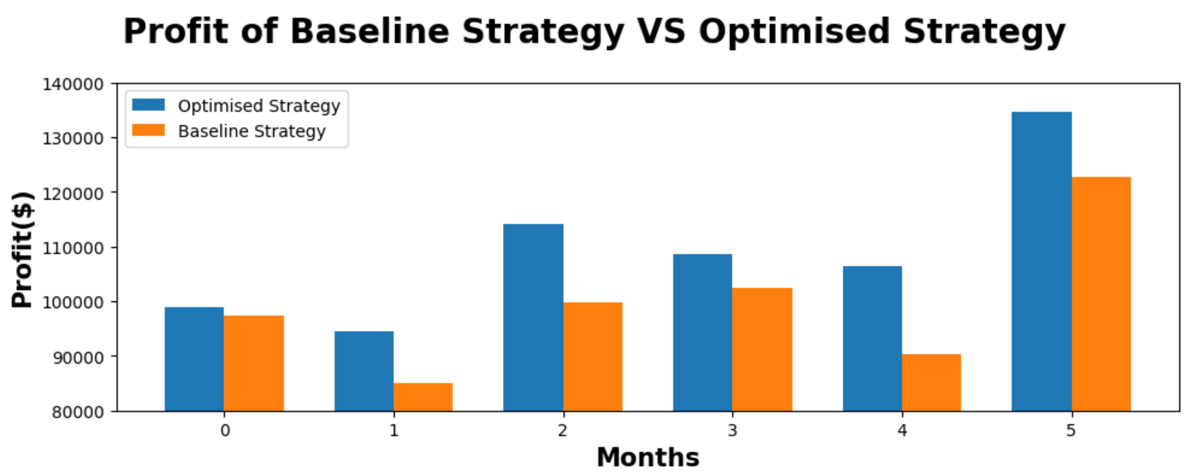

Figure 5.28: The comparison between the improved profit and the baseline profit. It can be seen that the optimized strategy is able to improve the profit stably.

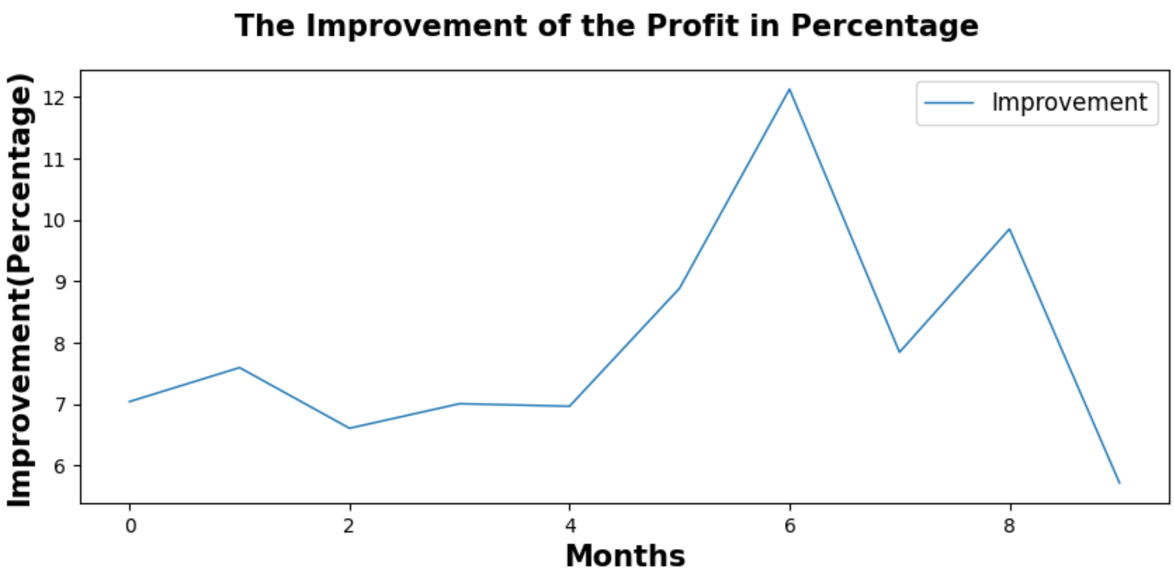

Figure 5.29: The percentage of profit improvement in each month. An average of 8% can be observed from the figure.

From Figure 5.28 it can be seen that all the profit from these 10 months are improved considerably. Figure 5.29 illustrates how much improvement is brought by the optimizer exactly, and an average of 7.96% of improvement can be calculated out.

5.5. Robust Optimization

In this section, how the robust optimization can help the model behave better under uncertainties from the fluctuation is studied.

5.5.1. Prediction Distribution

In order to have knowledge how much the yield can be in the target month, an ensemble of 50 LSTMs are applied to obtain the distribution. The results of the predictions are shown in Figure 5.30. It can be observed that after deleting the outsiders, the predictions are approximately under the normal distribution. Therefore, the integral of probability under normal distribution can be calculated to achieve the expectation of the profit.

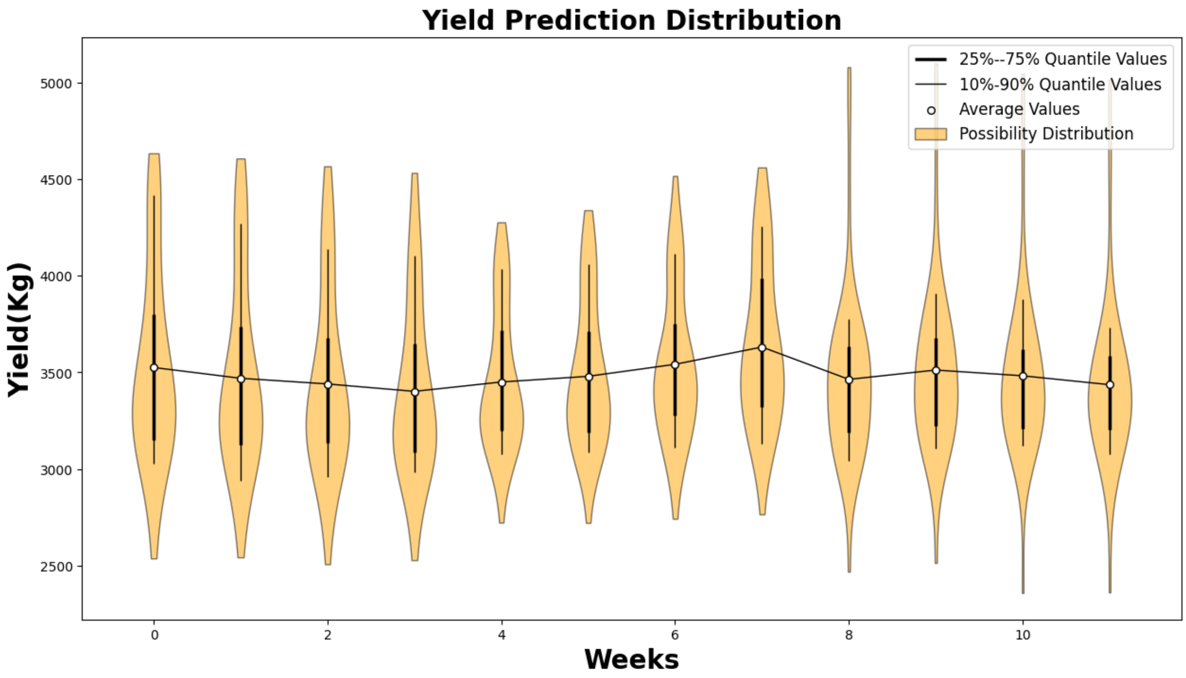

Figure 5.30: The distribution of the yield predictions made by 50 randomly initialized LSTMs. The distributions are close to Gaussian distribution after removing the outsiders.

5.5.2. Best Possible Profit

Figure 5.31 illustrates the optimized profits when the yield amount is equal to the minimum, average, maximum predictions and the true values. The profits optimized by the robust optimization model are also plotted for comparison. When being optimized using the robust optimization model, the yield is considered to be equal to the average yield prediction. It can be seen that the profit is proportional to the predicted yield amount. The profits optimized by the robust optimization model are a bit lower than the profits from the determined model as the robust optimization models are more conservative than the determined model.

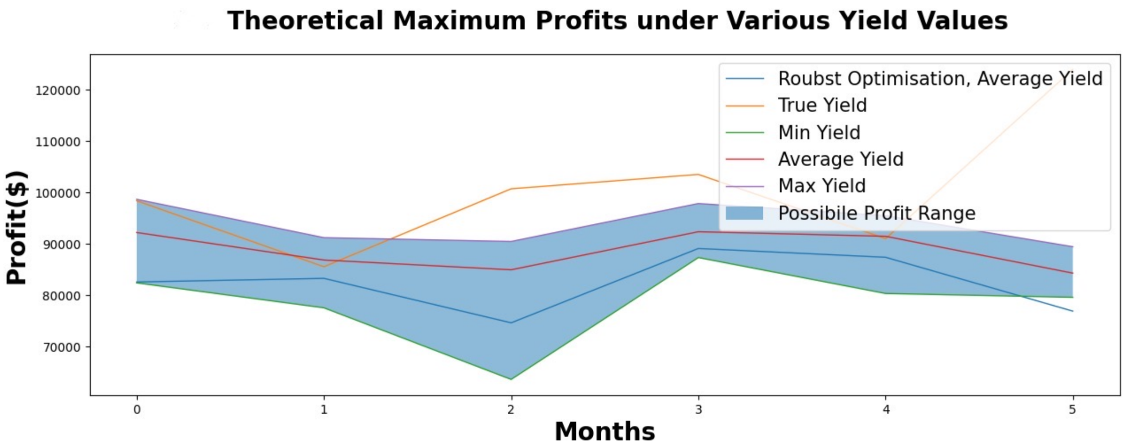

Figure 5.31: The theoretical maximum profits when the yield is equal to the minimum, average, maximum predictions and the true values. The profits optimized by the robust optimization model are also plotted for comparison.

5.5.3. Comparisons among Actual Profits

Figure 5.32 compares the profits when executing the contracts optimized by the model assuming yield is equal to the average predictions under various yield amount scenarios. It can be seen that the profits are still proportional to the yield, following a similar shape as Figure 5.31. Noticeably, the profits optimized by the robust optimization perform better than the determined model in 5 months out of 6. The comparison between the profit of the robust optimization model and the determined model is specifically compared and is shown in Figure 5.33.

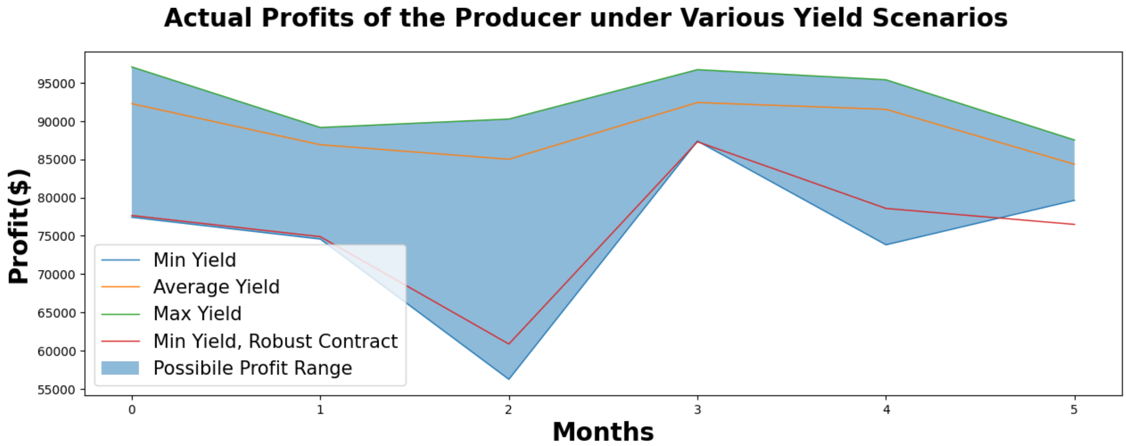

Figure 5.32: The comparison among the actual profits the producer can get under various yield situations after executing the contract optimized with assumption that the yield is equal to the average prediction.

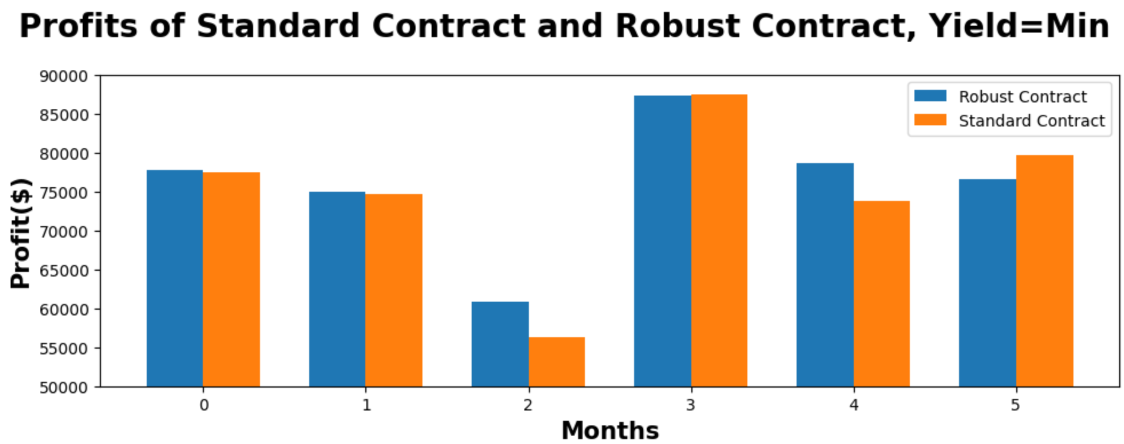

Figure 5.33: The specific comparison between the profit of the robust contract and the standard contract when the producer has a minimum yield.

The profit is improved by the robust optimization model by saving the shortage cost. The comparison of the shortage cost is shown in Figure 5.34.

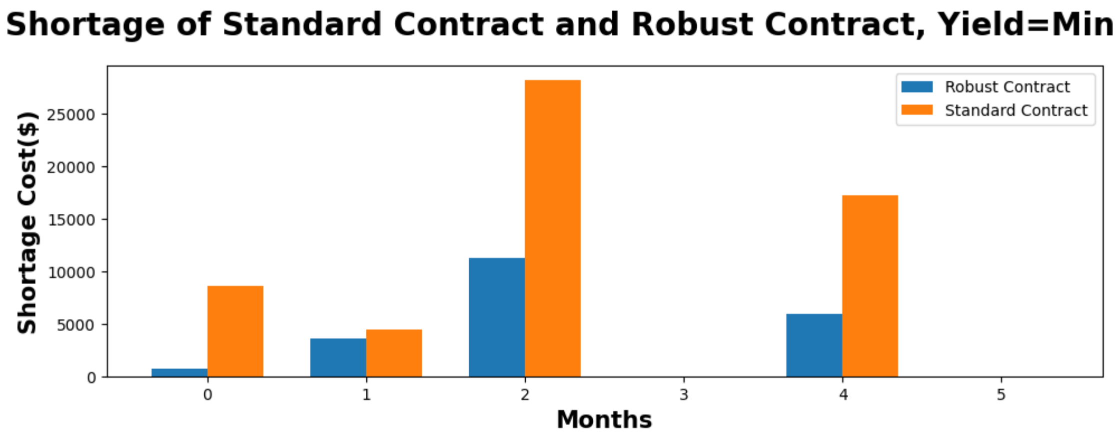

Figure 5.34: Comparison between the shortage cost in the standard and the robust contracts. It can be seen that the robust contract mitigates the shortage cost enormously.

It can be observed that among all the months the robust model suffers less from the shortage cost. Specifically, both of the models do not suffer from the shortage cost on the last month, which explains why the profit given by the robust model is less than the profit of the determined model on the last month shown in Figure 5.33.

Figure 5.35 illustrates how much profit is improved by the robust optimization model compared with the determined model in percentage. An average of 1.69% of profit improvement can be calculated out under this scenario.

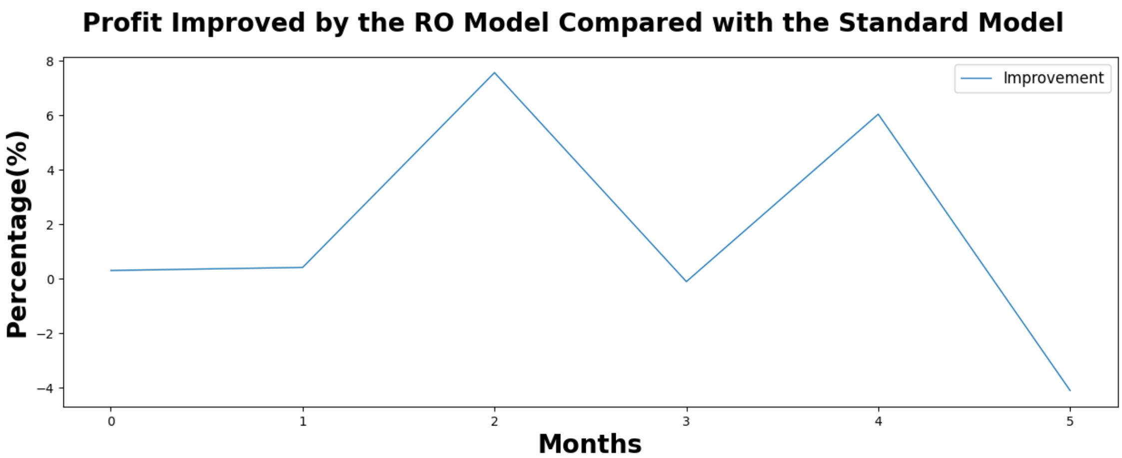

Figure 5.35: The profit improvement by the robust optimization model compared with the determined model in percentage.

5.5.4. Influence of Γ

How the value of Γ in the robust model affects the profit is studied in this section. The parameters affect the final profit by adjusting the conservation of the model. The worst-case yield value is decided by the Γ.

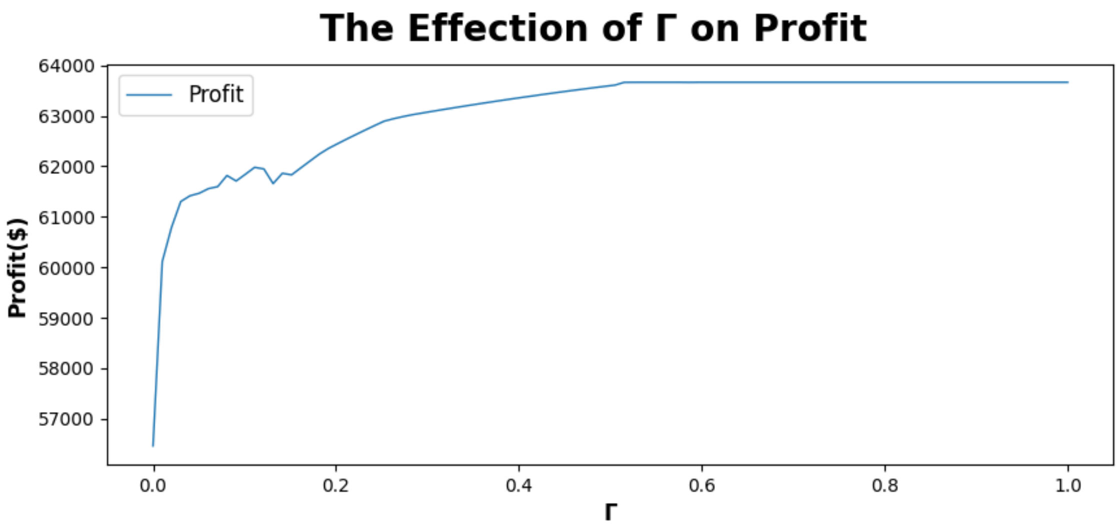

Figure 5.36: The profit increases with Γ if there is shortage cost.

Figure 5.36 shows the effect of Γ over the final profit when there is shortage cost on that month. As Γ increasing, the profit increases as well, because original shortage cost is covered by the increasing conservation controlled by Γ. However, when there is no shortage cost, as shown in Figure 5.37, the increase in Γ only decreases the profit. The reason is that there is no shortage cost needs to be covered and the increased conservation results in less profit.

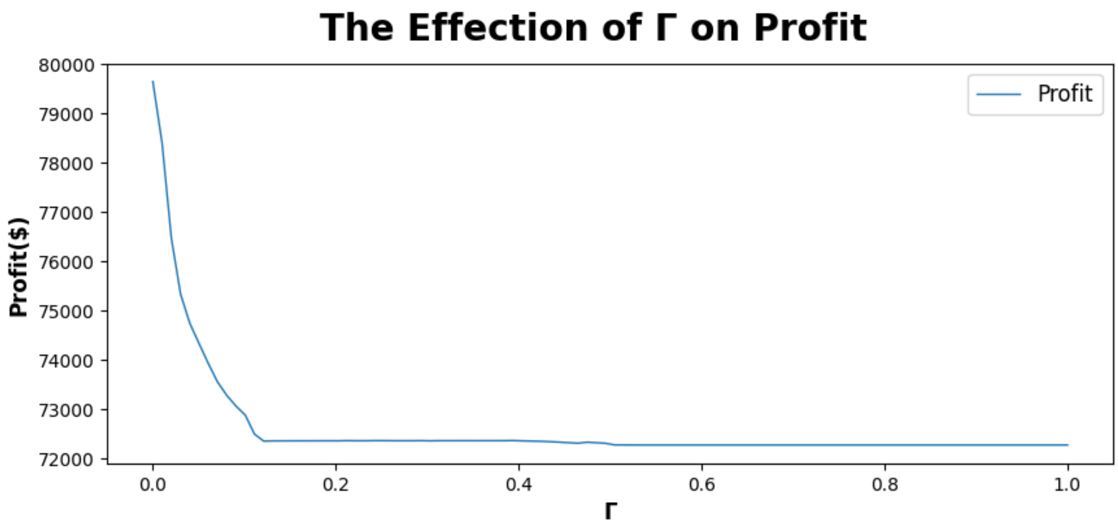

Figure 5.37: When there is no shortage cost, the profit decreases as the Γ increasing, raising the conservatism of the model.

5.5.5. Expectation of Profit

To study which value of Γ gives the best profit, the expectation of profit changing with Γ is studied. The multiplication between integral of probability density function until the worst case considered and the profit is calculated as the expectation. The expectation is shown in Figure 5.38, and it can be observed that under this scenario setting Γ as

1 gives the best profit expectation. It is because that the shortage cost is huge in this case and any risk of shortage is unworthy.

Figure 5.38: The expectation of profit changing with Γ. It can be seen that setting Γ as 1 promises the best performance.

5.6. Sensitivity Analysis

Finally, the model's sensitivity to the parameters is studied to analyze the importance of different parameters. To analyze the sensitivity, values of $D_{\cdot}^{prediction}$, $P_{\cdot}^{prediction}$, $Y_{\cdot}^{prediction}$ and parameters are set as Table 5.4 and Table 5.2.

It is found that the profit changes linearly with most of the parameters and change non-linearly with Cap_{truck}. Parameters having a linear sensitivity are listed in Table 5.8. The sensitivity of Cap_{truck} around the value in Table 5.2 is also listed in the table, with Figure 5.39 illustrating the sensitivity shape more clearly.

Table 5.8: Parameters sensitivity.

Name	Z_t	Z_h	Z_{truck}	σ_i	σ_t	σ_r	T_t
Sensitivity	-0.93%	-0.98%	-0.67	49.13%	24.32%	8.79%	-2.42%
Name	$w_{discount}$	$w_{shortage}$	w_{more}	w_{less}	Cap_{inv}	Cap_{trucks}	$limit$
Sensitivity	7.45%	-1.39%	-0.77%	72.40%	1.99%	9.02%	3.15%

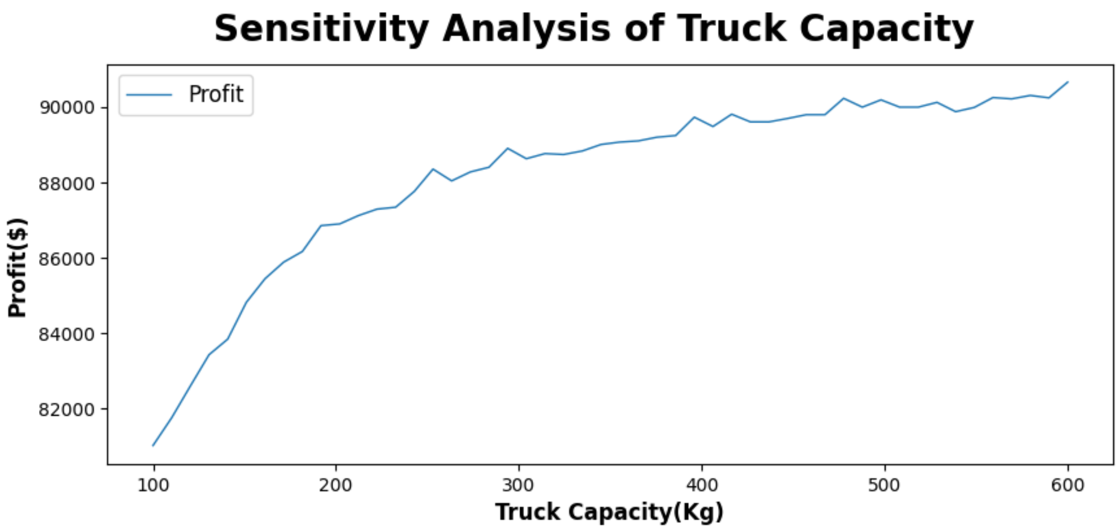

Figure 5.39: The sensitivity of the truck capacity Cap_{trucks}. The higher the capacity is, the less trucks are needed. Therefore, the transportation fee is reduced and the profit is increased.

Figure 5.39 shows the sensitivity of the truck capacity Cap_{trucks}. If the capacity is huge, the model is not sensitive to this parameter as only $1-2$ trucks are enough to deliver all the strawberries. There are steps can be observed in the shape, which is resulted by the fact that the number of trucks has to be integral.

From the results it can be concluded that the model is sensitive to the deterioration speed of the strawberries, especially when they are stored in the producer's inventory and the trucks. Therefore, devices slowing down the deterioration speed of the strawberries are expected to tremendously reduce the cost. It is found that w_{less}, which indicates how the price increases if the contract demand is lower than the prediction, also affects the profit dramatically. However, in the reality the parameter is set to a reasonable constant value and will not cause serious troubles.

6 Further Improvement

This chapter introduces some possible improvements to improve the performance of the model.

6.1. Sufficient Training Data

It is impossible to obtain weekly or monthly training data for strawberry demand and yield without the assistance from the third party agencies. The loss of the LSTMs can be reduced further as long as sufficient data is obtained to support more complex LSTMs neural networks.

6.2. Case Study

Because of a similar problem met in Section 6.1, it is not possible to evaluate the model using the real data. Therefore, data from the related companies are expected to further verify the reliability of the model. For now the model is only evaluated using the strategy that most people would choose by their intuition.

Similarly, the parameters for the models can also be set to be close to the reality. The parameters for the model are currently set to the state from which more interesting trade-offs can be seen.

6.3. Stochastic Optimization

Compare with robust optimization, which only considers about the worst cases, the stochastic utilizes the distribution of the uncertainty more sufficiently. The profit expectation can be drawn in this way, also making the result meaningful and providing the model with robustness.

6.4. Multi-agent Model

The model for now only considers about the profit of the producer. However, the behaviors of the retailer can also be considered about and the profit of the whole supply chain can be optimized in this case.

7 Conclusion

The thesis utilizes the prediction of the strawberry market and yield as guidance information supporting the MILP model to make decisions. As a result, it can be seen that the optimizer successfully raises the profit of the producer around 8% as demonstrated in Section 5.4.

The robust optimization helps the producers suffer less from the shortage cost induced by the fluctuation of the yield prediction by 1.7%. However, the robust optimization can play a more important role if the shortage does happen. Otherwise, it causes loss in profit because of its conservatism.

Sensitivity analysis is conducted and it is found that the profit is quite sensitive to the deterioration speed of strawberries when being kept in the inventories or in the trucks. Therefore, it is recommended that the producer and the retailer should use better cooling devices to prevent the strawberries from perishing.

The LSTMs in the whole model play an important role in predicting the strawberries market information and the yield, helping the optimizer make decision. The uncertainty distribution drawn from the predictions of the ensembles of LSTMs is also essential in terms of the robust optimization.

To improve the work, more reliable data is expected to train the ensembles of LSTMs for more accurate predictions. A case study is also needed to set the parameters as the reality and to analyse the performance of the model in the reality. Stochastic optimization can also be applied to replace the robust optimization part. Finally, the behaviors of the retailer can be considered about making the model a multi-agent model. In this case, the whole supply chain can be optimized and more complex trade-offs can be explored.

Bibliography

[1] J. Gustavsson, C. Cederberg, U. Sonesson, R. Van Otterdijk, and A. Meybeck, "Global food losses and food waste," 2011.

[2] L. S. Kantor, K. Lipton, A. Manchester, and V. Oliveira, "Estimating and addressing america's food losses," *Food Review/National Food Review*, vol. 20, no. 1482-2016-121447, pp. 2–12, 1997.

[3] A. A. Kader, "Increasing food availability by reducing postharvest losses of fresh produce," in *V International Postharvest Symposium 682*, pp. 2169–2176, 2004.

[4] D. Gunders and J. Bloom, "Wasted: How america is losing up to 40 percent of its food from farm to fork to landfill," 2017.

[5] K. Kelly, R. Madden, J. P. Emond, and M. C. do Nascimento Nunes, "A novel approach to determine the impact level of each step along the supply chain on strawberry quality," *Postharvest Biology and Technology*, vol. 147, pp. 78–88, 2019.

[6] T.-D. Nguyen, T. Nguyen-Quang, U. Venkatadri, C. Diallo, and M. Adams, "Mathematical programming models for fresh fruit supply chain optimization: A review of the literature and emerging trends," *AgriEngineering*, vol. 3, no. 3, pp. 519–541, 2021.

[7] J. V. Caixeta-Filho, "Orange harvesting scheduling management: a case study," *Journal of the Operational Research Society*, vol. 57, no. 6, pp. 637–642, 2006.

[8] C. D. Rocco and R. Morabito, "Production and logistics planning in the tomato processing industry: A conceptual scheme and mathematical model," *Computers and electronics in agriculture*, vol. 127, pp. 763–774, 2016.

[9] O. Ahumada and J. R. Villalobos, "A tactical model for planning the production and distribution of fresh produce," *Annals of Operations Research*, vol. 190, no. 1, pp. 339–358, 2011.

[10] K. H. Widodo, H. Nagasawa, K. Morizawa, and M. Ota, "A periodical flowering–harvesting model for delivering agricultural fresh products," *European Journal of Operational Research*, vol. 170, no. 1, pp. 24–43, 2006.

[11] Y. Y. Dubinin, "Integrated modeling of solutions in the system of distributing logistics of a fruit and vegetable cooperative," *Business: Theory and Practice*, vol. 15, no. 4, pp. 362–370, 2014.

[12] H. Grillo, M. Alemany, A. Ortiz, and V. S. Fuertes-Miquel, “Mathematical modelling of the order-promising process for fruit supply chains considering the perishability and subtypes of products,” *Applied Mathematical Modelling*, vol. 49, pp. 255–278, 2017.

[13] T. Itoh, H. Ishii, and T. Nanseki, “A model of crop planning under uncertainty in agricultural management,” *International Journal of Production Economics*, vol. 81, pp. 555–558, 2003.

[14] B. Kazaz, “Production planning under yield and demand uncertainty with yield-dependent cost and price,” *Manufacturing & Service Operations Management*, vol. 6, no. 3, pp. 209–224, 2004.

[15] A. Bezat-Jarzębowska and W. Rembisz, “Efficiency-focused economic modeling of competitiveness in the agri-food sector,” *Procedia-Social and Behavioral Sciences*, vol. 81, pp. 359–365, 2013.

[16] D. Bertsimas and M. Sim, “The price of robustness,” *Operations research*, vol. 52, no. 1, pp. 35–53, 2004.

[17] A. L. Soyster, “Convex programming with set-inclusive constraints and applications to inexact linear programming,” *Operations research*, vol. 21, no. 5, pp. 1154–1157, 1973.

[18] D. Bertsimas and A. Thiele, “A robust optimization approach to supply chain management,” in *Integer Programming and Combinatorial Optimization: 10th International IPCO Conference, New York, NY, USA, June 7-11, 2004. Proceedings 10*, pp. 86–100, Springer, 2004.

[19] Y. Bassok and R. Anupindi, “Analysis of supply contracts with commitments and flexibility,” *Naval Research Logistics (NRL)*, vol. 55, no. 5, pp. 459–477, 2008.

[20] Y. Wu, “Linear robust models for international logistics and inventory problems under uncertainty,” *International Journal of Computer Integrated Manufacturing*, vol. 24, no. 4, pp. 352–364, 2011.

[21] A. Ben-Tal, B. Golany, A. Nemirovski, and J.-P. Vial, “Retailer-supplier flexible commitments contracts: A robust optimization approach,” *Manufacturing & Service Operations Management*, vol. 7, no. 3, pp. 248–271, 2005.

[22] I. Paliari, A. Karanikola, and S. Kotsiantis, “A comparison of the optimized lstm, xgboost and arima in time series forecasting,” in *2021 12th International Conference on Information, Intelligence, Systems Applications (IISA)*, pp. 1–7, 2021.

[23] İ. Kırbaş, A. Sözen, A. D. Tuncer, and F. Ş. Kazancıoğlu, “Comparative analysis and forecasting of covid-19 cases in various european countries with arima, narnn and lstm approaches,” *Chaos, solitons & fractals*, vol. 138, p. 110015, 2020.

[24] S. Athiyarath, M. Paul, and S. Krishnaswamy, “A comparative study and analysis of time series forecasting techniques,” *SN Computer Science*, vol. 1, no. 3, p. 175, 2020.

[25] “Producer price index by commodity: Farm products: Strawberries.” https://fred.stlouisfed.org/series/WPU01110222. Accessed: 2023-02-03.

[26] “Average price: Eggs, grade a, large (cost per dozen) in u.s. city average.” https://fred.stlouisfed.org/series/APU0000708111. Accessed: 2023-04-24.

[27] “Average price: Tomatoes, field grown (cost per pound/453.6 grams) in u.s. city average.” https://fred.stlouisfed.org/series/APU0000712311. Accessed: 2023-04-24.