Research on Pricing and Replenishment Strategies for Vegetable Commodities Based on an Enhanced Nonlinear Objective Programming Model

Siying Wang[1], Yang Jiao[2], Yijia Ma[1, *]

[1]School of Economics and Management, Communication University of China, Beijing, China, 100024

[2]School of Computer and Cyber Sciences, Communication University of China, Beijing, China, 100024

*Corresponding author: marika0915@163.com

Abstract. With the increasing consumption of fresh food ingredients in the Chinese market and the perishable nature of vegetable commodities, which represent the largest proportion, optimizing procurement strategies plays a crucial role in enhancing the profitability of supermarkets. This has garnered significant attention from scholars and practitioners. Drawing on principles from microeconomics, this study employs multiple linear regression to describe the relationship between pricing and sales volume. Additionally, time series forecasting methods are utilized. By considering both supermarket profitability and social welfare maximization as optimization objectives, single-objective optimization and multi-objective nonlinear programming models are established. Various decision strategies for supermarkets under different scenarios are discussed, and a satisfactory decision is obtained by solving the models using differential evolution algorithm. Through research on the expected conditions and data provided by real supermarkets, this paper provides guidance for intelligent supermarket procurement strategies and supply chain management in practical settings.

Keywords: Multiple Linear Regression, Time Series Forecasting, Objective Optimization, Nonlinear Objective Programming, Differential Evolution Algorithm.

1. Introduction

In recent years, with the rapid development of the Chinese economy and the continuous improvement of residents' income levels, the demand for fresh food ingredients by consumers has been steadily increasing. According to data from the China Statistical Yearbook, the per capita consumption of fresh food ingredients in the country has risen from 232.0 kilograms in 2013 to 274.7 kilograms in 2021, with vegetables comprising the largest share [1]. Therefore, the stability of the supply in the vegetable market and the reasonableness of prices have always been a subject of concern. Consumers are often most concerned about factors such as the variety of vegetable products available, the prices of individual items, and the freshness of vegetables [2]. Determining the optimal replenishment and pricing decisions for vegetable commodities to meet consumer demands and maximize profits has become a pressing issue for fresh food supermarkets. This holds significant importance in the context of the multi-billion-dollar fresh food industry.

The research related to procurement decisions for vegetable commodities has generated significant interest among scholars. Deterministic optimization methods, particularly linear programming and mixed integer programming, have been widely employed and favored by researchers [3]. Lotfi et al. proposed a mixed integer nonlinear objective programming model for replenishment planning and space allocation in retail supermarkets. This model considers various constraints such as budget, space, storage time, and replenishment frequency, while representing market demand through pricing and product display space [4]. Amorim et al. developed a multi-objective model integrating fresh product production and distribution operations aiming to reduce storage time. However, their model does not address consumer demand [5]. Moreover, their model does not specifically discuss the unique considerations for vegetables as a commodity or methods to handle fluctuations in specific

item demand. Tijun et al. introduced a dynamic programming model for agricultural income management which takes inventory into account and solves it using heuristic methods [6]. This model adequately considers the influence of freshness on vegetable prices but does not account for consumer preferences among different categories or the demand for a balanced diet. Feng approached the problem from the perspective of product spoilage and proposed a dynamic optimization model for maximizing total profit per unit time. However, similarly, the model did not discuss the differentiation between vegetable varieties [7]. Numerous domestic and international researchers have applied mathematical models like dynamic programming and integer programming to conduct in-depth studies on the perishability, storage solutions, and supply chain planning aspects of vegetables [8-10].

In summary, previous studies have predominantly focused on maximizing profits from a product inventory perspective, often considering only the influence of factors such as individual product price, demand, and freshness. This approach tends to overlook the interdependencies of complementary and substitute products on demand and pricing. Furthermore, the sole emphasis on profit maximization neglects the importance of establishing a favorable reputation for the business.

This research introduces a multivariate linear regression model to describe the relationship between product demand and prices, effectively resolving the issues of substitutability and complementarity. Subsequently, the basic model is expanded to establish a non-linear objective programming model, discussing multi-objective vegetable procurement strategies with a primary focus on maximizing social welfare. Additionally, this study addresses the specific characteristics of vegetable products by separately considering category-based and individual item-based decision-making scenarios, thereby broadening the applicability of the model and providing valuable guidance for vegetable retail procurement decision-making.

2. The objective planning model focused on maximizing profits

The data for this study were procured from the 2023 National College Student Mathematical Modeling Competition, Problem C, encompassing the sales data of a certain fresh produce supermarket over the past three years, which will be employed in subsequent model demonstrations.

To commence, it is essential to acknowledge that the planners of this supermarket are rational agents who, in their decision-making endeavors, are driven by the pursuit of self-profit maximization. During the procurement phase, the supermarket has the option to purchase either by vegetable category or individual vegetable item. Given that complexity in the supply chain often begets heightened managerial costs and increased risk, rational decision-makers exhibit a proclivity towards procuring by category to streamline the supply chain. Bearing in mind the perishable nature of fresh products, it is posited in this study that if various vegetable varieties remain unsold on a given day, they shall not be made available for sale on the subsequent day. It is assumed that past daily sales of various categories align with actual market demand, with the exclusion of unforeseen circumstances and other low-probability events that have not occurred in the preceding three years. Section 2.1 elucidates the decision strategy for replenishing stock for a single day, under the premise that profit maximization serves as the sole objective. Section 2.2 expounds upon the pricing strategy for the forthcoming day, while Section 2.3 presents the outcomes derived from the model's solution.

2.1. Replenishment Strategies for Various Categories

The supermarket determines the quantity of replenishment for each category based on their forecast of future demand. In an ideal context, we posit that the supermarket's replenishment quantity precisely aligns with their projected market demand. In accordance with the tenets of microeconomics, commodity prices influence commodity demand. Moreover, recognizing that different categories of vegetables may either complement or substitute for one another, the demand for specific vegetable categories is also influenced by the prices of related categories [11]. To comprehensively account for the interplay among commodities, this study establishes a multivariate linear regression model between demand and prices to assist in replenishment decision-making, which is formulated as

$$Q_i = a_1P_1 + a_2P_2 + ... + a_jP_j + b_i, i \le j, \quad (1)$$

Where the variable Q_i represents the total demand for the ith category of vegetables, while P_1, P_2, ..., P_j denote the pricing for categories 1, 2, …, j. The coefficients a_1, a_2, ..., a_j represent regression coefficients, and b_i stands for the bias term. Given the extensive time frame and substantial sample size in this particular case, the data adheres to a multivariate normal distribution. Furthermore, each day's observations are independent. Hence, it is entirely reasonable to establish a multivariate linear regression model using this dataset.

For this case, we employed the daily sales volume and pricing of various vegetable categories over the past three years as the training set. Initially, we computed the average daily pricing for each category, weighted by the sales volume of each transaction. Subsequently, we employed SPSS software to calculate a multivariate regression model for six categories, illustrating the model parameters, the coefficient of determination, and statistical significance, as presented in Table 1:

Table.1. Multivariate Regression Model Results

	a_1	a_2	a_3	a_4	a_5	a_6	b_i	R^2	Sig.
Q_1	-2.231	-1.702	0.21	0.964	1.454	-1.149	55.749	0.173	0.000
Q_2	-3.116	-13.419	1.585	-1.413	-2.765	4.176	288.738	0.078	0.000
Q_3	-0.541	-0.132	-2.26	-0.433	0.069	0.96	62.194	0.098	0.000
Q_4	1.276	2.185	-0.621	-0.456	0.344	1.234	-9.002	0.323	0.000
Q_5	-0.487	-0.997	-0.51	4.247	-2.849	-2.764	103.726	0.111	0.000
Q_6	-2.186	-4.019	-0.838	1.07	2.7	-5.066	134.728	0.154	0.000

While individual models may exhibit relatively low R-squared values, the significance levels for all six models are below 0.01, indicating a significant relationship between the independent and dependent variables. Therefore, it remains justifiable to predict sales using pricing for each product category. Based on the aforementioned analysis, the supermarket must first determine the pricing for each product category, subsequently influencing the replenishment quantities, which aligns with real-world scenarios.

2.2. Pricing Strategies for Each Product Category

When determining pricing, the supermarket aims to maximize total profit. In the case of j different vegetable categories, the profit function can be represented as

$$w = \sum_{i=1}^{j} (S_iP_i - C_iQ_i), \quad (2)$$

Where w represents the total profit, S_i is the sales volume for category i, and C_i is the unit procurement cost for category i.

Assuming, in an ideal scenario, that the supermarket accurately forecasts consumer demand i. e., $S_i = Q_i$, and the profit function can be transformed into $w = \sum_{i=1}^{j} Q_i(P_i - C_i)$. And through algebraic substitution, we obtain $w = \sum_{i=1}^{j} (F_i(P_1, P_2, ..., P_j) * (P_i - C_i))$, where $F_i(P_1, P_2, ..., P_j)$ is the multiplex linear regression mapping function of the price of all categories (P_1, P_2,..., P_j) obtained in Section 2.1. We consider that future procurement costs can be derived from a time series of past costs (as discussed in Section 2.3), which implies that profit is solely dependent on the decision variables P_1, P_2,..., P_3.

To ensure more reasonable pricing and replenishment strategies, we establish the following constraints. Firstly, the procurement quantity for each product category (i.e., predicted demand) must be greater than or equal to 0 and less than or equal to the maximum of past sales volumes.Secondly,

the pricing for each product category must fall within the range of past minimum and maximum prices. Given these constraints, we formulate the objective optimization model as

$$\max w = \sum_{i=1}^{j} (F_i(P_1, P_2, ..., P_j) * (P_i - C_i)), \tag{3}$$

$$\text{S.t.} \begin{cases} 0 \le \mathrm{F}_i(P_1, P_2, .., P_j) \le \max(\{S_i^t\}) \\ \min(\{P_i^t\}) \le P_i \le \max(\{P_i^t\}) \\ i, j = 1,2,3,... \end{cases},$$

Where $\{S_i^t\}$ represents the collection of daily sales of category i in the past three years, and $\{P_i^t\}$ represents the collection of daily pricing of category i in the past three years. The supermarket needs to find the best possible pricing strategy to maximize the objective function.

2.3. Model Solving

First, we employ the time series ARIMA model to forecast the procurement costs for each category for the next day in this case. This approach involves transforming the data into stationary data through differencing, and then regressing the dependent variable on its lagged values and random error terms. The key steps are outlined in Equation as

$$C^t = \sum_{i=1}^{p} \varphi_i C^{t-i} + \varepsilon_t - \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}, \tag{4}$$

Where φ_i represents the autoregressive coefficients, p is the autoregressive order, θ_i represents the moving average coefficients, q is the moving average order, and $\{\varepsilon_t\}$ is the white noise sequence.

The forecasted procurement costs for each category for the next day are presented in Table 2.

Table.2. The Forecasted Procurement Costs For Each Category

C_1	C_2	C_3	C_4	C_5	C_6
13.35	3.28	7.85	6.96	4.43	4.34

Subsequently, utilize the Differential Evolution algorithm to search for satisfactory solutions for pricing decisions. Differential Evolution is a heuristic optimization algorithm used to solve objective optimization problems. It exhibits strong global search capabilities, independence from derivatives, and robustness when dealing with a large number of decision variables. The fundamental steps of this algorithm are illustrated in Figure 1.

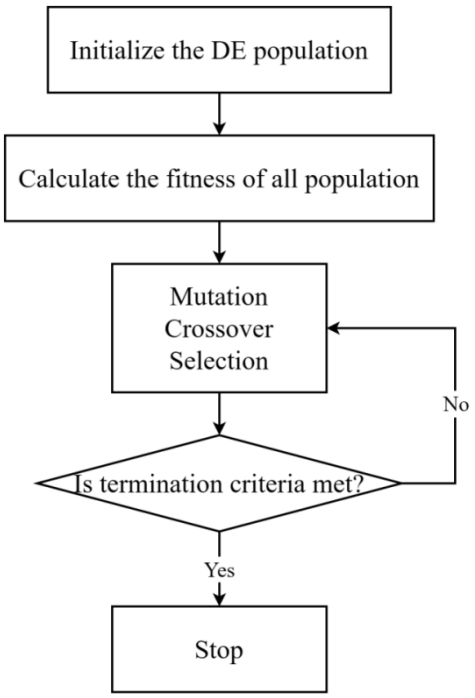

Figure 1. Steps of Differential Evolution Algorithm

We will use the scipy library in Python 3.8 to implement the differential evolution algorithm. We will set the population size to 50, the number of iterations to 100, the crossover probability to 0.4, and the mutation probability to 0.05. The initial pricing for each category will be randomly assigned between its minimum and maximum values. We will apply this algorithm under constraints to obtain a set of satisfactory solutions for the pricing, restocking strategies, and final profits of each vegetable category for the next day. The results are presented in Table 3:

Table.3. Pricing and Restocking Strategies for Each Category

Category	1	2	3	4	5	6	w
P_i	21.85	6.37	13.06	8.19	7.97	4.34	2195.56
Q_i	49.52	286.92	44.32	41.08	171.00	99.12	

Note that the solutions obtained using the differential evolution algorithm may be local optima rather than the optimal strategy. Supermarkets can make adjustments based on real-world conditions.

3. Model Extension: Considering Social Welfare in Multi-objective Optimization

Next, we will explore the case where maximizing social welfare is the decision objective. In the decision-making process for supermarkets, it is essential to consider various factors beyond their own economic interests, such as the overall social well-being. This approach not only shapes a positive corporate image that actively takes social responsibility but also helps increase market share and consumer loyalty.

Social welfare is typically composed of the benefits to supermarkets and consumer utility. In this case, we consider that consumer utility depends on the extent to which supermarkets satisfy diverse consumer demands. This means that supermarkets need to consider the actual demand for specific individual items, rather than only making restocking decisions based on categories. Supermarkets also need to weigh the importance of maximizing their own benefits and adequately meeting consumer demands. Therefore, the single-objective optimization problem described above can be transformed into a multi-objective optimization problem with two objectives. We will now modify

the model based on the real-world context, establish a goal programming model, and solve it to obtain a reasonable decision-making plan for the next day.

3.1. Modification of the Model Based on the Real-World Context

In order to simulate real-world situations more realistically and make the model more generalizable and practically valuable, we mainly consider the following scenarios in this study:

(1) Vegetables experience losses from purchase to sale, and the loss rates for each individual item are obtained from past experience statistics.

(2) In order to better manage shelf space, the minimum display quantity for each individual item should not be lower than a specified value.

(3) The total demand for each category can be predicted based on past sales data, and the sales volume of individual items is influenced by pricing. However, the specific actual demand quantity is unknown.

(4) In order to build a loyal customer base, while supermarket aims to satisfy consumers' demands for each category as much as possible, they also pursue maximum profits.

3.2. Goal Programming Model

Based on the modified model, we first update the profit expression as

$$w = \sum_{i=1}^{n} \sum_{j=1}^{mi} (P_{ij} \mathrm{S}_{ij} - C_{ij} I_{ij}), \tag{5}$$

P_{ij} represents the price of individual item j in category i, S_{ij} represents the sales volume of item j in category i, C_{ij} represents the unit procurement cost of item j in category i, and I_{ij} represents the procurement quantity of item j in category i (referred to as item ij hereafter). Since supermarkets aim to ensure that the procurement quantity meets the demand without missing out on opportunity costs, while avoiding excessive stock that leads to losses, we only consider the case where goods can be sold out after rational analysis. Thus, the expression for S_{ij} is given by:

$$\mathrm{S}_{ij} = I_{ij} * (1 - r_{ij}) \tag{6}$$

The variable r_{ij} denotes the loss rate of product ij, which falls within the range of [0,1].

According to the theoretical framework outlined in section 2.1, the sales of product ij continue to be influenced by its own pricing, as well as the pricing of complementary and substitute products. Consequently, it is plausible to construct a multivariate linear regression model($S_{ij} = G_{ij}(P_{11}, P_{12}, ..., P_{nm})$) to articulate these associations. Given the complexity of real-world operations involving a multitude of variables, the modeling process necessitates consideration of factors such as multicollinearity and the significance of variables.

The primary objective is to strive for consumer satisfaction across various product categories while avoiding deficits stemming from excessive replenishment—ensuring that sales precisely align with demand serves as the foremost goal. The secondary objective entails maximizing the profitability of the retail superstore. Additionally, absolute constraints are imposed, mandating that the minimum display quantity for each individual product should not fall below the stipulated value, denoted as 'k.' To safeguard the credibility of category demand forecasts, price fluctuations must be kept within reasonable bounds. This is achieved by confining the pricing of each individual product within the historical range of maximum and minimum values.In accordance with the above considerations, the objective programming model is formulated as

$$\min z = B_1 \sum_{i=1}^{n} (d_i^+ + d_i^-) + B_2 d_w^- \tag{7}$$

$$\text{S.t.}\begin{cases} \sum_{j=1}^{mi} S_{ij} + d_i^- - d_i^+ = Q_i \\ w + d_w^- - d_w^+ = M \\ S_{ij} \geq k \\ \min(\{P_{ij}^t\}) \leq P_{ij} \leq \max(\{P_{ij}^t\}) \\ P_{ij}, d_i^-, d_i^+, d_w^-, d_w^+ \geq 0 \\ i, j = 1,2,3,... \end{cases}$$

Where *minz* represents the directive from the decision-maker, urging a reduction in deviations from target values. This necessitates the adjustment of individual product pricing and replenishment quantities. B_1, B_2 serve as the priority factors, $B_1 \gg B_2 > 0$ ensuring the prioritized attainment of the primary goal. Positive deviation variables, d_i^+, d_w^+, denote the portion where decision values exceed the target value, while negative deviation variables, d_i^-, d_w^-, signify the portion where decision values fall short of the target. It is a fundamental principle that decision values cannot simultaneously exceed and fall short of the target, thus $d_i^+ * d_i^- = 0$, $d_w^+ * d_w^- = 0$ always holds true. *M* represents an infinitely positive constant, symbolizing the pursuit of maximizing profits. $\{P_{ij}^t\}$ denotes the set of daily pricing data for product ij over the preceding three years.

3.3. Model Solving

First, based on past experience, we predict the total demand for each category for the next day. Here, we use the SARIMAX method for forecasting, and the results are presented in Table 4. SARIMAX is a time series forecasting method that adds seasonality (S) and exogenous factors (X) to the differenced moving autoregressive integrated moving average (ARIMA) model. Compared to a regular ARIMA model, SARIMAX adequately considers the seasonal fluctuations in demand.

Table.4. Forecasted Total Demand for Each Category

Q1	Q2	Q3	Q4	Q5	Q6
19.423074	142.365901	21.893163	21.955683	86.270138	48.903092

Next, we determine the specific items that need replenishment. While ensuring that each category has at least one item, the supermarket prefers selecting those items with the highest profitability. Profitability is represented by the product of sales volume and profit margin, and the future profitability of each item is based on the total profitability in the past seven days. Following this principle, we ultimately select 27 items under the condition of k=2.5 as an example and solve for the optimal pricing and replenishment quantity for each item.

We use the ARIMA model to forecast the cost C_{ij} for each item i on the next day. We then establish a multivariate linear regression model between the actual sales volume of item ij and the prices of all items. In total, there are 27 model expressions. Taking item 11 as an example, we apply the stepwise regression method with a significance level of 0.05 for variable entry and a significance level of 0.1 for variable exclusion. After constructing the initial model, we eliminate variables with variance inflation factor (VIF) greater than 10 and select independent variables with eigenvalues greater than 1. The resulting multivariate linear regression model between sales volume and price is given by:

$$S_{11} = 0.054P_{11} + 0.009P_{21} + 0.015P_{26} - 0.013P_{61} + 0.004 \quad (8)$$

The model significance is 0.000, and the R-squared value is 0.887, indicating that the selected independent variables effectively explain the dependent variable. The standardized coefficients of the four independent variables are 0.866, 0.153, 0.134, and -0.169, respectively, indicating that the

pricing of an item itself has the greatest impact on sales volume, which is consistent with the real-world scenario. The remaining 26 models are derived in a similar manner.

Without loss of generality, let B_1=99999, B_2=1, and M=99999. After variable replacement, the pricing of the 27 items is considered as the optimization variable. The initial values are randomly assigned between the previous maximum and minimum values. Using the differential evolution method with a population size of 100, 1000 iterations, a crossover probability of 0.4, and a mutation probability of 0.05, we obtain a set of satisfactory solutions and their corresponding replenishment quantities as shown in Table 5:

Table.5. Pricing and Replenishment Strategies for Each Item

$Item_{ij}$	P_{ij}	I_{ij}	$Item_{ij}$	P_{ij}	I_{ij}
11	15.37	14.41	41	6.00	11.74
12	14.45	10.06	42	12.00	12.30
13	14.00	11.02	51	5.39	11.17
14	26.00	6.25	52	5.31	13.18
15	9.20	13.79	53	3.66	12.90
16	16.89	5.05	54	19.18	14.43
21	4.52	17.31	55	20.53	11.81
22	6.34	14.41	56	11.80	14.03
23	5.33	7.63	57	4.86	17.25
24	4.03	34.02	61	24.00	25.45
25	5.64	37.24	62	4.94	9.71
26	5.43	24.61	63	3.21	10.45
31	13.08	12.35	64	6.97	8.59
32	13.04	11.97			

Based on this strategy, the predicted profit for the next day is 998.27 yuan, which aligns with the actual profit performance of the supermarket in the dataset over the past seven days, demonstrating practical significance.

4. Conclusions

Fresh foodstuffs have evolved into indispensable commodities in the lives of consumers. The procurement determinations pertaining to vegetable commodities have likewise garnered considerable attention from an array of fresh produce retailers. In the quest to augment the existing knowledge framework governing procurement decisions and to furnish practitioners with intelligent procurement strategies, this study, grounded in the paradigm of goal programming, integrates multifaceted linear regression and time series forecasting methodologies. It conducts an exhaustive scrutiny of retailers' actual expectations and embarks upon the optimization of procurement strategies for vegetable commodities.

The research discerns that employing a multifaceted linear regression model to elucidate the relationship between vegetable product pricing and sales volumes possesses explicability. Retailers must, in consonance with actual circumstances, first determine product pricing and subsequently ascertain replenishment quantities. This corresponds to the empirical reality. With an eye towards retailers' expectations in practice, when the sole pursuit of profit maximization is in force, retailers are disposed to replenish vegetable categories as a whole to curtail management costs. The satisfactory solution yielded by the differential evolution method translates into a retailer's revenue amounting to 2195.56 yuan, surpassing the average revenue of the preceding week by a substantial margin of 221.28%. When retailers account for the diverse demands of consumers, they opt to procure individual vegetable items to enhance market share and consumer loyalty. Under these circumstances,

the decisional proceeds amount to 998.27 yuan, surpassing the average revenue of the preceding week by a notable 46.08%.

It is imperative to underscore that when considering societal welfare, retailers' proceeds fall short of those garnered when solely pursuing profit maximization. This indicates that decisions solely driven by the singular objective of economic profit may engender adverse ramifications on societal welfare. Lastly, the research findings also underscore that the decisional value predicated solely on profit maximization notably surpasses actual performance. This underscores that, in practical operations, retailers necessitate the holistic consideration of myriad factors in crafting judicious decisional strategies, transcending the mere pursuit of short-term economic profit.

References

[1] National Bureau of Statistics. China Statistical Yearbook 2022 [M]. Beijing: China Statistics Press, 2022.

[2] Juerkenbeck K S A .Importance of sensory quality signals in consumers ' food choice[J].Food Quality and Preference, 2021, 90(1).

[3] Nguyen T D , Nguyenquang T , Venkatadri U ,et al.Mathematical Programming Models for Fresh Fruit Supply Chain Optimization: A Review of the Literature and Emerging Trends[J]. 2021.

[4] Lotfi M M , Rabbani M , Ghaderi S F .A weighted goal programming approach for replenishment planning and space allocation in a supermarket[J].Journal of the Operational Research Society, 2011, 62(6):1128-1137.

[5] Amorim P , Almada-Lobo B .Multi-objective integrated production and distribution planning of perishable products[J].International Journal of Production Economics, 2012, 138(1):89-101.

[6] Fan T, Xu C, Tao F. Dynamic pricing and replenishment policy for fresh produce[J].Computers & Industrial Engineering, 2020, 139(Jan.):106127.1-106127.14.

[7] Feng L .Dynamic pricing, quality investment, and replenishment model for perishable items[J].John Wiley & Sons, Ltd, 2019(4).

[8] Ji Y , Du J , Han X ,et al.A mixed integer robust programming model for two-echelon inventory routing problem of perishable products[J].Physica A: Statistical Mechanics and its Applications, 2020, 548.

[9] Nguyen T D , Venkatadri U , Nguyen-Quang T ,et al.Optimization Model for Fresh Fruit Supply Chains: Case-Study of Dragon Fruit in Vietnam[J].AgriEngineering, 2019, 2(1):1-26.

[10] Gokarn S , Kuthambalayan T S .Creating sustainable fresh produce supply chains by managing uncertainties[J].Journal of Cleaner Production, 2019, 207(PT.1-1180):908-919.

[11] Chen C H , Zhang H , Migueis M ,et al.Principles of Microeconomics[M].Dryden Press, 2007.