Fresh Agricultural Products Cold Chain Location Selection in Context of Big Data

Lili Weng

Department of Management, Guangdong University of Foreign Studies South China Business College, China
Email: 893527419@qq.com

Abstract. In the domestic fresh market, agricultural products are one of the major consumer goods. The rational optimization of the fresh agricultural products cold chain distribution center location is of great significance in enhancing the efficiency of the entire cold chain logistics. Therefore, in the context of big data, this paper chooses the grey forecasting method with a higher prediction accuracy, applied GM(1,1) prediction model for fresh agricultural products demand forecast under the premise of comparing different forecast methods’ accuracy; Basing on the demand forecast, this paper formulates fresh agricultural products distribution center location model with the minimum total cost objective. In this objective, the thesis considers the cost of cargo damage and the penalty cost of violating the time window in order to improve customer satisfaction and loyalty, which is considered by a few researchers. Finally, a specific case study is designed for Q enterprise to solve the optimal problem and test the validity of the model algorithm.

1. Introduction

In the domestic fresh market, agricultural products are part of the major consumer products. Residents pay more attention to the quality of the agricultural products they buy. In 2019, China’s cold chain logistics market size was 367.3 billion Yuan, the total logistics of the whole society reached 304 trillion Yuan. It can be observed that the fresh farm products cold chain logistics market growth is huge.

However, the unreasonable choice of nodes in the cold chain distribution center has increasingly become one of the main factors affecting the progress of the cold chain logistics industry. The rational optimization of the fresh farm products cold chain distribution center location plays an important part in improving the efficiency of the cold chain logistics system of fresh agricultural products [1].

Furthermore, compared with traditional logistics, logistics in the context of big data pays more attention to consumer demand characteristics. And the consumer pays more attention to the quality and security. However, fresh farm products have the character of perishableness, a short expiration date. Therefore, they must be placed in the cold chain, in order to ensure their quality and nutrition. The cold chain distribution center for fresh agricultural products is the core of the fresh farm product supply chain. Its scientific location is conducive to improving system efficiency, thereby shortening the delivery time of fresh agricultural products and ensuring the quality of fresh agricultural products [2].

Therefore, based on the determining multiple demand points, this paper predicts their demand, and sets some demand points as distribution centers to meet the demand constraints of their related
demand points, and then constitutes a series of distribution service areas. On this basis, it seeks to minimize the cost of the distribution system.

2. Model and Algorithm

It is the NP problem that the location of logistics distribution center of cold chain distribution center of fresh agricultural products. With the rapid development of computers and big data, there has been more and more studies on the algorithm of distribution center location. Applying immune algorithm (IA) to the location selection problem of regional fresh agricultural product logistics distribution center can effectively improve the location selection efficiency [3]. At the same time, in the context of big data, people pay more attention to customer satisfaction, such as the rate of the damaged cargo, distribution speed, etc. Therefore, the establishment of this model considers the rate of damaged cargo. Moreover, this model also combines historical data to predict the demand with the help of big data technology, which will be more accurate in the location study of the distribution center.

2.1. Selection of Forecasting Methods for Demand of Fresh Agricultural Products

In the establishment of a cold chain distribution center for fresh agricultural products, the demand for fresh agricultural products at each demand point is a very important parameter, which may directly have an influence on the final location of the cold chain distribution center. The prediction methods can be subdivided into qualitative prediction methods and quantitative prediction methods commonly.

Because the qualitative forecasting method has its one-sidedness and low accuracy, the quantitative forecasting method is employed to the demand. Commonly, methods for demand forecasting include moving average forecasting method, grey forecasting method, regression forecasting, exponential smoothing forecasting, and artificial neural network forecasting [2, 4]. The artificial neural network prediction method requires more sample data to train and test. It will produce a large deviation when the number of samples is not sufficient. Therefore, when the sample data are too little, the artificial neural network prediction method is generally not considered for prediction. Because these analysis methods have different prerequisites and emphases, as well as have its own advantages and disadvantages in actual application, which is manifested by the large deviation between the predicted results and the actual demand values.

To enhance the validity of prediction, firstly this article select demand point 1’s the demand for fresh agricultural products from January 2019 to January 2020 and take the January to November 2019 demand data as the original data. Secondly, the method including the moving average forecasting method, regression forecasting method, exponential smoothing forecasting method, and artificial neural network forecasting is used to simulate the forecast of the demand in December 2019 and January 2020. Thirdly, it compares the demand value predicted by each forecast method with the actual significance and obtains the absolute value of the error of each predicting method as shown in table 1.

| Month     | Gray  | Regression | Exponential smoothing | Moving average |
|-----------|-------|------------|-----------------------|----------------|
| December 2019 | 2.88% | 6.80%      | 3.12%                 | 5.78%          |
| January 2020 | 12.14% | 16.21%     | 13.08%                | 18.23%         |
| Mean absolute error | 7.51% | 11.51%     | 8.10%                 | 12.01%         |

It can be seen from the error comparison that in the forecast of the demand for fresh products in a certain area, the average absolute error of the gray forecasting method is 7.51%, the prediction error is the smallest, and the prediction accuracy is the highest; The exponential smoothing forecasting method is next; the average absolute error of the moving average forecasting method is 12.01%, the prediction error is the largest, and the prediction accuracy is relatively poor. The main reason for the poor forecast results in January 2020 is that January is exactly the Chinese New Year, so the forecast value deviation will be larger.
Therefore, GM(1,1) with the highest demand forecasting accuracy is selected as the forecasting method for demand in this paper.

2.2. The Assumption of Distribution Center Location Problem
To build the model, the following assumptions are proposed [5]:
- The distribution capacity carried by the distribution center can always meet the total demand of its service demand point.
- Within the radiation range of the distribution center, the demand point can only be served by the corresponding distribution center.
- Only the transportation cost between the distribution center and the demand points it serves as the model’s service cost.
- The operating cost only considers the sum of the daily cost of the distribution center and the depreciation cost of vehicle use, and the remaining costs are not considered in this article.
- Taking cost reduction as the optimization goal, according to the principles of industrial engineering standardization, the operating expenses can be standardized and limited, so as to achieve effective cost control.
- The loss of fresh agricultural products during storage and handling is not considered.

2.3. Model Construction of Distribution Center Location Problem
Based on the above assumptions and references to relevant literature, the site selection factors considered in this article are relatively simple, but there is no lack of representativeness and versatility, there are four main parts: (1) Service charge between the distribution center and the demand points it serves; (2) Operating expenses of distribution center; (3) Cost of cargo damage; (4) The penalty cost for violating the time window. Establish the following location-allocation mathematical model:

Objective function:

\[
\min W = \sum_{i \in N} \sum_{j \in M_i} \omega_{d_{ij}} Z_{ij} + \sum_{i \in N} \sum_{j = 1}^{Z_{ij}} Z_{ij} X_j + \sum_{i \in N} \sum_{j \in M_i} \omega_{Z_{ij}} p(1 - e^{-\delta(t_i - t_j)}) + \phi(i) \tag{1}
\]

Restrictions:

\[
Z_{ij}, h_j \in \{0,1\} \quad Z_{ij} \leq h_j, i \in N \quad j \in M_i \tag{2}
\]

\[
\sum_{j \in M_i} Z_{ij} = 1 \quad i \in N \tag{3}
\]

\[
\sum_{j \in M_i} h_j = p \quad p \leq n \tag{4}
\]

\[
d_{ij} \leq s \tag{5}
\]

\[
N = \{1,2,\ldots,n\} \tag{6}
\]

\[
M_i \subseteq N \quad (7)
\]

\[
\phi(i) = \begin{cases} 
M & t_{ij} \leq EET_i \\
\mu_{i}(ET_i - t_{ij}), & EET_i \leq t_{ij} \leq ET_i \\
0, & ET_i \leq t_{ij} \leq LT_i \\
\mu_{i}(t_{ij} - LT_i), & LT_i \leq t_{ij} \leq ELT_i \\
M & t_{ij} \geq ELT_i 
\end{cases} \quad (8)
\]
Equation (1) is the total cost of all distribution centers and constitutes the objective function of the model [5, 6]. The first part represents the model’s service cost, the second part represents the model’s operating cost, the third part represents the model’s cargo damage cost, and the last part represents the model’s penalty cost for violating the time window [7-9].

- \( \omega_j \) is the demand at the demand point;
- \( d_{ij} \) is the distance between the logistics distribution center and the demand point;
- \( X_j \) is the operating expenses of the distribution center;
- \( p \) is the product price;
- \( \delta \) is the decay rate;
- \( t_{ij} \) is the time point of the refrigerated truck from the distribution center \( i \) to the demand point \( j \);
- \( t_i \) is the time when the refrigerated truck departs from the distribution center \( i \).

Equations (2) and (3) are the mathematical language of hypothesis (2), which ensures that each demand point has and can only be served by its corresponding distribution center, \( Z_{ij}, h_j \) are two 0-1 variables, when both are 1, the service relationship between the distribution center \( j \) and the demand point \( i \) is represented by \( Z_{ij} \), the demand point \( j \) is set to the logistics distribution center and is represented by \( h_j \), when both are 0, it means the opposite.

Equation (4) specifies that among \( n \) demand point sets, \( p \) distribution centers are selected.

Equations (5) indicates that the service scope of the distribution center can cover all the demand points it serves.

Equations (6) and (7) represent the demand points and the number of distribution centers, respectively.

Equation (8) represents the penalty cost constraint for violating the time window [6, 10].

Among them, \([ET, LT]\) represents the best service time window of delivery vehicles to the demand point \( j \), which is no additional penalty costs are required during this period; \([EET, ET]\) and \([LT, ELT]\) are the acceptable time windows of the demand point \( j \). At this stage, the delivery of the goods to the terminal retailer needs to pay a certain penalty cost for violating the time window constraints, \( \mu_1, \mu_2 \) are the penalty factors for violating the time window for early arrival and late arrival; \( M \) indicates that when the delivery early time and late time are too large, an infinite cost will be incurred due to violation of the time window constraint.

3. Introduction of IA for Location Problem of Distribution Center

Among many intelligent algorithms, the immune algorithm (IA) is generated relatively late. It is an algorithm proposed by T-Fukuda et al in 1998. IA is gradually used to solve polymorphic optimization problems [11]. Similar to the genetic algorithm imitating the genetic evolution law of the biological world, the immune algorithm is also inspired by the theory of the biological immune system. Inspired by the diversity of generation and maintenance mechanism of the human immune system, the group of the algorithm presents diversity. In this way, the old and difficult problem in the general optimization process-the "precocity" problem can be solved [3, 7]. Thus, when dealing with the multi-peak function optimization problem, this paper believes that the use of immune algorithm will get the ideal global best optimal solution.

4. Example Simulation

This article takes Q enterprise as an example to analyze the optimization needs of Q companies’ cold chain logistics layout. And it will analyze the location and path optimization of the cold chain distribution center of Q enterprises and give the final plan. The Q enterprise is a cold chain logistics
A company operating fresh products. The company needs to deliver fresh raw products to various demand centers before 9 o’clock every day. In this way, the morning peak in the city can be avoided, and the supply of fresh products in each demand center can be guaranteed. The case scale selected 6 from 20 demand points as logistics distribution centers.

First, according to the historical data of each demand point, use the gray prediction method to predict the demand of each demand point, and organize the node information of each demand point, including the latitude and longitude of each demand point and the service time window, the specific information is shown in Table 2 (due to the space limitations, only the data of 8 demand points are shown).

| ID | Longitude       | Forecast demand | Time window   | Acceptable time window |
|----|----------------|-----------------|---------------|------------------------|
| 1  | 113.809218,23.304102 | 1.5            | 5:30-7:30    | 5:00-8:00             |
| 2  | 113.627244,23.29754   | 0.5            | 5:00-7:00    | 4:30-7:00             |
| 3  | 113.610153,23.220344 | 1.0            | 6:00-7:00    | 5:30-7:30             |
| 4  | 113.495754,23.448393 | 1.5            | 5:00-8:00    | 5:00-8:30             |
| 5  | 113.221944,23.396941 | 2.0            | 4:50-6:30    | 4:30-8:30             |
| 6  | 113.291165,23.217944 | 2.0            | 6:30-7:30    | 6:30-8:00             |
| 7  | 113.200927,23.378581 | 1.6            | 5:30-7:00    | 5:00-7:00             |
| 8  | 113.629773,23.13791  | 1.0            | 6:00-7:30    | 5:30-7:30             |

According to the distribution center location model, the calculation example is solved according to the immune algorithm steps, the parameter size of the algorithm is 50, the memory capacity is 10, the number of iterations is 100, the crossover probability is 0.5, the mutation probability is 0.4, and the diversity evaluation parameter is 0.5, \( \delta \) is 0.02 yuan/(ton·min).

Based on the same conditions, the genetic algorithm (GA) is used to solve this model. The comparison chart is shown in Figure 1. It is obvious from the figure that the results obtained by using IA are more accurate than those obtained by GA. In addition, IA is superior to genetic algorithms in terms of convergence speed, operation speed, and optimization ability. So it was solved many times with MATLAB, among the many distribution center location plans, (20,16,1,4,10,18) is the optimal solution. Figures 2 and 3 are the IA algorithm convergence diagram and the logistics distribution center location plans respectively.

5. Conclusion
Big data technology is vital to the location selection of distribution centers. Big data mining and big data analysis technologies should serve the location selection of distribution centers, so as to obtain a
more scientific and long-term location selection. The location selection of the regional fresh agricultural product logistics distribution center studied in this paper takes the immune algorithm as the core, Matlab as the carrier, and big data technology as the auxiliary, thus forming a complete set of distribution center location selection scheme, provide a scientific and reasonable solution for the location selection of the distribution center of the fresh agricultural products industry.

Acknowledgments
Work at the school-level youth projects of research on influential factors and countermeasures of Guangdong Port logistics development under the Maritime Silk Road Initiative is supported by the Guangdong University of Foreign Studies South China under contracts 20-003C.

References
[1] Hua X, Hu X and Yuan W 2016 Research optimization on logistics distribution center location based on adaptive particle swarm algorithm Optik International Journal for Light & Electron Optics S0030-4026 16306544.
[2] Pei Y H 2016 The Study on Location of Cold Chain Distribution of Fresh Agricultural Products Based on Demand Forecast (Chongqing Technology and Business University) chapter 1 pp 1-2.
[3] Wang C Y and Liu Z N 2014 An immunity-based algorithm for distribution center location Advanced Materials Research 971-973 1537-1542.
[4] Mi W H, Yue X W, Shao J F and Qian W M 2018 Research on selection of logistics distribution center based on an adaptive immune algorithm Value Engineering.
[5] Shi F, Wang H and Yu L 2015 Analysis of 30 cases of MATLAB intelligent algorithm (Beijing University of Aeronautics and Astronautics Press) chapter 12 pp 118-128.
[6] Chen G H 2019 Prediction of the demand for cold chain logistics of fresh agricultural products in Henan based on the gray model Modern Trade Industry 040 (011) 56-58.
[7] Yi L 2019 Research on location routing problem for fresh agri-products cold chain distribution center based on bi-level programming Doctoral Dissertation.
[8] Amorim P and Almada-Lobo B 2014 The impact of food perishability issues in the vehicle routing problem Computers & Industrial Engineering 67 (Jan) 223-233.
[9] Ya-Ping L I 2013 Study on distribution routing optimization problems of cold chain logistics of fresh agricultural products Value Engineering.
[10] Wang Y, Assogba K, Liu Y, Ma X, Xu M and Wang Y 2018 Two-echelon location-routing optimization with time windows based on customer clustering Expert Systems with Application 104 (Aug) 244-260.
[11] Rong A, Akkerman R and Grunow M 2011 An optimization approach for managing fresh food quality throughout the supply chain International Journal of Production Economics 131 421-429.