

Research on cold chain logistics demand forecast based on LSTM model

Xin Feng *

School of Economics and Management, Southwest Petroleum University, Chengdu, China

* Corresponding author: (Email: 1778790683@qq.com)

Abstract: Cold chain logistics demand is affected by multiple factors and exhibits significant nonlinear, dynamic and time series dependence characteristics. Traditional forecasting methods are difficult to accurately capture its long-term changing patterns, resulting in large deviations in forecast results. In order to improve the accuracy of cold chain logistics demand forecasting and optimize industry resource allocation, infrastructure construction and development planning, this article takes the total national cold chain logistics demand as the research object, selects 12 core indicators from macroeconomics, population consumption, industry development and other dimensions as influencing factors, collects relevant statistical data from 2005 to 2024, and builds a cold chain logistics demand forecasting model based on long short-term memory network (LSTM). The core parameters of the LSTM model are optimized based on time series characteristics, and the prediction accuracy of the model is verified through three indicators: root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). Research has confirmed that the LSTM model has excellent applicability and accuracy in cold chain logistics demand forecasting, and its prediction results can effectively support the scientific and intelligent development of the cold chain logistics industry.

Keywords: Cold chain logistics; Demand forecast; LSTM long short-term memory network.

1. Introduction

As a core component of the modern logistics system, cold chain logistics is a key link in ensuring the quality, safety and efficient circulation of fresh agricultural products, quick-frozen foods, pharmaceutical cold chain products and other materials. Its development level is directly related to people's livelihood security and the process of industrial upgrading. In recent years, driven by the upgrading of residents' consumption structure, the rise of new consumption models and the promotion of rural revitalization strategies, my country's cold chain logistics industry has entered a new stage of large-scale development. my country's total cold chain logistics demand has increased from 82 million tons in 2005 to 365 million tons in 2024. It has achieved more than 4 times growth in two decades, and the market revenue has exceeded 530 billion yuan. The industry as a whole has shown a steady growth trend. At the same time, the national level also attaches great importance to the development of the cold chain logistics industry. The "14th Five-Year Plan for Cold Chain Logistics Development Plan" clearly proposes to build a full-chain, networked, high-standard modern cold chain logistics system. This plan puts forward higher requirements for the industry's infrastructure layout and service capabilities.

However, the generation of cold chain logistics demand is not the result of a single factor, but is affected by the complex coupling of multi-dimensional variables such as macroeconomics, population structure, residents' income, and industrial layout. Its change pattern therefore exhibits significant nonlinear and time-dependent characteristics. Traditional qualitative forecasting methods and linear regression methods are difficult to adapt to this complex feature, resulting in insufficient forecasting accuracy. The lack of forecasting accuracy directly leads to an imbalance in cold chain resource allocation, which is manifested in the coexistence of excess facilities in some regions, idle capacity

and insufficient supply in core consumption areas. This contradiction has become a key bottleneck restricting the high-quality development of the industry. In this context, exploring accurate prediction methods that adapt to cold chain logistics demand characteristics and building scientific and effective prediction models are of great theoretical value and practical significance for optimizing the layout of cold chain infrastructure, rationally allocating transportation capacity, and achieving efficient resource allocation.

Judging from the current research status, cold chain logistics demand forecasting has always been a research hotspot in the field of logistics. The development of the foreign cold chain logistics industry started early, and the relevant research system is more mature. Scholars focus on the identification of demand influencing factors and the application of time series forecasting methods, which not only confirms the significant correlation between macroeconomics, consumer market characteristics and cold chain demand, but also verifies the feasibility of intelligent algorithms in forecasting. Domestic research has gone through an evolution process from traditional models to deep learning methods. Existing research generally recognizes the driving role of economic development level, urbanization rate, fresh food e-commerce development and other factors on cold chain demand. Among them, long short-term memory network (LSTM) is an advantageous algorithm for processing time series data, and its application value in logistics demand forecasting has been widely paid attention to. Ran Maoliang [3] proposed the EEMD-LMD-LSTM-LEC deep learning model for short-term logistics demand forecasting based on the characteristics of short-term logistics demand data, but failed to give full play to the advantages of the LSTM model in capturing the long-term dependencies of time series.

Based on the above research status and industry pain points, this article will systematically build an LSTM cold chain logistics demand forecast model, fully explore the model's

unique advantages in processing nonlinear and long-term data, and help the industry achieve optimal resource allocation and high-quality development through empirical analysis and accuracy verification.

2. Relevant theoretical and methodological foundations

2.1. Characteristics of Cold Chain Logistics Demand

Cold chain logistics demand refers to the total demand for cold chain logistics services from social production and consumption activities within a certain period of time under a specific socio-economic environment. Its core performance is the total amount of materials that need to be circulated through professional cold chain facilities such as cryogenic refrigeration and freezing [4]. Cold chain logistics demand is different from ordinary logistics demand and has significant industry characteristics: first, the diversity of driving factors, cold chain logistics demand is affected by multi-dimensional factors such as economy, population, consumption, industry, etc., and the factors are coupled with each other, forming a complex non-linear relationship; second, the time series nature of data, cold chain logistics demand data changes with time. The passage of time shows obvious continuity and trend, and there is a significant dependence between annual demand data; third, the stability and volatility of growth coexist. Overall, the demand for cold chain logistics shows a steady growth trend with economic development. At the same time, there are small short-term fluctuations affected by factors such as peak consumption seasons, policy adjustments, and market environment.

The above characteristics of cold chain logistics demand determine that its prediction work cannot use simple linear methods, but requires the selection of intelligent algorithms that can handle nonlinear, long-term, and multi-variable coupled data. The LSTM long short-term memory network can exactly meet this prediction demand.

2.2. LSTM Long Short-Term Memory Network Theoretical Basis

2.2.1. Core Principles of LSTM Model

Long short-term memory neural network (LSTM) is an improved deep learning algorithm of recurrent neural network (RNN). Its core design is to solve the gradient disappearance and gradient explosion problems of traditional RNN when processing long sequence data. Traditional RNN neurons have a simple structure and can only capture short-term dependencies of data. For long-term data such as cold chain logistics needs, it is difficult to effectively remember early key information, which leads to a decrease in prediction accuracy. The LSTM model achieves selective memory and forgetting of key information in long sequence data by introducing two core structures: gating mechanism and cell state, and can accurately capture the long-term dependence of time series data. This is also its core advantage in adapting to cold chain logistics demand forecasting.

Cell state is the core information channel throughout the entire LSTM network, responsible for storing the temporal characteristic information of data. Its update and transmission rely on the synergy of the three core gating structures of forget gate, input gate, and output gate [5]. As shown in Figure 2-1. In this repeated transfer process, there are 4 interactive activation functions, 3 Sigmoid and 1 tanh. Sigmoid converts

the incoming value to between 0 and 1, and tanh converts it to between -1 and 1.

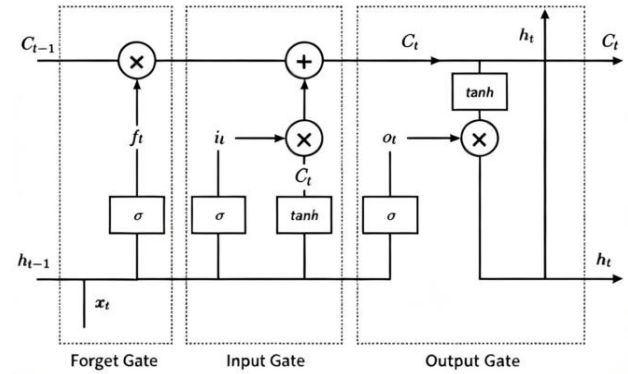


Figure 1. Structure of Long Short-Term Memory Network (LSTM)

2.2.2. The Core Structure and Operation of the LSTM Model

Specifically, the core function of the forget gate is to filter historical information. By reading the input data at the current moment and the output data at the previous moment, a value between 0 and 1 is output as the retained weight of the historical information. The lower the weight, the less important the information is. In this way, the invalid interference information in the cold chain logistics demand data is filtered and the core trend characteristics are retained.

The core function of the input gate is to screen and incorporate new effective information to complete the cell state update. Its work is divided into two steps. First, the Sigmoid function is used to determine the information dimensions that need to be updated, and then the tanh function is used to generate a new candidate information vector. Finally, the two are combined to complete the cell state update, ensuring the accurate inclusion and prediction timeliness of new feature information for cold chain logistics needs.

The core function of the output gate is to generate the current output result based on the updated cell state, determine the output dimension of the cell state through the Sigmoid function, and then normalize the cell state through the tanh function. Finally, the two are multiplied to obtain the output value, ensuring the rationality and accuracy of the output result, and fitting the actual changing laws of cold chain logistics needs. The LSTM model achieves accurate processing of long-term data through the gating mechanism. It can not only remember the long-term growth trend of cold chain logistics demand, but also capture short-term fluctuation characteristics. It is one of the best algorithms for cold chain logistics demand forecasting.

2.3. Model Prediction Accuracy Evaluation Index

In order to objectively and quantitatively evaluate the prediction effect and accuracy of the LSTM cold chain logistics demand forecast model constructed in this article, the three most commonly used evaluation indicators in time series prediction in the field of deep learning are selected, namely root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). The three indicators are all negative indicators. The smaller the calculation result value is, the smaller the deviation between the model's predicted value and the true value is, and the higher the model's prediction accuracy is. Among them, MAPE is the relative error index, which can intuitively reflect the accuracy of prediction. It is the core evaluation index of

this article. The specific calculation formula is as follows:

Root mean square error (RMSE): By calculating the square sum of the average difference between the predicted value and the true value and then taking the square root, it reflects the average degree of deviation between the predicted value and the true value. The smaller the value, the higher the accuracy.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Mean absolute error (MAE): Calculates the average of the absolute values of the difference between the predicted value and the true value. It is not sensitive to outliers and can stably reflect the prediction deviation.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Mean Absolute Percent Error (MAPE): Calculate the average of the absolute value of the difference between the predicted value and the true value as a percentage of the true value and convert it into a percentage. It can intuitively reflect the relative accuracy of the prediction and is a core evaluation indicator.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100$$

In the formula, n is the number of samples, y_i is the true value of cold chain logistics demand, and $\hat{y}_i$ is the predicted value of cold chain logistics demand output by the model.

3. Construction of Cold Chain Logistics Demand Forecast Model Based on LSTM

3.1. Predictive Indicator Selection and Data Sources

3.1.1. Core Influencing Factors and Indicator Selection

The formation of cold chain logistics demand is affected by multi-dimensional factors, and the rationality of indicator selection directly determines the prediction accuracy of the model. This article combines the development characteristics of the cold chain logistics industry with existing research results, follows the principles of scientificity, comprehensiveness, quantification, and accessibility, and selects a total of 12 core indicators as input variables for the model from the three dimensions of macroeconomics, population and consumption, and industry development, and uses the total national cold chain logistics demand as the only output variable of the model to build a complete cold chain logistics demand prediction indicator system. The specific indicators are as follows.

Macroeconomic dimension: Gross domestic product (GDP), primary industry added value, and tertiary industry added value proportion. This dimension index reflects the macroeconomic development level and industrial structure characteristics, and is the core basis for the growth of cold chain logistics demand.

Population and consumption dimension: urban population, per capita disposable income of residents, and total retail sales of consumer goods. This dimension index reflects the size of the consumer group, consumption power and overall consumption demand, and directly affects the terminal

demand of cold chain logistics.

Industry development dimensions: fresh food e-commerce transaction volume, total cold storage capacity, refrigerated truck ownership, Internet penetration rate, total agricultural product production, and cold chain logistics policy dummy variables. This dimension indicator reflects the development level of the cold chain logistics industry, supporting facility construction and market supply capabilities, and is the direct driving factor for cold chain logistics demand.

3.1.2. Data Source

This article selects 20 years of annual data from 2005 to 2024 as a research sample. The data comes from statistical yearbooks, industry reports and policy documents issued by the National Bureau of Statistics, China Federation of Logistics and Purchasing, China Cold Chain Logistics Alliance and other institutions.

3.2. Data Preprocessing

In order to avoid problems such as inconsistent dimensions, large numerical differences, outliers, and missing values in the original data that lead to slow model convergence, low prediction accuracy, and even fitting deviations, the original data is standardized and preprocessed before building the model. The preprocessing steps are as follows:

Data cleaning: Use the Laida criterion (3σ) to detect outliers in the original data, eliminate extreme data that deviate from the normal range, and ensure the integrity and continuity of the data.

Data normalization: Since the dimensions and numerical ranges of each indicator are quite different, the Min-Max normalization method is used to map all data to the $[0, 1]$ interval to eliminate the influence of dimensions and improve the training efficiency and convergence speed of the model. The normalization formula is: $x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$. Among them, x is the original data, $x_{\min}$ and $x_{\max}$ are the minimum and maximum values of the indicator respectively, and x' is the normalized standard data.

Data set division: Following the basic principles of time series prediction, 14 years of data from 2005 to 2018 are divided into model training sets for parameter learning and fitting of the model; 6 years of data from 2019 to 2024 are divided into model test sets to verify the generalization ability and prediction accuracy of the model. The sample ratio of the training set to the test set is 7:3 to comply with the data set division specifications of deep learning.

4. Empirical Analysis and Result Verification

4.1. Sample Data Analysis

From 2005 to 2024, the total demand for cold chain logistics in the country has shown a sustained and stable growth trend. The data chart shows that it has increased from 82 million tons in 2005 to 365 million tons in 2024. The average annual growth rate during the two decades has reached 8.3%, with no obvious negative growth or cliff-like fluctuations. The overall growth trend is highly consistent with the pace of macroeconomic development and residents' consumption upgrades. Judging from the changing characteristics of the input indicators, the GDP increased from 18.73 trillion yuan to 130.58 trillion yuan, the urban population increased from 562 million to 937 million, and the per capita disposable income of residents increased from

10,493 yuan. To 49,969 yuan, the transaction volume of fresh food e-commerce has achieved leapfrog growth from 0 to over 3 trillion yuan. All core indicators have shown a steady growth trend, which is significantly consistent with the growth trend of cold chain logistics demand, further verifying the rationality of indicator selection.

4.2. Analysis of Prediction Accuracy and Fitting Effect of LSTM Model

The preprocessed training set data was input into the LSTM model for training. After iteratively optimizing the parameters, the model had excellent convergence, no overfitting or underfitting, and good generalization ability. Input the preprocessed test set data into the trained model to obtain the predicted value of national cold chain logistics demand from 2019 to 2024. By calculating three evaluation indicators of RMSE, MAE, and MAPE, the prediction accuracy of the model is quantitatively evaluated. At the same time, by drawing the fitting curve between the predicted value and the real value, the fitting effect of the model can be intuitively analyzed.

4.2.1. Core Influencing Factors and Indicator Selection

After calculation, the three core evaluation index results of the LSTM model constructed in this article in cold chain logistics demand forecasting are: RMSE=0.036, MAE=0.029, MAPE=1.28%. Judging from the index values, the three indicators are all at extremely low levels. Among them, the core indicator MAPE is only 1.28%, indicating that the average prediction relative error of the model is less than 2%, and the prediction accuracy reaches the high standard of industry applications. This result fully proves that the LSTM model can accurately capture the time series characteristics and nonlinear change patterns of cold chain logistics demand, and the prediction results are highly reliable.

4.2.2. Fitting Effect Analysis

Judging from the fitting curve between the predicted value and the real value, there is almost no obvious deviation or lag between the predicted value curve output by the LSTM model and the real value curve. As shown in Figure 4-1. In years when cold chain logistics demand fluctuates slightly, the model can accurately capture the fluctuation characteristics, and the fluctuation range of the predicted value is consistent with the real value; in years when cold chain logistics demand grows rapidly, the model can accurately predict the growth rate without excessive overestimation or underestimation.

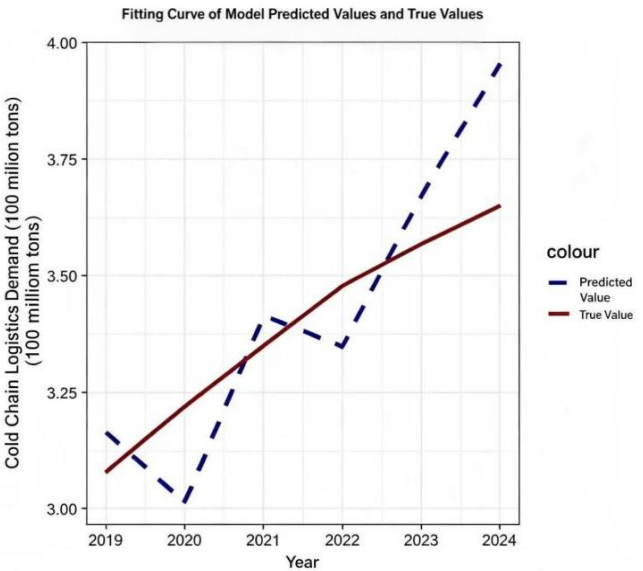


Figure 2. Fitting graph of predicted values and actual values from 2019 to 2024

Even if the growth pace of cold chain logistics demand is slightly adjusted due to the impact of the external environment from 2022 to 2024, the model can still accurately fit the changing characteristics, and the error between real demand and predicted demand is very small, as shown in Table 4-1. The excellent performance of the LSTM model in cold chain logistics demand prediction is further verified.

Table 1. Error between Predicted Values and Actual Values from 2019 to 2024

Year	Actual Demand	Forecasted Demand	Error
2019	3.08	3.16	0.085
2020	3.22	3.01	0.208
2021	3.35	3.42	0.067
2022	3.48	3.35	0.130
2023	3.57	3.67	0.103
2024	3.65	3.95	0.304

5. Conclusion

Based on the statistical data from 2005 to 2024, this article constructed an LSTM cold chain logistics demand forecast model and completed empirical analysis. It was concluded that cold chain logistics demand is positively affected by multi-dimensional factors such as macroeconomics and population consumption. The gating mechanism of the LSTM model can accurately capture the long-term trends and short-term fluctuations in demand. It has strong adaptability, good stability, and has practical application value.

Based on the research conclusions, in the future, the industry can further scientifically plan the layout of cold chain infrastructure and optimize resource allocation; accelerate the digital transformation of the industry and promote the application of intelligent technology and prediction models; improve the policy and data support system to help industry standardization, intelligence, and green development.

The sample span of this study is limited. In the future, higher frequency data can be collected to improve the prediction granularity, and an optimization model of emergency impact factors can be introduced. At the same time, an LSTM hybrid model can be constructed to improve prediction accuracy, and an exclusive prediction model can be built for cold chain subdivisions such as agricultural products

and medicine to provide more accurate decision support.

References

- [1] Li, G. X., Ma, W. B., & Xia, G. E. (2021). Research on logistics demand forecast model based on deep learning. *Journal of Systems Science*, 29(2), 85–89. (In Chinese)
- [2] Ran, Y. Z. (2021). Logistics demand forecast in Guizhou Province based on deep learning model. *Logistics Engineering and Management*, 43(6), 19–20+18. (In Chinese)
- [3] Ran, M. L., Chen, Y. R., & Yang, X. B. (2022). Short-term logistics demand forecast based on EEMD-LMD-LSTM-LEC deep learning model. *Control and Decision*, 37(10), 2513–2523. (In Chinese)
- [4] Wang, M. T. (2023). Research on cold chain logistics demand forecast [Master's thesis]. Shanxi University of Finance and Economics. (In Chinese)
- [5] Wang, Z. Y., & Zhang, Z. Q. (2024). Comparative study on the performance of LSTM and GRU models in predicting logistics demand in Hubei Province. *Logistics Engineering and Management*, 46(4), 10–14. (In Chinese)
- [6] Guo, W. Y., Li, Y. T., Bai, K. C., et al. (2025). Cold chain logistics demand forecast for fresh products based on deep learning. *Logistics Research*, (2), 75–83. (In Chinese)
- [7] Li, S. C., & Ye, J. (2022). Analysis and prediction of agricultural product cold chain logistics demand based on gray regression model. *Highway Transportation Technology*, 39(5), 166–174. (In Chinese)
- [8] Zhang, Z. (2025). Research on Cold Chain Logistics Demand Forecast of Fresh Agricultural Products in Hubei Province [Master's thesis]. Jiangnan University. (In Chinese)
- [9] Wei, X. (2024). Research on demand forecasting of cold chain logistics of agricultural products based on artificial neural network [Master's thesis]. Shenyang University of Technology. (In Chinese)
- [10] Duncombe, J. U. (1959). Infrared navigation—Part I: An assessment of feasibility (Periodical style). *IEEE Transactions on Electron Devices*, ED-11, 34–39.
- [11] Du, L. J., & Zhang, Y. L. (2025). Analysis of high-quality development paths of cold chain logistics. *China Storage and Transportation*, (10), 117. (In Chinese)