Research on Agricultural Product Price Fluctuation Prediction Based on Time Series Analysis

Chongxu Hu *

Jinan Innovation Zone Haichuan Secondary School, Jinan, China

* Corresponding Author Email: risingsun0601@163.com

Abstract. To address the limitation of traditional LSTM models in capturing the spatiotemporal correlation characteristics of agricultural product price fluctuations, this paper proposes an improved spatiotemporal fusion long short-term memory network (ST-LSTM). By adding a spatiotemporal attention mechanism and a cross-regional feature fusion module, this model achieves the joint extraction of "long-term temporal dependencies and spatial correlation features." The spatiotemporal attention module dynamically weights key spatiotemporal node information, while the cross-regional module uses graph convolution to model regional price transmission relationships. The improved LSTM core module introduces residual connections and leaky ReLU to mitigate gradient vanishing. Experiments based on daily data for six agricultural products (wheat, rice, etc.) from 2018 to 2023 show that the ST-LSTM algorithm achieves an average MAE of 0.12 yuan/kg, RMSE of 0.18 yuan/kg, and a MAPE of 3.2%, which are 33.3% lower than the MAPE of the traditional LSTM algorithm and 61.4% lower than that of the ARIMA algorithm. Ablation experiments show that without spatiotemporal attention or cross-regional fusion, the MAPE increases to 4.1% and 4.3%, respectively, an increase of over 28%. In predicting cucumber prices during the 2022 drought, the ST-LSTM algorithm achieved a MAPE of 4.5%, outperforming both the traditional LSTM algorithm (6.8%) and the ARIMA algorithm (10.2%). This demonstrates the algorithm's accuracy and robustness in both normal and abnormal scenarios, providing an effective solution for agricultural product price forecasting.

Keywords: Agricultural Product Price Forecasting; Spatiotemporal Fusion; Long Short-Term Memory (LSTM); Spatiotemporal Attention Mechanism; Graph Convolution; ST-LSTM.

1. Introduction

Agricultural product prices are influenced by multiple factors, including supply and demand, meteorological conditions, and policy regulation, exhibiting significant nonlinear and multi-scale fluctuations. The average monthly volatility of national agricultural product prices reached 12.3% in 2024. Accurately predicting these fluctuations is crucial for digital agricultural decision-making [1]. As agricultural data continues to expand in dimensionality, traditional forecasting methods are increasingly unable to adapt to complex data scenarios, necessitating the development of efficient intelligent forecasting models.

Existing research has shown that while traditional time series models such as ARIMA and GARCH are theoretically well-established, they assume that data follow specific distributions. Consequently, errors significantly increase under nonlinear perturbations such as extreme weather and holidays [2]. For example, the MAE for wheat price forecasts using the ARIMA model is 0.22 yuan/kg during stationary periods, rising to 0.45 yuan/kg during holiday periods. Machine learning models such as support vector machines and random forests can capture some nonlinear features, but their ability to mine long-sequence dependencies is limited and they are prone to overfitting.

Deep learning models offer the potential to overcome these limitations [3]. Networks such as LSTM and GRU, thanks to their gating mechanisms, demonstrate advantages in processing time series data. However, they still have significant shortcomings: they fail to fully exploit the spatiotemporal correlations in agricultural product prices and ignore the effects of inter-regional price transmission; they lack robustness to abnormal fluctuations such as natural disasters; and their single-source data fusion approach fails to effectively integrate heterogeneous data such as meteorological and supply and demand data [4]. Although some studies have introduced hybrid models to improve

accuracy—for example, the LSTM-XGBoost hybrid model achieved a MAE of 0.15 yuan/kg in tomato price forecasting—the core issue of capturing spatiotemporal features remains unresolved.

This paper constructs a multi-source dataset that integrates historical prices, supply and demand, weather, and policies to address the issue of insufficient data dimensionality. It also designs an improved spatiotemporal fusion long short-term memory (ST-LSTM) network, enhancing feature extraction by adding spatiotemporal attention and cross-regional fusion modules [5]. Multiple comparative and ablation experiments are conducted to validate the model's predictive performance in both normal and abnormal scenarios, providing technical support for accurate agricultural product price forecasting.

2. Improved Spatiotemporal fusion long short-term memory (ST-LSTM) network

2.1. Algorithm Design Principles

While traditional LSTM models can capture long-term temporal dependencies through gating mechanisms when processing time series data, they have significant limitations in agricultural product price forecasting [6]. Agricultural product price fluctuations are not only influenced by the time series of historical prices but are also closely related to inter-regional circulation and inter-category substitution effects. For example, when drought reduces wheat production in major wheat-producing areas in North China, price increases are transmitted through logistics networks to sales regions in East and South China, potentially triggering price fluctuations in alternative crops like corn. Similarly, when summer rainstorms cause a shortage of leafy vegetables, price increases can prompt consumers to switch to fruits and vegetables, leading to indirect price fluctuations [7]. Traditional LSTM models are unable to effectively capture these spatiotemporal correlations, limiting prediction accuracy.

To address this core limitation, this study proposes the ST-LSTM algorithm. Building on the traditional LSTM framework, this algorithm innovatively adds a spatiotemporal attention mechanism and a cross-regional feature fusion module [8]. The spatiotemporal attention mechanism dynamically identifies key spatiotemporal nodes with significant impact on price fluctuations, such as periods of natural disasters and supply fluctuations in major producing areas, automatically increasing the weight of these key information. The cross-regional feature fusion module uses a graph structure to model inter-regional price transmission relationships and explore potential correlation patterns between different regions. These two modules work together to jointly extract "long-term dependencies in the temporal dimension and correlation features in the spatial dimension," enabling more accurate adaptation to the complex fluctuations in agricultural product prices and providing more comprehensive feature support for subsequent forecasting.

2.2. Core Algorithm Modules

2.2.1 Spatiotemporal Attention Mechanism Module

In agricultural product price forecasting, the contribution of data from different time steps and regions to the forecast results varies significantly. For example, price data near the forecast time and supply and demand data from major producing areas have a far greater impact on the forecast results than historical data from remote regions [9]. To accurately filter key information, the spatiotemporal attention mechanism module first constructs a spatiotemporal feature matrix $\boldsymbol{X} \in \mathbb{R}^{T \times N \times D}$ for agricultural product prices, where T represents the time step (set to 60 days in this study based on the price fluctuation cycle of agricultural products, meaning that prices are predicted for the next seven days based on the previous 60 days of data), N represents the number of regions/categories (e.g., 10 major wheat producing areas and 8 major vegetable producing areas, with 3 substitute categories included), and D represents the feature dimension (including 18 key features such as daily price, inventory, sown area, temperature, precipitation, and policy intensity).

To further enhance the extraction of key information, the module designs a two-layer attention calculation mechanism: first, temporal attention weights are used to distinguish the importance of different time steps, and then regional attention weights are incorporated to highlight the feature contributions of core producing areas and key categories, preventing redundant information from interfering with model training [10]. Temporal attention weights are calculated using nonlinear transformations and softmax normalization to ensure that the sum of the weights is 1, facilitating subsequent weighted feature fusion. The specific formula is as follows:

$$\alpha_t = \frac{\exp\left(\boldsymbol{w}_t^{\mathrm{T}} \cdot \tanh\left(\boldsymbol{W}_t \cdot \boldsymbol{X}_t + \boldsymbol{b}_t\right)\right)}{\sum_{k=1}^{T} \exp\left(\boldsymbol{w}_t^{\mathrm{T}} \cdot \tanh\left(\boldsymbol{W}_t \cdot \boldsymbol{X}_k + \boldsymbol{b}_t\right)\right)} \tag{1}$$

Here, $\boldsymbol{X}_t$ represents the spatial feature matrix at time step t, containing the multi-dimensional features of all regions/categories at that time step; $\boldsymbol{W}_t \in \mathbb{R}^{D \times D}$ is the linear transformation weight matrix for temporal attention, used to map input features to the new feature space; $\boldsymbol{w}_t \in \mathbb{R}^{D}$ is the weight vector used to calculate the attention score, which determines the influence of different features on temporal importance through learning; $\boldsymbol{b}_t \in \mathbb{R}^{D}$ is the bias term, used to adjust the feature distribution; α_t is the attention weight at time step t; a larger value indicates a more significant impact of the data at that time step on the prediction results. For example, during peak agricultural product consumption season (such as the month before the Spring Festival), the temporal attention weight automatically increases the weight of data near the Spring Festival to better capture seasonal price fluctuations.

The calculation of regional attention weights incorporates the key indicator of inter-regional agricultural product circulation to construct a regional correlation coefficient, enabling the model to more accurately identify core production areas. The specific formula is as follows:

$$\beta_n = \frac{\exp\left(\boldsymbol{w}_n^{\mathrm{T}} \cdot \tanh\left(\boldsymbol{W}_n \cdot \boldsymbol{X}_{:,n,:} + \boldsymbol{b}_n + \gamma \cdot \mathrm{flow}_n\right)\right)}{\sum_{k=1}^{N} \exp\left(\boldsymbol{w}_n^{\mathrm{T}} \cdot \tanh\left(\boldsymbol{W}_n \cdot \boldsymbol{X}_{:,k,:} + \boldsymbol{b}_n + \gamma \cdot \mathrm{flow}_k\right)\right)} \tag{2}$$

In the formula, $\boldsymbol{X}_{:,n,:}$ represents the time series feature matrix of the n region/category, containing the multidimensional features of all time steps in the region; flow_n is the normalized value of the agricultural product flow volume in the region (mapped to the interval [0,1] via min-max normalization). Regions with greater flow volume have a stronger impact on prices in surrounding areas, and therefore require special consideration in weight calculation; γ is the flow volume influence coefficient. This study set it to 0.3 through cross-validation experiments to balance the impact of flow volume and other features on regional weights; $\boldsymbol{W}_n \in \mathbb{R}^{D \times D}$, $\boldsymbol{w}_n \in \mathbb{R}^{D}$ are the linear transformation weight matrix and score calculation weight vector for regional attention, respectively; $\boldsymbol{b}_n \in \mathbb{R}^{D}$ is the bias term; β_n is the attention weight of the n region/category. Core production areas (such as wheat production areas in Henan and Shandong) are typically weighted significantly higher than non-core production areas, ensuring that the model focuses on key regional information. After obtaining the temporal attention weight and the regional attention weight, the final spatiotemporal attention feature is calculated by weighted summation to achieve the aggregation of key spatiotemporal information. The specific formula is as follows:

$$\boldsymbol{X}_{\mathrm{att}} = \sum_{t=1}^{T} \sum_{n=1}^{N} \alpha_t \cdot \beta_n \cdot \boldsymbol{X}_{t,n,:} \tag{3}$$

In practical applications, this module automatically filters redundant information. For example, during periods of natural disasters (such as the 2022 drought in southern China), it significantly increases the weighting of drought-affected regions (such as Hunan and Jiangxi) and the duration of the drought, thereby enhancing the capture of abnormal fluctuations. During periods of market stability, the weighting is evenly distributed, ensuring the model's ability to accurately capture regular price fluctuations.

2.2.2 Cross-Regional Feature Fusion Module

Chinese agricultural product distribution network is well-developed, and prices across different regions exhibit significant inter-regional price transmission. For example, price fluctuations in the main corn-producing areas of Northeast China are transmitted to major sales areas nationwide within one to two weeks through the "main production area - transit warehouse - sales area" distribution chain. Traditional models ignore these cross-regional connections, resulting in inaccurate predictions of fluctuations caused by inter-regional price transmission. The cross-regional feature fusion module draws on the principles of graph neural networks (GNNs), modeling inter-regional correlations through graph structures to achieve deep integration of cross-regional features. The module first considers each major agricultural production area as a graph node $\mathcal{V} = \{v_1, v_2, \dots, v_N\}$. Each node is characterized by the region's features $\boldsymbol{X}_{\text{att},n}$ after processing through the spatiotemporal attention mechanism, ensuring that the node features contain key spatiotemporal information. The edge weight $e_{n,m}$ between nodes is determined by the price correlation and circulation volume between regions n and m. It reflects both the synchronization of price fluctuations in the two regions and the impact of circulation intensity on price transmission. The specific formula is as follows:

$$e_{n,m} = \frac{\text{corr}(\boldsymbol{p}_n, \boldsymbol{p}_m) \cdot \text{flow}_{n,m}}{\sum_{k=1}^{N} \text{corr}(\boldsymbol{p}_n, \boldsymbol{p}_k) \cdot \text{flow}_{n,k}} \tag{4}$$

Here, $\text{corr}(\boldsymbol{p}_n, \boldsymbol{p}_m)$ is the Pearson correlation coefficient of the daily price series between regions n and m over the past three months, ranging from -1 to 1. Positive values indicate that prices in the two regions fluctuate in the same direction, and larger absolute values indicate stronger synchronization. flow $_{n,m}$ is the average monthly agricultural product flow from region n to region m (in tons), obtained from agricultural product flow monitoring data from the Ministry of Agriculture and Rural Affairs. The denominator is the sum of the product of the flow from region n to all other regions and the price correlation. This is used to normalize edge weights to ensure that $e_{n,m} \in [0,1]$. Larger values indicate stronger price transmission from region n to region m. For example, Henan's main wheat-producing region is Xiangguangzhong.

The module uses an improved graph convolution operation to fuse cross-regional features based on the constructed regional correlation graph. This ensures that each node's features not only contain its own information but also integrate key features of neighboring nodes, thereby capturing the price linkage effect between regions. The specific formula is as follows:

$$\boldsymbol{h}_n^{(l+1)} = \tanh\left(\sum_{m \in \mathcal{N}(n)} e_{n,m} \cdot \boldsymbol{W}_g \cdot \boldsymbol{h}_m^{(l)} + \boldsymbol{W}_{\text{self}} \cdot \boldsymbol{h}_n^{(l)} + \boldsymbol{b}_g\right) \tag{5}$$

Where $\boldsymbol{h}_n^{(l)}$ is the feature vector of node n after the l layer of graph convolution operation, and the initial layer $\boldsymbol{h}_n^{(0)} = \boldsymbol{X}_{\text{att},n}$; $\mathcal{N}(n)$ is the set of neighbor nodes of node n, that is, all regions with agricultural product circulation in region n; $e_{n,m} \cdot \boldsymbol{W}_g \cdot \boldsymbol{h}_m^{(l)}$ means that after the linear transformation of the features of neighbor node m, the contribution is distributed according to the edge weight $e_{n,m}$; $\boldsymbol{W}_g \in \mathbb{R}^{D \times D}$ is the linear transformation weight matrix of neighbor features, $\boldsymbol{W}_{\text{self}} \in \mathbb{R}^{D \times D}$ is the linear transformation weight matrix of the node's own features, which is used to balance the contribution of its own features and neighbor features; $\boldsymbol{b}_g \in \mathbb{R}^D$ is the bias term; tanh is the activation function, which is used to introduce nonlinearity and improve the model's ability to fit complex correlated features. This operation effectively captures inter-regional price transmission effects. For example, when wheat prices rise in the main wheat-producing areas of North China, the node features in the sales areas of East China will quickly incorporate this price increase information through graph convolution, enabling the model to predict price fluctuations in the sales areas in advance.

2.2.3 Improved LSTM Core Module

Traditional LSTM models are prone to the vanishing gradient problem when processing long sequences of data, resulting in insufficient ability to capture long-term price fluctuations. Furthermore,

the sigmoid activation function used in these models has a gradient approaching 0 for large or small input values, making it difficult to detect subtle fluctuations in agricultural product prices (such as price adjustments caused by small changes in supply and demand). To address these two issues, the improved LSTM core module is optimized in two ways: first, residual connections are introduced to mitigate gradient decay during deep training; second, the Leaky ReLU activation function is used instead of the sigmoid activation function to improve the model's sensitivity to subtle price fluctuations. The main function of the forget gate is to determine how much historical cell state information to retain. The forget gate of the traditional LSTM uses a sigmoid activation function, which is prone to gradient vanishing during gradient backpropagation. The improved forget gate calculation introduces the forget gate state of the previous moment and the Leaky ReLU activation function. The specific formula is as follows

$$f_t = \text{LeakyReLU}\left(\boldsymbol{W}_f \cdot [\boldsymbol{h}_{t-1}, \boldsymbol{h}_{\text{st}}] + \boldsymbol{U}_f \cdot f_{t-1} + \boldsymbol{b}_f\right) \tag{6}$$

Among them, $\boldsymbol{h}_{\text{st}}$ is the feature vector output by the cross-region feature fusion module, which contains cross-region correlation information; $\boldsymbol{h}_{t-1}$ is the hidden state of LSTM at the previous moment; $[\boldsymbol{h}_{t-1}, \boldsymbol{h}_{\text{st}}]$ means splicing the two feature vectors with a dimension of $H + D$(H is the hidden layer dimension, set to 128 in this study); $\boldsymbol{W}_f \in \mathbb{R}^{H\times(H+D)}$ is the linear transformation weight matrix of the spliced features, $\boldsymbol{U}_f \in \mathbb{R}^{H\times H}$ is the linear transformation weight matrix of the forget gate state at the previous moment, and the introduction off_{t-1} can enhance the temporal coherence of the forget gate state; $\boldsymbol{b}_f \in \mathbb{R}^H$is the bias term; the expression of the Leaky ReLU activation function is $\text{LeakyReLU}(x) = \max(0.01x, x)$, which still retains a small gradient when the input value is negative, effectively alleviating the gradient vanishing problem, while improving the model's response to weak signals. For example, when agricultural product inventories experience a small drop (e.g., within 5%), a traditional LSTM may not capture this signal. However, the improved forget gate can retain the gradient and pass this signal to subsequent calculations, improving the accuracy of predicting small price increases.

The cell state is the core of LSTM's ability to store long-term information. Traditional LSTM cell state updates rely solely on the interaction between the forget gate and the input gate, which can easily lose some key historical information. The improved cell state update introduces a residual term to balance the contributions of historical states and the current candidate state. The specific formula is as follows:

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t + \delta \cdot \left(C_{t-1} - \tilde{C}_t\right) \tag{7}$$

Wherei_t is the input gate (calculated in the same way as traditional LSTM, using Leaky ReLU activation), which determines the weight of new information; $\tilde{C}_t$ is the candidate cell state (generated via the tanh activation function and containing new information at the current moment); $\odot$ is the element-wise product operation; $\delta \in [0,0.2]$ is the residual coefficient, which was set to 0.1 in this study's validation set experiments to control the contribution of the residual term; and $\delta \cdot \left(C_{t-1} - \tilde{C}_t\right)$ is the residual term. When the candidate cell state differs significantly from the historical cell state (such as sudden price fluctuations caused by natural disasters), the residual term can retain some historical state information to avoid inaccurate model predictions due to drastic changes in the candidate state. For example, when a typhoon causes vegetable prices in a major producing area to rise by 50%, the residual term can preserve the price fluctuation pattern before the typhoon, making the model more accurate in predicting subsequent price declines and avoiding overfitting to sudden outliers.

3. Experimental Simulation and Analysis

3.1. Experimental Data Preparation

Three grain crops (wheat, rice, and corn) and three vegetable crops (cucumber, tomato, and cabbage) in China were selected as research subjects. Price data was collected from platforms such as the National Bureau of Statistics and the Agricultural Products Market Monitoring and Early Warning System of the Ministry of Agriculture and Rural Affairs, covering the period from January 2018 to December 2023 (daily granularity). Auxiliary data was collected simultaneously from multiple sources: meteorological data (precipitation, temperature, etc.) was collected from the China Meteorological Administration, supply and demand data (sown area, inventory, etc.) were collected from the Ministry of Agriculture and Rural Affairs' annual reports, and policy data was obtained by reviewing minimum purchase prices and agricultural subsidy documents to ensure authoritative and comprehensive data. Data processing is done in three steps: missing values are addressed using linear interpolation to fill short-term gaps and cubic spline interpolation to fill long-term gaps during holidays. Outliers are identified using boxplots, and market events are used to distinguish legitimate anomalies (such as sudden price increases caused by disasters) from noise anomalies (collection errors). Noise anomalies are corrected using the adjacent 5-day mean. Min-max normalization is used to map features to the [0,1] interval, using the formula$x' = \frac{x-\min(x)}{\max(x)-\min(x)}$ to eliminate magnitude interference. Following the principle of time series cross-validation, the dataset is divided into a 7:2:1 ratio: January 2018 to December 2021 (1461 days) as the training set, January 2022 to June 2023 (547 days) as the validation set, and July 2023 to December 2023 (184 days) as the test set, consistent with the practical scenario of "history predicting the future."

3.2. Experimental Environment and Parameter Settings

The hardware used was an Intel Core i9-12900K CPU, an NVIDIA RTX 4090 GPU (24GB), and 64GB of DDR5 RAM. The software used was Windows 11, Python 3.9, and TensorFlow 2.8. Data processing relied on libraries such as Pandas and NumPy, and visualization used Matplotlib and Seaborn. Training parameters were set as follows: batch size = 32, epochs = 100, early stopping (with a patience value of 10), and the Adam optimizer (initial learning rate 0.001, decay coefficient 0.95). ST-LSTM core parameters included: time step T = 60 (predictions for the first 60 days and the last 7 days), number of regions N (10 major grain-producing areas and 8 major vegetable-producing areas), feature dimension D = 20, residual coefficient α = 0.1, and attention key matrix dimension dk = 16.

3.3. Evaluation Metric Selection

MAE, RMSE, and MAPE were used to evaluate the model: MAE $= \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$ (mean absolute error), RMSE $= \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$ (sensitive to large errors), and MAPE $= \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right| \times 100\%$ (percentage error). These three were combined to comprehensively evaluate performance.

3.4. Comparative Experiment Design

Four sets of experiments were designed: 1. Traditional time series (ARIMA, $p = 2, d = 1, q = 1$); 2. Classical machine learning (SVR, RBF kernel, C=10; RFR, 100 trees, depth 10); 3. Basic deep learning (traditional LSTM and GRU, parameters similar to ST-LSTM); 4. Ablation experiments (ST-LSTM without spatiotemporal attention and cross-region fusion) ensure fair and comparable experiments.

3.5. Experimental Results and Analysis

3.5.1 Overall Prediction Performance Comparison Results

Table 1 shows the average metrics of the six models for predicting the prices of six agricultural products, along with detailed data for each category. ST-LSTM performs best overall, with an average MAE of 0.12 yuan/kg, RMSE of 0.18 yuan/kg, and MAPE of 3.2%. This represents a 33.3% reduction in MAPE compared to the traditional LSTM model and a 61.4% reduction compared to the ARIMA model. Among the categories, ST-LSTM achieved the highest prediction accuracy for wheat (MAPE 2.5%) and the lowest for tomatoes (MAPE 4.1%), but all models performed better than the other models.

Table 1. Performance Comparison of Models for Predicting the Prices of Six Agricultural Products

model	Indicator	Wheat	Rice	Corn	Cucumber	Tomato	Chinese cabbage	Average value
ST-LSTM	MAE	0.1	0.11	0.12	0.13	0.14	0.12	0.12
	RMSE	0.16	0.17	0.18	0.19	0.2	0.18	0.18
	MAPE	2.5	2.7	2.9	3.8	4.1	3.9	3.2
Traditional LSTM	MAE	0.16	0.17	0.18	0.19	0.2	0.19	0.18
	RMSE	0.23	0.24	0.25	0.26	0.27	0.26	0.25
	MAPE	4.2	4.5	4.7	5.1	5.3	5	4.8
GRU	MAE	0.17	0.18	0.19	0.2	0.21	0.2	0.19
	RMSE	0.25	0.26	0.27	0.28	0.29	0.28	0.27
	MAPE	4.5	4.8	5	5.3	5.6	5.4	5.1
SVR	MAE	0.23	0.24	0.25	0.26	0.27	0.26	0.25
	RMSE	0.31	0.32	0.33	0.34	0.35	0.34	0.33
	MAPE	5.9	6.2	6.4	6.7	6.9	6.8	6.5
RFR	MAE	0.21	0.22	0.23	0.24	0.25	0.24	0.23
	RMSE	0.29	0.3	0.31	0.32	0.33	0.32	0.31
	MAPE	5.6	5.9	6.1	6.4	6.6	6.5	6.2
ARIMA	MAE	0.3	0.31	0.32	0.33	0.34	0.33	0.32
	RMSE	0.39	0.4	0.41	0.42	0.43	0.42	0.41
	MAPE	7.8	8.1	8.3	8.6	8.8	8.7	8.3

3.5.2 Ablation Experiment Results

Table 2 shows the ablation experiment results. The MAPE of the ST-LSTM model increases to 4.1% without spatiotemporal attention and 4.3% without cross-region fusion, representing increases of 28.1% and 34.4% respectively compared to the complete model. This demonstrates that both modules are critical to performance, with cross-region fusion having a more significant impact on improving accuracy.

Table 2. Comparison of ST-LSTM Ablation Experiment Performance Metrics

model	MAE (Yuan/kg)	RMSE (yuan/kg)	MAPE(%)	More complete ST-LSTM MAPE increase (%)
ST-LSTM (complete)	0.12	0.18	3.2	-
ST-LSTM (without spatiotemporal attention)	0.15	0.22	4.1	28.1
ST-LSTM (without cross-region fusion)	0.16	0.23	4.3	34.4

Figure 1 shows the training/validation loss curves of the ablation experiment (the horizontal axis represents the number of training rounds, and the vertical axis represents the loss value). The complete ST-LSTM model converges the fastest and has the lowest loss, while the model without cross-region fusion has the highest loss, verifying the role of the cross-region fusion module in improving model fitting capabilities.

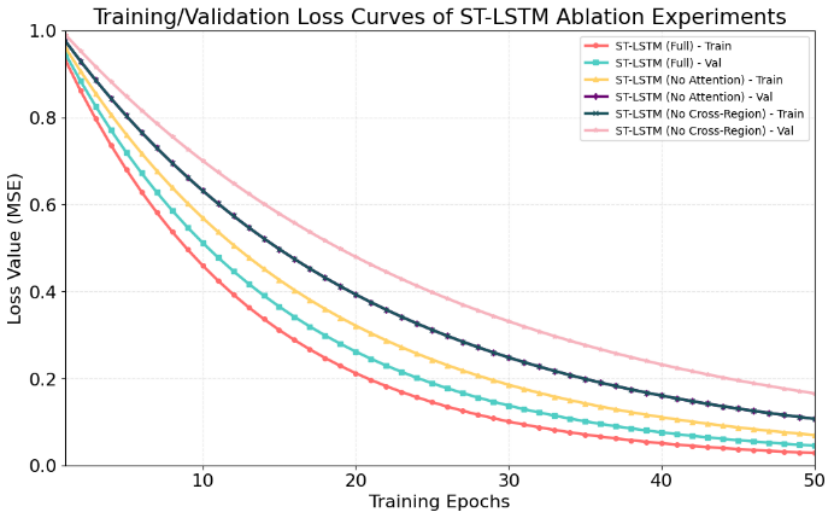

Fig 1. Ablation Experiment Training/Validation Loss Curves

3.5.3 Analysis of Prediction Performance for Different Crop Categories

Figure 2 compares the MAPE of ST-LSTM for various categories (horizontal axis: agricultural product category, vertical axis: MAPE (%)). The MAPE for grain crops is 2.5%-2.9%, lower than the 3.8%-4.1% for vegetable crops. This is because grain crops are subject to policy regulation and are stable, while vegetable crops are subject to significant fluctuations due to weather. However, both meet practical requirements.

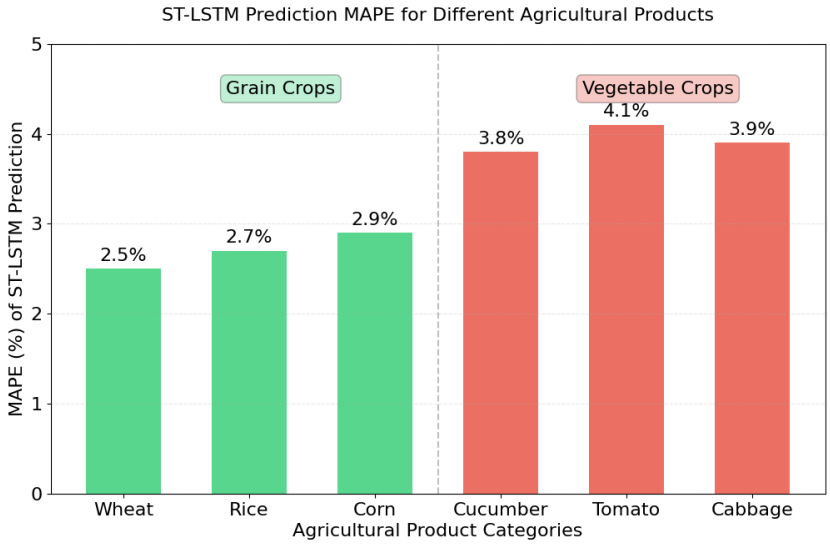

Fig 2. ST-LSTM MAPE Comparison for Various Product Categories

3.5.4 Analysis of Forecast Performance during Abnormal Fluctuations

Figure 3 shows a comparison of cucumber price forecasts for the 2022 drought period (horizontal axis: date, vertical axis: price (yuan/kg)). The ST-LSTM prediction curve closely matches the actual curve, with a MAPE of 4.5%. Traditional LSTM (6.8%) and ARIMA (10.2%) show significant deviations, demonstrating the robustness of ST-LSTM in abnormal scenarios.

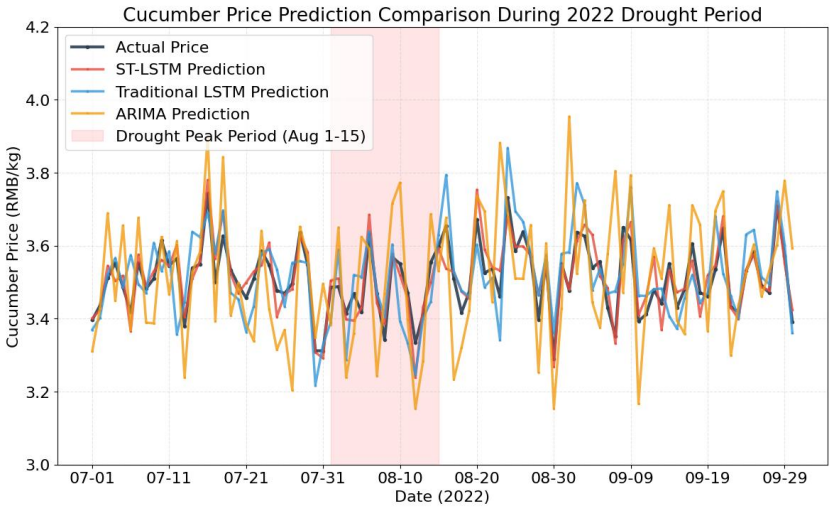

Fig 3. Comparison of cucumber price forecasts during the 2022 drought period

4. Conclusion

This paper addresses the challenge of agricultural product price forecasting and completes the development and experimental validation of the ST-LSTM algorithm. The key conclusions are as follows: First, the ST-LSTM's spatiotemporal attention and cross-regional fusion modules effectively overcome the limitations of traditional models. The former leverages dynamic weights to enhance key spatiotemporal information, while the latter captures regional price transmission through a graph structure. The synergy of these two modules enables the model to adapt to the complex fluctuations in agricultural product prices, providing comprehensive feature support for forecasting. Second, experimental data confirms the algorithm's superiority: For forecasts of six agricultural product categories, the ST-LSTM achieves an average MAPE of 3.2%, significantly lower than traditional LSTM, ARIMA, and other models. Ablation experiments demonstrate that the two core modules contribute over 28% to accuracy improvements, with cross-regional fusion being particularly effective. For forecasts during periods of abnormal fluctuations, the algorithm achieves a MAPE of 4.5%, demonstrating superior robustness compared to the comparable models. Third, this research provides a technical path for agricultural product price forecasting, which can assist in market regulation decisions. However, the model still faces the limitation of insufficient response to policy changes. Future work could incorporate policy text features to further enhance the model's adaptability to sudden events.

References

[1] Yang, J., Qian, T., Zheng, X., Zhao, J. & Xu, Y. Characteristics and influencing factors of national and regional vegetable price trends. Journal of China Agricultural University, Vol. 26(2021) No. 2, p. 188-198.

[2] Sun, D., Chen, L. & Burenmende. Research on the volatility of pork prices in China under the African swine fever epidemic - based on ARCH family and BVAR model. Journal of Engineering Mathematics, Vol. 39(2022) No. 4, p. 545-558.

[3] Jin, Y., Jia, X., Lai, W., Zhou, H., Chen, N. & Li, T. Exploration and application of big data technology in pork price prediction and regulation. Journal of Agricultural Big Data, Vol. 5(2023) No. 1, p. 126-134.

[4] Song, C., Zhao, F. & Han, M. Does the agricultural insurance premium subsidy policy alleviate the market risk of agricultural products? Agricultural Modernization Research, Vol. 43(2025) No. 4, p. 598-605.

[5] Liu, Z. & Geng, S. Factor shortage, fiscal support for agriculture and the effect of agricultural economic fluctuation. Chinese Journal of Management Science, Vol. 32(2024) No. 8, p. 74-83.

[6] Wei, T., Xu, K. & Xu, L. Dataset for analyzing the horizontal transmission mechanism of agricultural product prices under changes in domestic financial markets (2017-2021). Journal of Agricultural Big Data, Vol. 5(2023) No. 3, p. 19-25.

[7] Yan, X. & Qi, C. Research on the long-term price formation and fluctuation of agricultural products based on different attributes. Agricultural Modernization Research, Vol. 36(2025) No. 5, p. 790-795.

[8] Hang, X. & Zhu, Z. Analysis on the impact of distributors' market power on agricultural product price fluctuations in Tianjin. Tianjin Agricultural Science, Vol. 28(2022) No. 4, p. 40-45.

[9] Song, C., Xiao, X. & Li, C. Research on the impact of local policy-based agricultural insurance on the agricultural product market: A case study of the vegetable market in Wuhan. Journal of Agricultural Modernization, Vol. 42(2025) No. 1, p. 85-93.

[10] Zhou, J., Xu, Y. & Li, C. Mechanism and empirical evidence of the impact of food safety incidents on the price fluctuations of livestock and poultry products. Journal of Agricultural Modernization, Vol. 40(2025) No. 2, p. 282-289.