Proceedings of ICFTBA 2024 Workshop: Finance in the Age of Environmental Risks and Sustainability
DOI: 10.54254/2754-1169/94/2024OX0204

Metal Futures Prices Prediction Based on LSTM-GRU Hybrid Model

Boang Liu[1,a,*]

[1]*Queen Mary School Hainan, Beijing University of Posts and Telecommunications, County Street 663, Ling Shui, China*
a. 2022lba@bupt.edu.cn
**corresponding author*

Abstract: Accurately predicting the prices of metal futures such as aluminum, copper, lead, nickel, tin, and zinc is crucial for investors, manufacturers, and policymakers due to their significant impact on economic activities and industrial processes. This paper presents a comprehensive study on the prediction of futures prices using a hybrid model combining Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) neural networks. The research utilizes historical data from 2014 to 2024, including daily open, high, low, close prices, trading volumes, and return. Data preprocessing involved interpolation for missing values and normalization using MinMaxScaler. The model's performance was evaluated with various accuracy metrics, such as Root Mean Squared Error (RMSE), Mean Squared Error (MSE), Mean Absolute Error (MAE) ect, showing high accuracy for lead, tin, and zinc futures in both train and test datasets. However, the model performed well only on the train data for aluminum, copper, and nickel, indicating the need for more iterations, additional factors, or the consideration of unexpected events. The results highlight the efficacy of the LSTM-GRU hybrid model in futures price prediction, with significant implications for informed investment decisions. Future research should consider cross-market analysis and additional factors to enhance model accuracy further.

Keywords: Metal Futures, LSTM-GRU Hybrid Model, Price Prediction, Market Analysis.

1. Introduction

Futures are financial derivatives that allow trading parties to agree to buy or sell a certain amount of underlying assets at a specific point in the future at an agreed price. In particular, metal futures such as aluminum, copper, lead, nickel, tin, and zinc play a crucial role in the global economy. These metals are integral to various industries, including construction, manufacturing, and technology, making their price stability and predictability essential for economic planning and investment strategies. Accurate forecasting of metal futures prices is vital for stakeholders to make informed decisions, manage risks, and optimize their operations.

In financial markets, accurate prediction of metal futures prices is crucial for investors, producers and traders. Traditional time series forecasting methods, such as Auto-Regressive Moving Average Model (ARIMA) and exponential smoothing, often fall short in capturing the complex, non-linear patterns present in financial data. In recent years, deep learning models, particularly those based on recurrent neural networks (RNNs), have shown promise in addressing these challenges.

DOI: 10.54254/2754-1169/94/2024OX0204

LSTM stands for Long Short-Term Memory. It was created by Hochreiter & Schmidhuber and later developed and popularized by many researchers [1]. It's a type of RNN architecture used in the field of deep learning. LSTMs are designed to overcome the limitations of traditional RNNs, particularly the problem of long-term dependencies, where the influence of a given input diminishes over time [2]. GRU is a variant of RNN in deep learning, proposed by KyungHyun Cho et al. in 2014. Similar to Long Short-Term Memory Networks, GRU is also designed to address the long-term dependency problem in traditional RNNs [3]. LSTM models are particularly well-suited for time series forecasting due to their ability to learn long-term dependencies, making them ideal for predicting the highly volatile future market.

This study explores the application of an LSTM-GRU hybrid model for predicting futures prices, leveraging the strengths of both architectures to ensure predictive performance. This model has two advantages. On one hand, the LSTM-GRU hybrid model could be designed to handle different types of dependencies. For example, LSTM could be used to handle long-term dependencies, and GRU could be used to handle the short-term dependencies. On the other hand, GRUs are more computationally efficient than LSTMs. The hybrid model could use a combination of LSTM and GRU units within the same network, allowing the network to employ LSTM units where long-term memory is required and GRU units where computational efficiency is needed.

2. Literature Review

2.1. Time Series Forecasting in Financial Markets

Time series forecasting in financial markets has been extensively studied. Classical methods, including ARIMA, have been widely used but often lack the ability to handle non-linearities in the data. Machine learning approaches, such as support vector machines (SVM) and decision trees, have provided some improvements but still face limitations in dealing with sequential dependencies [4]. For example, SVM and decision trees are typically designed for static and structured data, making it difficult to directly process sequential data such as text, time series, and other time-dependent data. And these methods often struggle to capture long-term dependency relationships because they do not have built-in memory mechanisms to effectively process time series information in data.

2.2. Deep Learning Approaches

The advent of deep learning has revolutionized time series forecasting. LSTM and GRU, both variants of RNNs, have been particularly effective in capturing long-term dependencies in sequential data. LSTM networks, introduced by Hochreiter and Schmidhuber, use a gating mechanism to control the flow of information, thereby addressing the vanishing gradient problem. GRUs, proposed by Cho et al, simplify the LSTM architecture while maintaining similar performance [5].

2.3. Hybrid Models

Hybrid models combining different neural network architectures have been explored to leverage their complementary strengths. For instance, combining CNNs with RNNs has been effective in various domains, such as Natural Language Processing (NLP), computer vision and stock price prediction [6]. However, the combination of LSTM and GRU for financial time series forecasting remains relatively underexplored. As is known, LSTM effectively preserves and utilizes the long-term information of historical data by introducing memory units and gating mechanisms such as input gates, forget gates, and output gates. GRU also uses gating mechanisms such as update gates and reset gates, which can efficiently capture short-term dependencies in time series data. This study aims to

DOI: 10.54254/2754-1169/94/2024OX0204

fill this gap by investigating the performance of an LSTM-GRU hybrid model in predicting metal futures prices.

3. Data and Methodology

3.1. Data Collection and Preprocessing

This paper utilizes historical data for aluminum, copper, lead, nickel, tin, and zinc futures from 2014 to 2024, all the data comes from https://cn.investing.com/commodities/metals which is a professional website for investors. The dataset includes daily open, high, low, close prices, volume of trades and return of each day. The data was preprocessed to ensure consistency and completeness. Missing values were handled through interpolation, and features were normalized using MinMaxScaler to facilitate model training. In addition, the author established several 3X2 regions to display the close price, return and return distribution of six futures, showing as Figure 1 and Figure 2. From the Figure 1, it can be seen that all futures' close prices experience rapid growth in early 2022. According to news reports, the geopolitical conflict between Russia and Ukraine erupted at that time. Affected by this event, the market's concerns about supply chain disruptions continue to heat up, and commodities related to conflicts such as copper, nickel, aluminium, etc. are highly sought after by capital. Figure 2 indicates that nickel exhibits the smallest return fluctuation. The return fluctuations of the other five futures prices are mostly between minus 0.05 and plus 0.05.

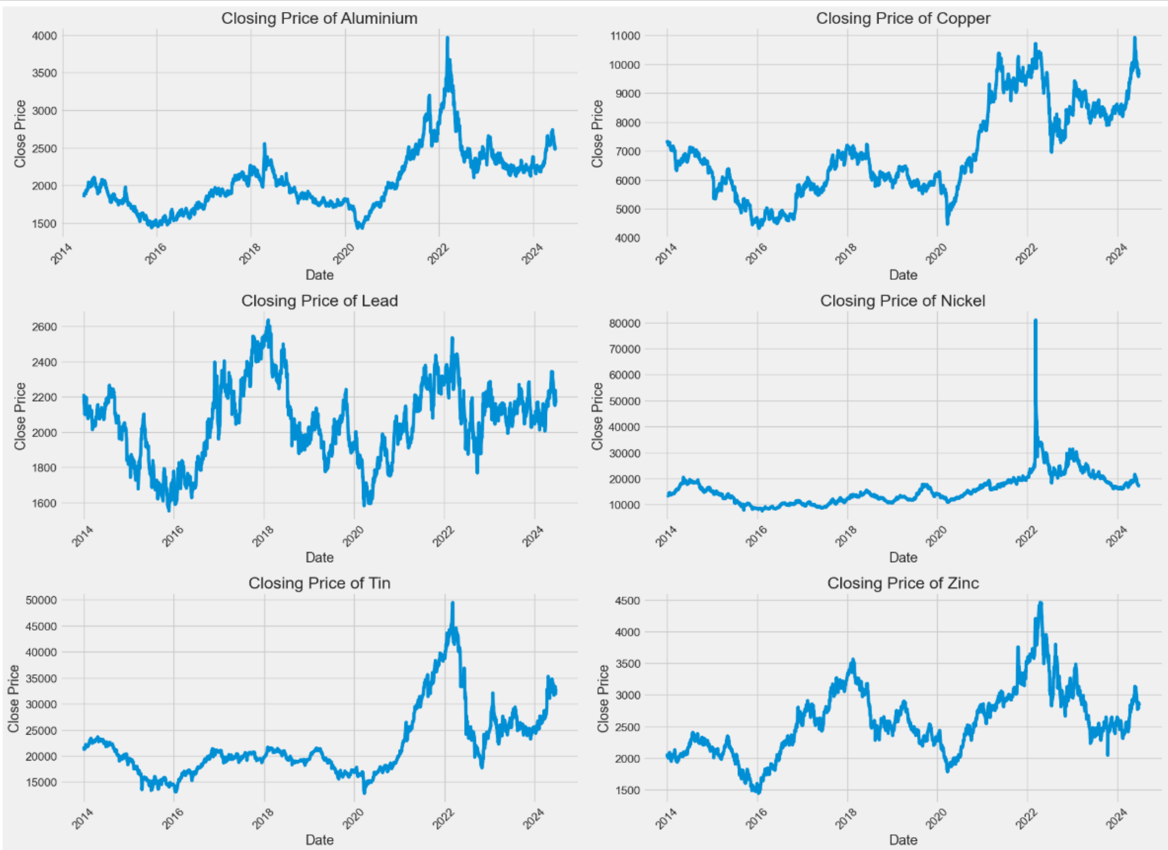

Figure 1: Close Price of Aluminium, Copper, Lead, Nickel, Tin, Zinc Futures

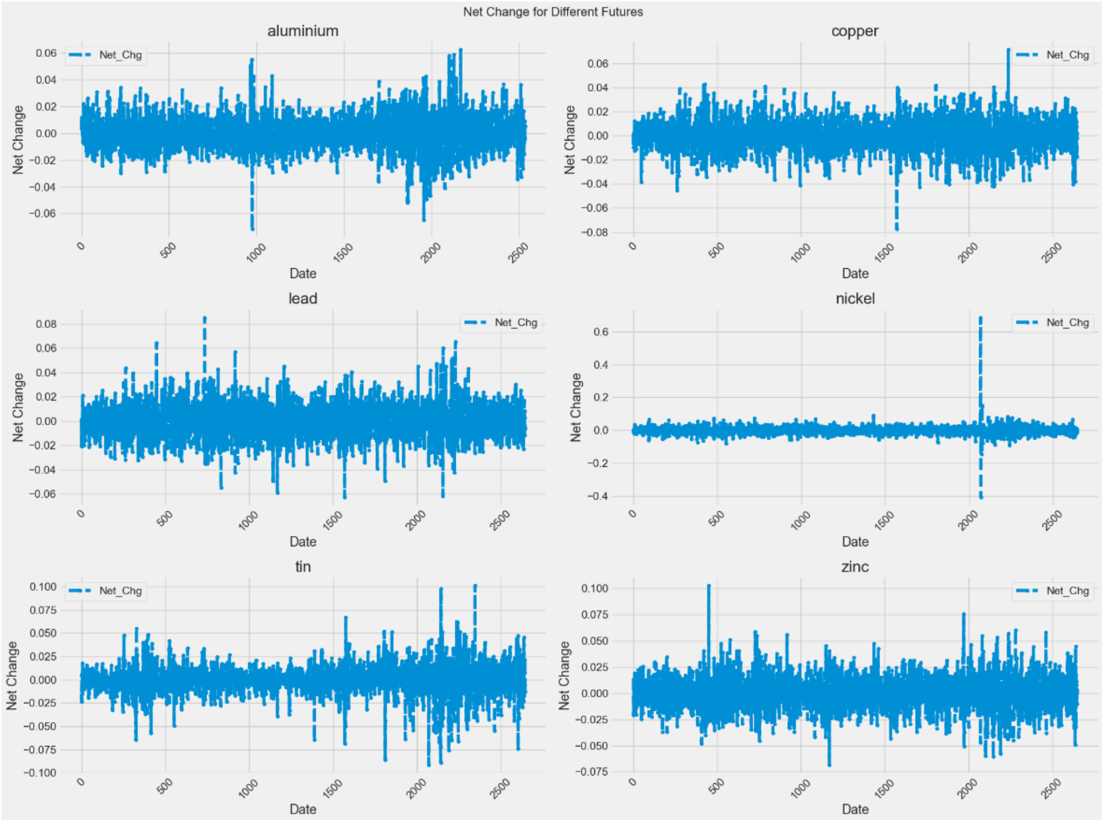

Figure 2: Return of Aluminium, Copper, Lead, Nickel, Tin, Zinc Futures

What's more, the author put the close prices of six types of futures into a plot and use different colors to represent different futures for easy observation of their changes, such as green line means copper and yellow line means tin, showing as figure 3. From Figure 3, it shows that the aluminium, zinc, copper and lead futures were quite stable during last ten years. Instead, the tin and nickel futures changed rapidly during last decade, particularly between 2020 and 2023. Maybe due to the COVID-19 epidemic, the soaring demand for tin and nickel has led to a rapid increase in their prices.

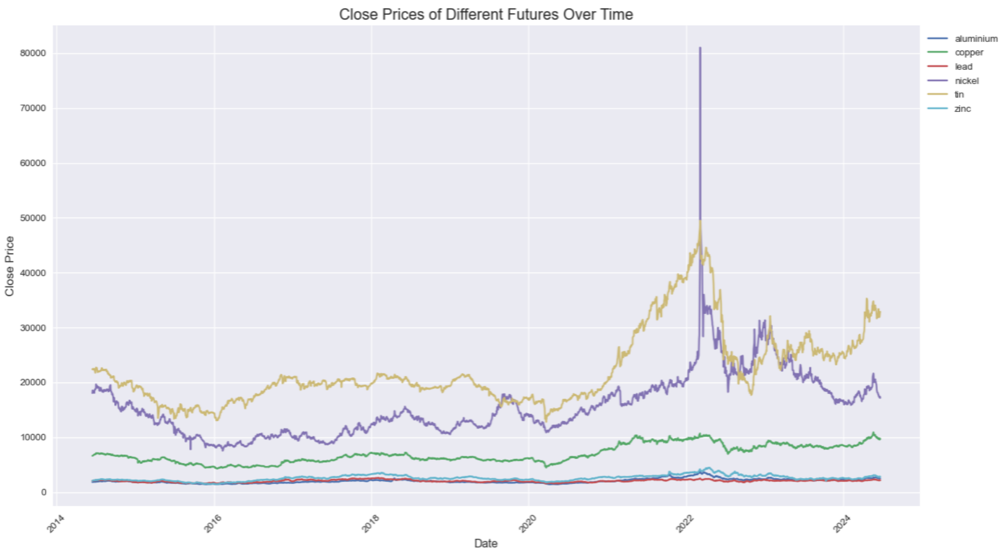

Figure 3: Close Price of Different Futures over Time

DOI: 10.54254/2754-1169/94/2024OX0204

Two heatmaps of the correlation matrix were drawn using the seaborn library to visually display The correlation between six futures closing prices and the correlation between six futures returns, showing as Figure 4 and Figure 5 below. The color intensity of the heatmap indicates the strength of the correlation, with the redder of color showing the stronger correlation of two futures. If the color intensity is 1, it shows red color and means the two futures are highly positively correlated. But if the color intensity is -1, it shows blue color and means the two futures are highly negatively correlated.

Based on Figure 4, which shows the correlation matrix of close price for different futures, it indicates that the close price of copper and aluminium, tin and aluminium, zinc and aluminium, copper and tin has very high correlation. And the close price between nickel and lead has lowest correlation.

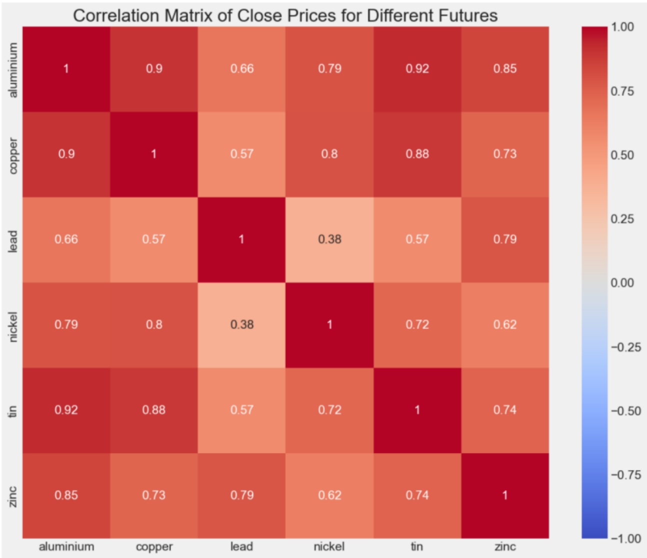

Figure 4: Correlation Matrix of Close Prices for Different Futures

Figure 5 shows the correlation matrix of returns for different futures. It indicates that the net changes between these six futures are not very related. There is a certain correlation only between lead and copper, tin and copper, but not strong.

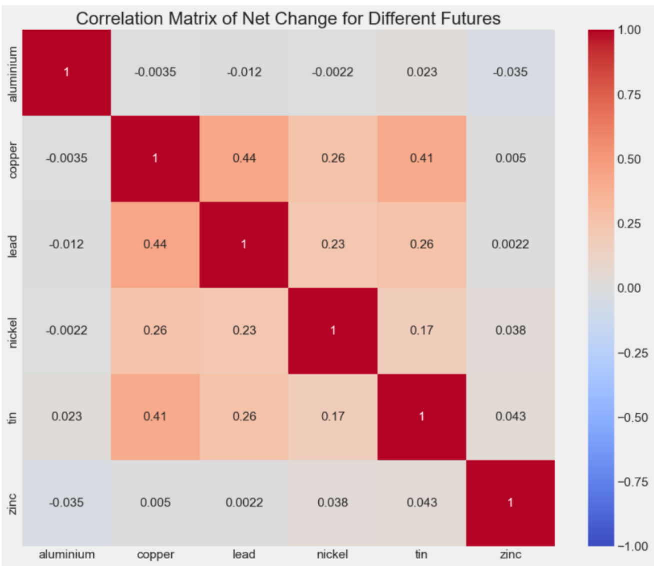

Figure 5: Correlation Matrix of Returns for Different Futures

DOI: 10.54254/2754-1169/94/2024OX0204

3.2. Model Architecture

Table 1 below shows the summary information of the hybrid model, including the name of each layer, output shape, number of parameters, and total parameters. The hybrid model combines two LSTM layers followed by two GRU layers, and a dense layer at last. This architecture is designed to capture both short-term and long-term dependencies in the data. It starts with two LSTM layers to capture long-term dependencies. Then two subsequent GRU layers are added to capture additional patterns. Finally, a dense layer at the end produces the final output.

Table 1: Summary Information of the Model

Layer(type)	Output Shape	Param #
lstm(LSTM)	(None, 15, 32)	4352
lstm_1(LSTM)	(None, 15, 32)	8320
gru(GRU)	(None, 15, 32)	6336
gru_1(GRU)	(None, 32)	6336
Dense (Dense)	(None, 1)	33
Total params		25377

3.3. Model Training

The author divided the dataset into two parts, 60% data for training and the rest 40% data for testing. The author uses the fit function in Keras to train a neural network model, which iterates 100 times over the entire training set. At each iteration, the number of samples used by the model to update weights is 5. There are several benefits of using a mini batch of data with a size of 5. Firstly, small batch training is more efficient than single sample training. By processing multiple samples simultaneously, the parallel computing power of modern hardware such as GPUs can be utilized to accelerate model training speed. Secondly, updating weights with the average gradient of multiple samples in each iteration helps reduce variance during training and makes the training process more stable. This stability can help the model converge to a better solution faster. Thirdly, small batch training helps the model better generalize to unseen data. By calculating gradients and updating weights on multiple samples, the model is more likely to learn the overall features of the dataset rather than the specific noise of individual samples.

3.4. Model Evaluation

The performance of hybrid LSTM-GRU models is assessed using various evaluation metrics, including RMSE, MSE, MAE, Variance Regression Score, R2 Score, Mean Geometric Deviation (MGD) and Mean Percentage Deviation (MPD). By evaluating these metrics for both training and testing data, insights into the LSTM and GRU models' performance in terms of accuracy, error magnitude, and predictive power can be gained. These metrics collectively provide a comprehensive assessment of how well the models generalize to unseen data and handle the prediction task at hand.

3.4.1. RMSE

RMSE measures the average magnitude of the errors between predicted values and actual values. It penalizes larger errors more heavily, giving a sense of the model's overall accuracy [7]. Lower RMSE values indicate better model performance in terms of predictive accuracy.

DOI: 10.54254/2754-1169/94/2024OX0204

3.4.2. MSE

MSE calculates the average squared differences between predicted and actual values [8]. It provides a measure of the model's prediction error. Like RMSE, lower MSE values signify better model performance with smaller prediction errors.

3.4.3. MAE

MAE computes the average absolute differences between predicted and actual values [9]. And it gives a direct measure of average prediction error magnitude. Also like RMSE, lower MAE values can indicate better predictive accuracy, providing insight into the model's average performance in terms of absolute errors.

3.4.4. Variance Regression Score

The variance score calculates the proportion of the variance in the dependent variable that is predictable from the independent variable. When a score closer to 1 indicates a stronger correlation and better prediction effect [10].

3.4.5. R2 Score (Coefficient of Determination):

R2_score quantifies the proportion of the variance in the dependent variable that is predictable from the independent variable [11]. It provides a measure of how well future samples are likely to be predicted by the model. Like the variance score, R2 score closer to 1 indicates better model performance in explaining the variability of the data about the mean.

3.4.6. MGD

MGD calculates the geometric mean of the ratio of predicted to actual values. It's a multiplicative error measure, so a lower MGD indicates better model accuracy, especially for applications sensitive to proportional errors.

3.4.7. MPD

MPD calculates the average percentage deviation between predicted and actual values. It's a measure of relative prediction error, also a lower MPD indicates better model accuracy, particularly in terms of percentage deviation from actual values.

4. Results

4.1. Performance Metrics

The table 2 below displays all the evaluation results for each futures. For example, focus on the R2_score, the R2_store of the training data portion for the six futures is greater than or very close to 0.99. And the R2_store of the testing data portion for the lead, tin and zinc futures is greater than 0.94, while the R2_store of the testing data portion for the aluminium, copper and nickel futures is less than 0.86. The table 2 shows that the hybrid model has a good fitting effect on both the train data and test data of futures of lead, tin and zinc. However, it only has a good fitting effect on the train data of futures of aluminium, copper and nickel, has a poor fitting effect on the test data. Perhaps there are the following reasons: First of all, the number of iterations is too small or the sample space is too small. What's more, perhaps there are too few factors to consider, and additional factors such as politics and economic development need to be added. Last, failure to consider unexpected events.

DOI: 10.54254/2754-1169/94/2024OX0204

The experimental results demonstrate that the effectiveness of the hybrid model in capturing long-term and short-term dependencies.

Table 2: Evaluation Results of Aluminium, Copper, Lead, Nickel, Tin, Zinc Futures

Name	Indicator	Train data	Test data
alu	RMSE	22.04	147.48
alu	MSE	485.55	21750.17
alu	MAE	17.41	71.45
alu	variance_score	0.99	0.88
alu	R2_score	0.99	0.86
alu	MGD	0.00014	0.00268
cop	RMSE	66.81	465.17
cop	MSE	4463.06	216386.62
cop	MAE	49.39	361.96
cop	variance_score	0.99	0.93
cop	R2_score	0.99	0.84
cop	MGD	0.00014	0.00258
cop	MPD	0.78	23.60
lead	RMSE	24.50	26.65
lead	MSE	600.44	710.04
lead	MAE	18.22	20.16
lead	variance_score	0.99	0.98
lead	R2_score	0.99	0.98
lead	MGD	0.00014	0.00016
lead	MPD	0.30	0.34
nic	RMSE	250.38	2079.11
nic	MSE	62692.40	4322710.50
nic	MAE	187.13	576.70
nic	variance_score	0.99	0.85
nic	R2_score	0.99	0.85
nic	MGD	0.00043	0.01803
tin	RMSE	205.67	1713.01
tin	MSE	42299.90	2934409.06
tin	MAE	150.81	1004.41
tin	variance_score	0.99	0.96
tin	R2_score	0.99	0.95
tin	MGD	0.00014	0.00217
tin	MPD	2.44	78.85

DOI: 10.54254/2754-1169/94/2024OX0204

Table 2: (continued).

zinc	RMSE	30.43	57.42
zinc	MSE	925.83	3296.87
zinc	MAE	23.18	39.27
zinc	variance_score	1.00	0.99
zinc	R2_score	1.00	0.99
zinc	MGD	0.00016	0.00038
zinc	MPD	0.38	1.11

Figure 6 displays the comparison between original close price of each future and train predicted close price and test predicted close price. The difference between different futures can be clearly and intuitively seen from Figure 6, and the same conclusion can be drawn as Table 2.

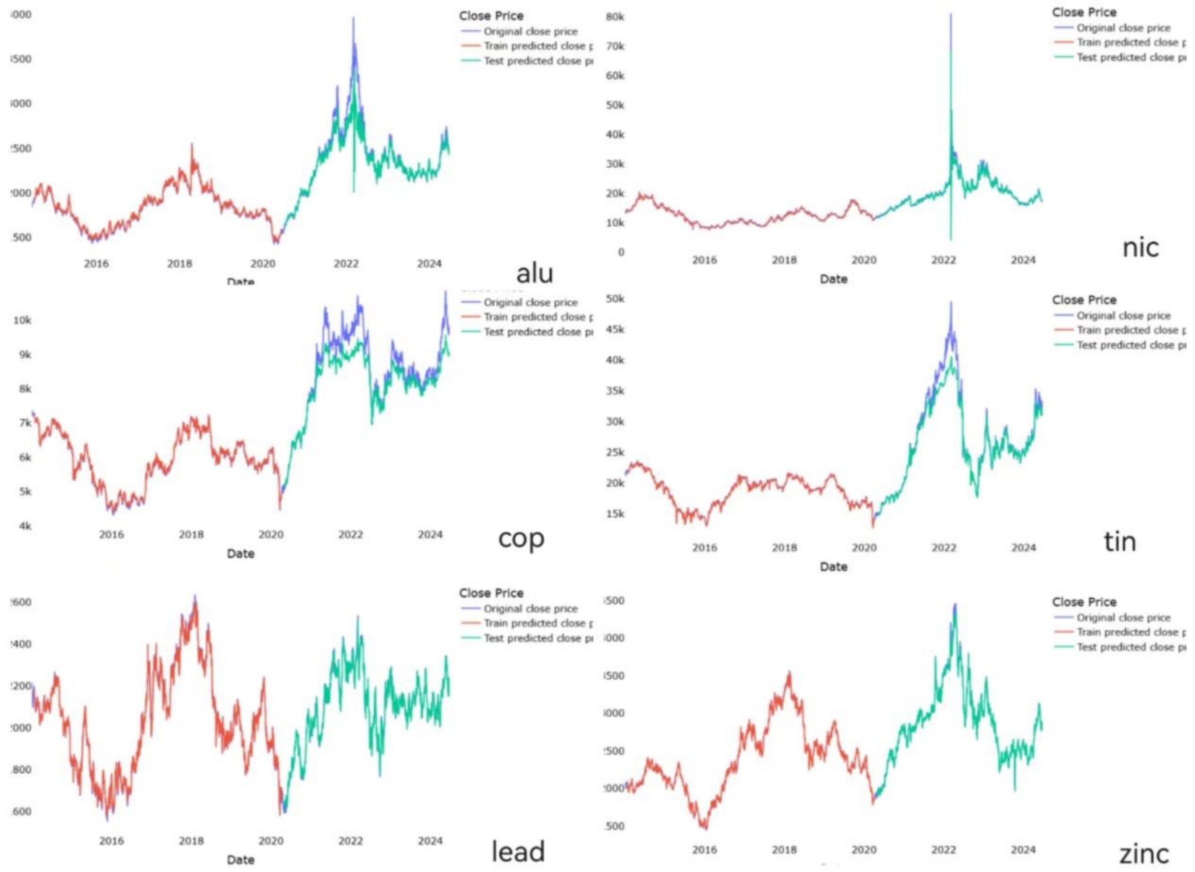

Figure 6: Prediction Price VS Original Price

5. Discussion

5.1. Model Refinement and Optimization

Future work should focus on increasing the number of iterations and expanding the sample space to improve the model's performance, especially for commodities like aluminum, copper, and nickel.

Additionally, exploring advanced optimization techniques and hyperparameter tuning could further enhance predictive accuracy.

5.2. Incorporating Additional Factors

To improve the model's robustness, incorporating more diverse factors such as political events, economic indicators, and unexpected events could be beneficial. This would help in capturing the broader context affecting futures prices.

5.3. Exploring Other Hybrid Architectures

While the LSTM-GRU hybrid model has shown promising results, experimenting with other combinations of neural network architectures, such as integrating CNNs or Attention Mechanisms, could provide further improvements.

5.4. Cross-Market Analysis

Extending the study to include other markets and financial instruments could validate the model's applicability across different domains. Analyzing cross-market interactions and their impact on futures prices could offer deeper insights.

By addressing these areas, future research can build on the findings of this study to develop more accurate and reliable futures price prediction models, ultimately benefiting market participants in making informed investment decisions.

6. Conclusion

This study demonstrates the efficacy of a hybrid LSTM and GRU neural network model in predicting futures prices for various metal futures, including aluminum, copper, lead, nickel, tin, and zinc. By leveraging the strengths of both LSTM and GRU architectures, the hybrid model effectively captures both short-term and long-term dependencies in the data, resulting in high predictive accuracy. The analysis also highlights the significant correlations between the close prices of different futures, offering valuable insights for market participants. However, the model's performance varies across different commodities, suggesting that further refinements are necessary to enhance its generalizability. Future research should address issues such as insufficient iterations, limited sample size, and the exclusion of external factors like political and economic developments, which can improve the accuracy of model predictions. Additionally, exploring other hybrid neural network architectures and conducting cross-market analysis are also crucial to enhancing the model's robustness and applicability. Overall, the LSTM-GRU hybrid model presents a promising approach for futures price prediction, assisting market participants in making informed investment decisions.

References

[1] *Ghosh, S., Examining crude oil price - Exchange rate nexus for India during the period of extreme oil price volatility. Applied Energy, 166, 2016, 113-120.*

[2] *Shi, P., Song, S., Lang, Q., et al., Futures Price Forecasting Based on LSTM Model Using Public Opinion Sentiment Analysis. In: Technical Committee on Control Theory, Chinese Association of Automation (Eds.), Proceedings of the 42nd Chinese Control Conference (15), 2023, 7,7-12.*

[3] *Lei, Z., Chen, Y., Bouri, E., Time-varying jump intensity and volatility forecasting of crude oil returns. Energy Economics, 70, 2018, 40-49.*

[4] *Longyue, L., Ying, H., Research on short-term forecasting methods of gold futures prices—Analysis of COMEX gold futures price data based on CEEMDAN and LSTM models. Price Theory and Practice, 2023, 09, 164-168.*

DOI: 10.54254/2754-1169/94/2024OX0204

[5] Gong, H., Xing, H., Yu, Y., Liang, Y. A combined model using secondary decomposition for crude oil futures price and volatility forecasting: Analysis based on comparison and ablation experiments. Expert Systems With Applications, 120(2019), 370-379.

[6] Researchers from University of Oviedo Detail New Studies and Findings in the Area of Support Vector Machines. PM10 concentration forecasting in the metropolitan area of Oviedo (Northern Spain) using models based on SVM, MLP, VARMA and ARIMA: A … Ecology Environment & Conservation, 24(2), 2018, 312-320.

[7] Wei, W. Research on Multivariate Futures Price Forecasting Based on IMAT-LSTM Model. Academic Journal of Computing & Information Science, 6(1), 2023, 15-26.

[8] Ceperic, E., et al. Machine Learning; Findings from E. Ceperic et al Has Provided New Information about Machine Learning (Short-term forecasting of natural gas prices using machine learning and feature selection algorithms). Energy Weekly News, 2019, 45-47.

[9] Lianlian, F., Xinsheng, D., Yuehui, D. Ensemble empirical mode decomposition-based preprocessing method with Multi-LSTM for time series forecasting: a case study for hog prices. Connection Science, 31(4), 2019, 379-393.

[10] Sun, J.Y., Bing, G.Y. A new method for predicting Shanghai copper futures prices integrating multi-source data information. Operations Research and Management Science, 1-8 [2024-07-08]. Available at: http://kns.cnki.net/kcms/detail/34.1133.G3.20240628.1434.004.html.

[11] Mostafa, M., Dipu, K.H.M., Farnad, N., Abbas, K. Neural network-based interval forecasting of construction material prices. Journal of Building Engineering, 30, 2020, 101-110.