Print ISSN: 2288-4637 / Online ISSN 2288-4645
doi:10.13106/jafeb.2022.vol9.no10.0159

# An Empirical Analysis of Sino-Russia Foreign Trade Turnover Time Series: Based on EMD-LSTM Model*

**Jian GUO[1], Kai Kun WU[2], Lyu YE[3], Shi Chao CHENG[4], Wen Jing LIU[5], Jing Ying YANG[6]**

**Received: September 15, 2022 Revised: November 26, 2022 Accepted: December 05, 2022**

## Abstract

The time series of foreign trade turnover is complex and variable and contains linear and nonlinear information. This paper proposes pre-processing the dataset by the EMD algorithm and combining the linear prediction advantage of the SARIMA model with the nonlinear prediction advantage of the EMD-LSTM model to construct the SARIMA-EMD-LSTM hybrid model by the weight assignment method. The forecast performance of the single models is compared with that of the hybrid models by using MAPE and RMSE metrics. Furthermore, it is confirmed that the weight assignment approach can benefit from the hybrid models. The results show that the SARIMA model can capture the fluctuation pattern of the time series, but it cannot effectively predict the sudden drop in foreign trade turnover caused by special reasons and has the lowest accuracy in long-term forecasting. The EMD-LSTM model successfully resolves the hysteresis phenomenon and has the highest forecast accuracy of all models, with a MAPE of 7.4304%. Therefore, it can be effectively used to forecast the Sino-Russia foreign trade turnover time series post-epidemic. Hybrid models cannot take advantage of SARIMA linear and LSTM nonlinear forecasting, so weight assignment is not the best method to construct hybrid models.

**Keywords:** Sino-Russia Foreign Trade Turnover, Time Series, Weight Assignment Hybrid Model, Forecast Accuracy

**JEL Classification Code:** C32, F14, F17, P45

## 1. Introduction

China and Russia have maintained close strategic collaboration in international and regional affairs, which has strongly maintained peace and stability in the region and the world. China has held on to its position as Russia's top trading partner for 12 consecutive years, and Russia ranks 10th among China's major trading partners. Russia is a major energy exporter in the world, while China is a leading industrial power in the world, importing large amounts of energy from Russia. There is no doubt that energy is a key part of the bilateral trade and economic relations between China and Russia. Over the last 10 years, China's GDP has ranked approximately 2nd and Russia's at 11th all over the world. The economic sanction of European and American countries is the most direct factor limiting the growth of Russia's GDP and foreign trade turnover.

To ease the stress on the Russian economy from the economic sanctions of Europe and American countries, the "Look East" strategy has been proposed and implemented, which has provided an opportunity to strengthen economic and trade cooperation with East Asian countries. So, Russia strengthened its economic and trade cooperation with China. Sino-Russia foreign trade turnover from 2013 to 2020 was

***Acknowledgments:**
[1] The authors want to thank our editor and reviewers for their valuable comments and advice.
[2] This research was funded by the China Scholarship Council (grant 201908090322).
[1]First Author and Corresponding Author. Faculty of Economics, Peoples' Friendship University of Russia, Russia [Postal Address: Miklukho-Maklaya 6, Moscow, 117198, Russia]
Email: guojianels@126.com
[2]Institute of Economics and Management, Lviv Polytechnic National University, Ukraine [Postal Address: 5 Mytropolyta Andreia St., Lviv, 999146, Ukraine] Email: juziwukaikun@163.com
[3]Institute of Industrial Management, Economics, and Trade, Peter the Great St. Petersburg Polytechnic University, Russia [Postal Address: Politekhnicheskaya Ulitsa, 29, St Petersburg, 195251, Russia] Email: lyu.ye.91@mail.ru
[4]Faculty of Economics, Peoples' Friendship University of Russia, Russia [Postal Address: Miklukho-Maklaya 6, Moscow, 117198, Russia] Email: dzftg12345@gmail.com
[5]Faculty of Economics, Moscow State University, Russia [Postal Address: Ulitsa Kolmogorova, 1, Moscow, 119991, Russia]
Email: liuw4311@gmail.com
[6]Faculty of Economics, Peoples' Friendship University of Russia, Russia [Postal Address: Miklukho-Maklaya 6, Moscow, 117198, Russia] Email: unnayang@yandex.ru

$714.9 billion. The main export and import commodities are minerals (33% of total trade), machinery, and equipment (29% of total trade) (RU-STAT, 2021). Energy cooperation is always the most important, fruitful, and wide-ranging area of practical cooperation between China and Russia. Among Russian major trade partners, China is the most essential one, and the "Look East" strategy shows remarkable results. The Sino-Russia foreign trade turnover time series, which is the main external expression of the import and export markets, presents a very complex operating mechanism. At the same time, the time series of foreign trade turnover implicitly shows the current situation and future change trend of Sino-Russia foreign trade development. Therefore, it is particularly important to understand its fluctuation pattern to optimize the structure of import and export trade, increase foreign exchange reserves and create more employment in economic activities.

## 2. Literature Review

Traditional time series models include the grey prediction model (GM(1,1)), holt-winters (HW), auto-regressive (AR), moving average (MA), vector auto-regression (VAR), autoregressive moving average (ARMA) model, seasonal auto-regressive integrated moving average (SARIMA) models, etc. Scholars use traditional time series models to conduct in-depth studies of foreign trade turnover forecasts. For instance, Zhang et al. (2013) improved on the traditional GM(1,1) model to increase the forecast accuracy and finds that there is still much potential for China-Russia trade cooperation. The ARIMA model is a classical time series analysis method with high forecast accuracy in the short term and has been widely used in the field of foreign trade markets. The ARIMA model is used by Farooqi (2014) to forecast Pakistan's total exports and imports from 1947 to 2013, and he finds that the ARIMA (2, 2, 2) model is suitable for forecasting annual imports and the ARIMA (1, 2, 2) model is suitable for forecasting annual exports.

The application of the ARIMA model greatly contributed to forecasting research in financial markets and itself became a classic control model for various new forecasting methods (Kim &Won, 2018; Suh et al., 2014). Mladenovic et al. (2016), used the seasonal Holt-Winters and ARIMA methods to analyze export trend changes for the Republic of Serbia. It is confirmed that the two methods are consistent in predicting export trend changes, which provides a valid reference method for the analysis of other macroeconomic indicators. In another example, Urrutia et al. (2019) used the ARIMA model and the Bayesian Artificial Neural Network (BANN) model to forecast the turnover of Philippine imports and exports, as well as to compare the predictive performance of the two models. They conclude that there is no significant difference between the actual and predicted values of Philippine imports and exports, so the BANN model is the most appropriate model for forecasting the turnover of Philippine imports and exports.

With the research of deep learning, LSTM neural networks alleviate the gradient disappearance and explosion phenomenon that easily occurs during the training process of traditional recurrent neural networks (RNN) (Gers et al., 2000), and the forecasting advantage in foreign trade time series is gradually highlighted. Many scholars have studied foreign trade turnover time series by using LSTM models. Qu et al. (2019) proposed forecasting the trend of trade turnover using a recurrent neural network (RNN) based on LSTM. The MSE exceeds that of the traditional time series model at 12.2% and proves the effectiveness of the LSTM-RNN model. The LSTM neural network is also applied in fields such as agricultural crop yield forecasting (Haider et al., 2019) and stock price forecasting (Nelson et al., 2017; Gyamerah & Korda, 2021), which shows superior predictive performance in comparison with other machine learning methods.

Because of the complexity and cyclical characteristics of foreign trade turnover time series, single models have gradually revealed shortcomings in long-term forecasting. To compensate for the shortcomings of single models, various combined models have been developed. Yu et al. (2022) combined the SARIMA model with the LSTM model and innovatively constructed the SALSTM model to forecast economic indicators related to the competitiveness of export products and proved its effectivity and superiority by the mean square error (MSE), which provides a reference basis for national trade policy formulation. In another study, the ARIMA-LSTM hybrid model is used by Dave et al. (2021) to forecast the turnover of Indonesian exports. In addition, the ARIMA-LSTM hybrid model is also applied to forecast stock prices (Alzheev & Kochkarov, 2020; Zainuri et al., 2021). Zhang and Sun (2019) examined the combined method of hybrid models and constructed the ARIMA-SVM hybrid model by the weight assignment method to forecast and analyze Chinese foreign trade turnover from 1980 to 2016. The result shows that the forecast accuracy of the hybrid model is higher than that of the single model.

Although the LSTM model demonstrates superiority in the forecasting of time series, because the series is not transformed from non-stationary to stationary, the forecast accuracy still needs to be improved. The empirical mode decomposition (EMD) method can decompose nonlinear and non-smooth data into smooth data, which is a good selection for data preprocessing. The original time series is decomposed by EMD into a fixed number of intrinsic mode functions (IMFs) at different frequencies, and then the decomposed signal is simulated using other models. In the end, all the simulated results are superimposed to improve the forecast accuracy of the hybrid models (Li et al, 2018; Solanki et al., 2020). As an example, Qin et al. (2018) use the hybrid model of integrated empirical modal decomposition (EEMD) and support vector regression (SVR) to forecast the Chinese consumer price index (CPI) and verify that the hybrid model has higher forecast accuracy than the EEMD single model. Shu and Gao (2020) use EMD and a convolutional neural network (CNN) to preprocess the original stock price series, which is modeled by the LSTM neural network and obtain the final predictive price after linear transformation. The hybrid network can achieve better performance by modeling different frequencies than other models. However, there are fewer studies that combine the SARIMA model with the EMD-LSTM model to forecast trade turnover time series.

To summarize the above studies, it can be found that the SARIMA model has advantages in short-term forecasting and that the LSTM neural network algorithm has more potential for application in the economic field. The hybrid model has a higher forecast accuracy than the single model. Therefore, this paper proposes the following hypotheses:

***H1:*** *The EMD algorithm can effectively resolve the hysteresis phenomenon generated by the LSTM model and improve the forecast accuracy;*

***H2:*** *The weight assignment hybrid model combines the linear forecasting advantage of the SARIMA model with the nonlinear forecasting advantage of the EMD-LSTM model, so it has the highest forecast accuracy of all models;*

***H3:*** *Weight assignment is an efficient method to construct the hybrid model of foreign trade turnover time series.*

## 3. Data and Model Specification

### 3.1. Data Source and Assessment Method

This paper uses monthly historical data on Sino-Russia foreign trade turnover from 2013 to 2021 as a research sample; the December 2021 trade turnover is not yet available, so the 2021 average turnover is used instead. The foreign trade turnover from 2013–2019 is the training set, and the rest is the test set. The ratio of the training set to the test set is 7:2, and the predicted values of the test set are compared with the actual values for validation. The mean square error (MSE), mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE) are used as indicators for the evaluation of the model. The mathematical formula is:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2 \tag{1}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{2}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2} \tag{3}$$

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{4}$$

where $y_i$ is the actual value, $\hat{y}_i$ is the predicted value, and $n$ is the sample size.

### 3.2. Relevant Models

#### 3.2.1. The SARIMA Model

The SARIMA model is a combination of seasonal differencing and an ARIMA model for forecasting time series data with cyclical characteristics (John & Patrick, 2016). The model expression is:

$$\text{SARIMA} = (p, d, q) \times (P, D, Q)s \tag{5}$$

where $p$ is the auto-regressive orders, $P$ is the seasonal auto-regressive orders, $d$ is the differential counts, $D$ is the seasonal differential counts, $q$ is the moving average orders, $Q$ is the seasonal moving-average orders, and $s$ is the number of periods.

The minimum Akaike information criterion (AIC) (Akaike, 1974) can effectively compensate for the subjectivity of the fixed order of auto-correlation function (ACF) and partial auto-correlation function (PACF) plots and can find the optimal hyper-parameters more quickly within a limited range of orders. In this paper, the model parameters are determined by the AIC values, and the mathematical formula is:

$$\text{AIC} = 2k - 2\ln(\hat{L}) \tag{6}$$

where $k$ is the number of parameters and $\hat{L}$ is the maximum likelihood function.

#### 3.2.2. The EMD Algorithm

The methods for transforming non-stationary time series into stationary time series include log transformation, exponential smoothing, difference, and decomposition. The EMD algorithm is a signal processing method that converts non-stationary and nonlinear data into stationary and linear data and is an important tool for exploring the hidden time series relationships in the data. It does not require a predetermined basis function and relies on the signal itself to adaptively decompose the IMF and residual terms for the time series (Li et al., 2019). The formula is:

$$X(t) = \sum\nolimits_{i=0}^{n} C_i(t) + R_n(t) \tag{7}$$

where $X(t)$ is the time series, $C_i$ is the IMF, and $R_n$ and is the residual term, which usually represents the overall trend of the time series.

### 3.2.3. The LSTM Model

The LSTM model adds an input gate, output gate, and forgetting gate to the neurons in the RNN. The forgetting gate is mainly used to decide which information needs to be retained and which information needs to be updated, so the training results are reflected in the cell state and output information. This structure is a better solution to the problem of gradient disappearance or gradient explosion in RNNs during training (Figure 1).

The data first pass through the forget gate, which determines the forgotten and retained information in the cell state $C_{t-1}$ at moment $t-1$. The activation function $\sigma$ maps $[h_{t-1}, x_t]$ to 0 and 1, where 0 means "completely discard" and 1 means "completely retain" (Ni et al., 2017). The output value $f_t$ of the activation function, $\sigma$ determines whether $[h_{t-1}, x_t]$ is forgotten or not. The mathematical expression for $f_t$ is:

$$f_t = \sigma\,(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{8}$$

where $h_{t-1}$ is the output information at moment $t-1$, $x_t$ is the input information at moment $t$, $W_f$ is the weight of $[h_{t-1}, x_t]$, $b_f$, is the bias amount, and σ is the activation function.

The data pass through the input gate, and the valid information contained in $[h_{t-1}, x_t]$ is determined. The output value $i_t$ of the activation function, σ determines the proportion of valid information. The amount of valid information in $[h_{t-1}, x_t]$ is calculated using the function to generate a new vector of candidate values $\tilde{c}_t$ (Song et al., 2020). The mathematical expressions for $i_t$ and $\tilde{c}_t$ are:

$$i_t = \sigma\,(W_t\,.\,[h_{t-1}, x_t] + b_i) \tag{9}$$

$$\tilde{c}_t = \tanh\,(W_c\,.\,[h_{t-1}, x_t] + b_c) \tag{10}$$

Update the cell state $C_{t-1}$ at moment $t-1$, and acquire the cell state $C_t$ at moment $t$. The mathematical expression for $C_t$ is:

$$C_t = f_t C_{t-1} + i_t \tilde{c}_t \tag{11}$$

where $f_t C_{t-1}$ is the retained information, and $i_t \tilde{c}_t$ is the added new information.

The data pass through the output gate, and the output value $o_t$ of the activation function σ a determines the proportion of the output information, and $C_t$ passes through the function to calculate the output information. The product of $\tanh(C_t)$ and $o_t$ is the output information $h_t$ at moment $t$. The mathematical equations for $o_t$ and $h_t$ are as follows:

$$o_t = \sigma\,(W_o\,.\,[h_{t-1}, x_t] + b_o) \tag{12}$$

$$h_t = o_t \tanh(C_t) \tag{13}$$

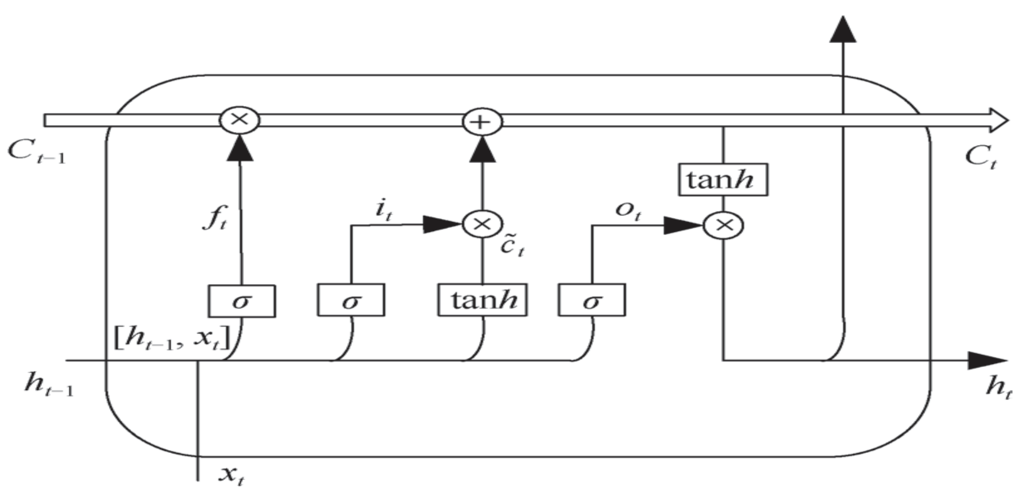


**Figure 1:** The Structure of the LSTM Neural Network

#### 3.2.4. Hybrid Model

The optimal hyperparameters of the SARIMA model were determined by grid search with cross-validation (GridSearchCV), and the predicted values of the SARIMA model were obtained. The EMD algorithm is used as a preprocessing method for signal feature extraction, and each IMF component is used as input information for the LSTM model to obtain the predicted values of the EMD-LSTM model. The SARIMA-EMD-LSTM hybrid model is construed by weight assignment (Figure 2). The mathematical formulation of the hybrid model is:

$$Y = a \,.\, Y_1 + b \,.\, Y_2 \tag{14}$$

where $Y$ is the forecasting value of the SARIMA-EMD-LSTM hybrid model, $Y_1$ is the forecasting value of the SARIMA model, $Y_2$ is the forecasting value of the EMD-LSTM model, and $a$, $b$ is the weighting factor.

The weight factors a and b are determined based on the MAE values of the EMD-LSTM model and SARIMA model, which take values in the range [0,1], $a + b = 1$, and the formula is:

$$a = \frac{\mathrm{MAE}_{\mathrm{SARIMA}}}{\mathrm{MAE}_{\mathrm{SARIMA}} + \mathrm{MAE}_{\mathrm{EMD-LSTM}}} \tag{15}$$

$$b = \frac{\mathrm{MAE}_{\mathrm{EMD-LSTM}}}{\mathrm{MAE}_{\mathrm{SARIMA}} + \mathrm{MAE}_{\mathrm{EMD-LSTM}}} \tag{16}$$

### 3.3. Statistical Treatment

This paper uses Python coding, and the IDE uses Python 2021 software. The SARIMA model is constructed by the arima_model package, and the EMD-LSTM model is constructed by the keras package and the PyEMD package.

## 4. Results and Discussion

### 4.1. Results of the SARIMA Model

The Seasonal and Trend decomposition using Loess (STL) method is used to decompose the Sino-Russia foreign trade turnover time series to obtain the decomposition values of the seasonal effect, trend effect, and residual volatility effect. The decomposition result shows that there is a clear trend in the time series, with a downward trend from 2014–2016 and an upwards trend from 2016–2020, which requires a differencing operation. Meanwhile, there is a stable cyclicality in trade turnover, with a cycle of 12 months. The Sino-Russia foreign trade volume time series has both

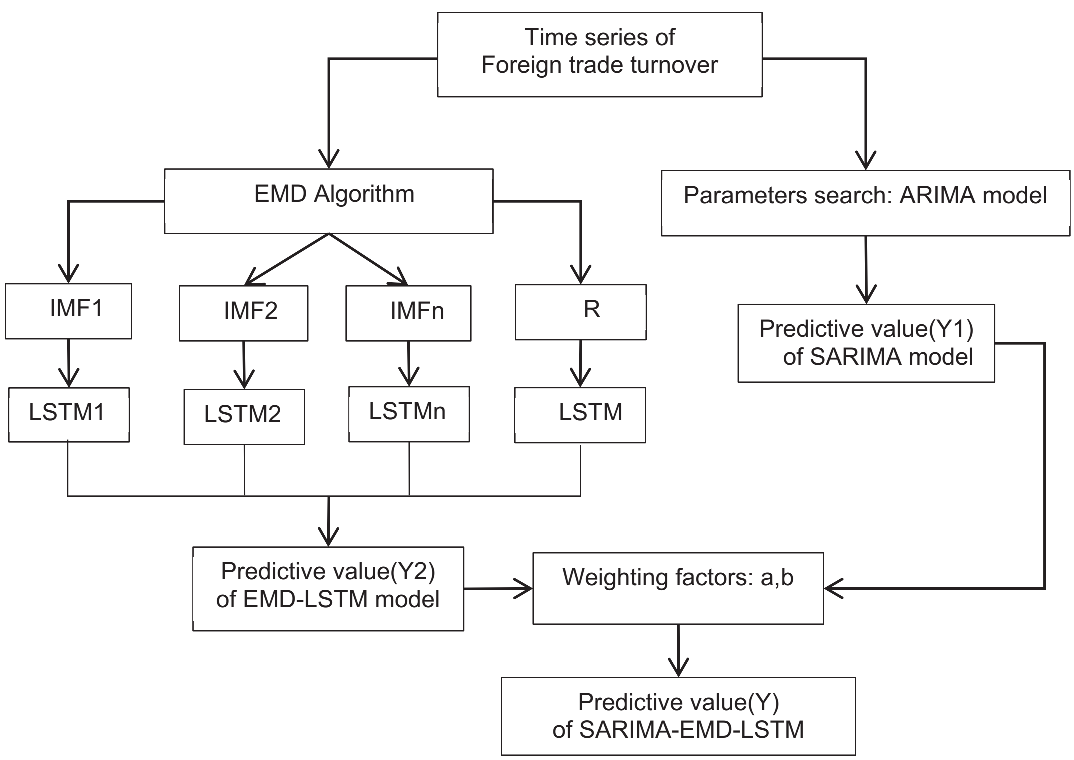


**Figure 2:** Illustration of the SARIMA-EMD-LSTM Hybrid Model Structure

seasonal and nonseasonal information, so it needs to be fitted and forecasted using the SARIMA model. After the second-order differences are applied to the time series, the results of the Augmented Dickey-Fuller test (ADF) and Ljung-Box test (LB) show that they are smooth and have no white noise.

Optimal hyper-parameters are determined by using GridSearchCV in machine learning. The search for parameters is performed in two stages, first determining the optimal parameters for the ARIMA model and then determining the seasonal parameters.

Figure 3 shows that the ACF is trailing, with the auto-correlation falling into the 95% confidence interval at $p = 2$, and the PACF is also trailing, with the auto-correlation falling into the 95% confidence interval at $q = 5$. The range of values for the parameters of the ARIMA model is set as follows: $p \in [1,2]$, $d = 2$, $q \in [1, 5]$. When the parameters are the following options: (1,2,5), (2,2,5), (2,2,4), and (2,2,1), the value of AIC is less than 530. Only the ARIMA(2,2,1) model has all coefficients with $p$ values significant at the 0.05 level. The range of values for the seasonality parameter in the grid search is set as $P \in [0, 4]$, $D = 2$, $Q \in [0, 6]$. When the ARIMA model is added to the seasonal parameters $P = 1$, $D = 2$, $Q = 0$, the model is significantly optimized, and the AIC value is significantly reduced from 529.45 to 377.68. Therefore, the optimal model obtained by the two-stage search for parameters is SARIMA(2,2,1) × (1,2,0,12).

Figure 4 (Normal Q-Q) shows that the residuals are approximately on a straight line, which indicates that the residuals are normally distributed; Figure 4 (Correlogram) shows that all auto-correlations are within the 95% confidence interval. Therefore, the residual series of the SARIMA model is white noise.

As shown in Table 1, all coefficients of the model are statistically significant, and therefore, the SARIMA(2,2,1) × (1,2,0,12) model can be used to predict the Russian trade turnover time series.

## 4.2. The EMD-LSTM Model Results

Each IMF represents the signal functions of each layer after the original signal has been decomposed by EMD. As shown in Figure 6, each IMF satisfies the constraints (Huang et al., 1998).

As shown in Figure 5, IMF1 exhibits strong nonlinear characteristics and contains a large amount of irregular noise. Moreover, it has a high instantaneous frequency and an insignificant periodicity, so it can be regarded as a high-frequency function of the load series. IMF2 demonstrates significant cyclical characteristics, so it can be seen as

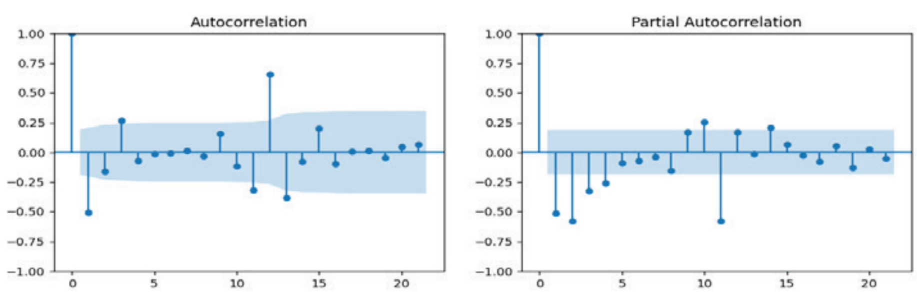


**Figure 3:** ACF and PACF Plots

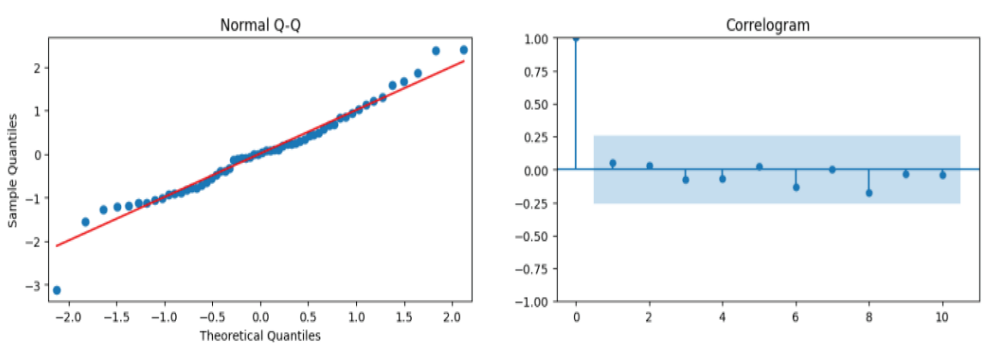


**Figure 4:** The Residual Information of the SARIMA Model

**Table 1:** Statistical Significance Test of SARIMA Model Coefficients

| Parameters | Coef | Std Err | z | P >\| z | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| ar.L1 | −0.5103 | 0.137 | −3.727 | 0.000*** | −0.779 | −0.242 |
| ar.L2 | −0.3926 | 0.149 | −2.626 | 0.009*** | −0.686 | −0.100 |
| ma.L1 | −0.6847 | 0.142 | −4.824 | 0.000*** | −0.963 | −0.406 |
| ar.S.L12 | −0.4686 | 0.143 | −3.274 | 0.001*** | −0.749 | −0.188 |
| sigma2 | 30.4494 | 5.837 | 5.216 | 0.000*** | 44.792 | 41.891 |

Note: Standard errors in parentheses, *$p < 0.1$, **$p < 0.05$, and ***$p < 0.01$. Source: Composed by authors.

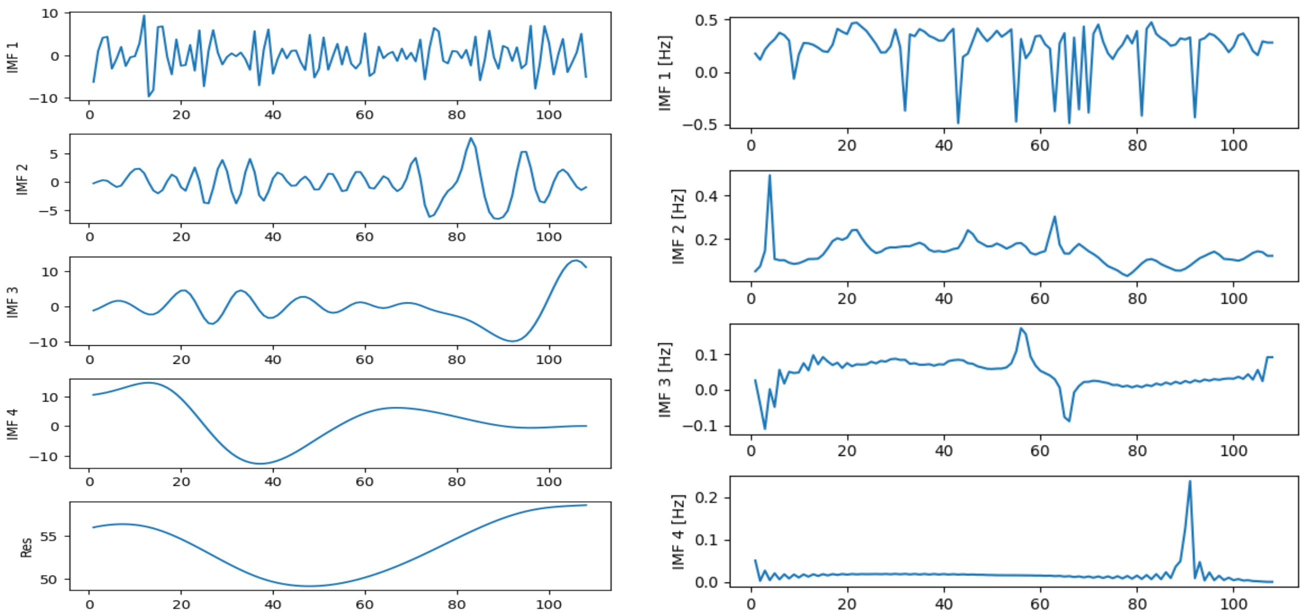


**Figure 5:** Illustration of EMD Decomposition

the cyclical function of the load series. IMF4 has a low instantaneous frequency and an insignificant periodicity, so it can be seen as a low-frequency function of the load series. The residual shows a significant change trend, so it is considered to be the trend function of the load series.

Each IFM and residual is used as input information to the LSTM model for prediction. The dataset is normalized in the range [0,1] using MinMaxScaler from the scikit-learn library. The equation is:

$$X_{\text{norm}}^{(i)} = \frac{X^{(i)} - X_{\min}}{X_{\max} - X_{\min}} \tag{17}$$

The dataset is also transformed into a supervised learning dataset, and the value at moment $t$ is predicted by the value at moment $t$−1. By stacking LSTM hidden layers to increase the depth of the neural network, it can effectively reduce neurons and improve training efficiency and forecast accuracy. In this paper, we use the Keras deep learning framework to quickly build a stacked LSTM neural network model. The sequential model is built by adding an LSTM layer (with 256 units), a dropout layer (Krizhevsky et al., 2017), three LSTM hidden layers (with 128, 64, and 32 units), and finally, a dense layer to aggregate its dimensionality to 1. Moreover, when dropout is 0.5, it generates the most network structures randomly, which effectively enhances the generalization ability of the model and prevents overfitting (Song et al., 2021; Ding, 2022).

The activation function *Sigmoid* can easily cause problems such as gradient explosion and gradient disappearance, so *Relu* is used for the activation function. The optimization training is performed using the Adam optimizer, which can adaptively adjust the learning rate (Yi et al., 2020). The time step is set to 1, the feature vector is also 1, and the pitch size is 40. After 300 iterations, the loss rate tends to a stable value close to 0, and convergence is reached. The time series is reconstructed to obtain the predicted values of the EMD-LSTM model. Finally, the predicted values of the SARIMA-EMD-LSTM hybrid model are obtained based on the weighting factors.

As shown in Table 2, we can make the following conclusions:

1. The SARIMA model can only capture linear information but not nonlinear information. The forecast result shows that the RMSE is 8.933681, and the MAPE is 14.4404%, which is the worst forecast accuracy of all models, indicating that the long-term forecasting effect of the SARIMA model is not significant.
2. The forecast accuracy of the EMD-LSTM model improved by 2.862% over that of the LSTM model, and this conclusion is consistent with ***H1***.
3. The EMD-LSTM model has the highest forecast accuracy of all models: the MAPE is 7.4304%, which is 2.6398% higher than that of the SARIMA-EMD-LSTM model. This result indicates that the hybrid model does not exploit the linear predictive advantage of the SARIMA model with the nonlinear predictive advantage of the LSTM model. Therefore, this conclusion is inconsistent with ***H2.***
4. The forecast accuracy of the SARIMA-EMD-LSTM hybrid model is lower than that of the EMD-LSTM model, and the forecast accuracy of the SARIMA-LSTM model is lower than that of the LSTM model, which indicates that the hybrid model constructed by the method of weight assignment cannot take advantage of the SARIMA linear and LSTM nonlinear forecasting, so the weight assignment is not the best-combined method for constructing the hybrid model. This conclusion is inconsistent with ***H3***. The hybrid model is generated to effectively compensate for the limitations of the single model, but it does not mean that the single model forecasting is certainly worse than the hybrid model. The reason is that the forecast accuracy depends on the combined method of the hybrid model.
5. The SARIMA model accurately captures the fluctuation pattern of the time series from November 2020 to October 2021. Furthermore, the predicted value for January 2021 is the same as the true value. However, the abnormal decline in trade turnover due to COVID-19 is not reflected in the SARIMA model, which is a shortcoming of the SARIMA model fitting and forecasting.
6. The EMD algorithm effectively resolves the phenomenon of hysteresis generated by the LSTM model. This conclusion is consistent with Hypothesis 1, and it can accurately capture the fluctuation pattern of the Russian trade turnover time series. The ARIMA-LSTM and SARIMA-EMD-LSTM models are significantly influenced by the SARIMA model, as seen in Figure 6(c), with almost the same predicted fluctuation pattern of the time series.

## 5. Conclusion and Limitations

The Sino-Russia foreign trade turnover time series is complex and has significant seasonal fluctuations, which are transformed into a smooth series by differencing methods. In this paper, the two-stage grid search is used to find the optimal hyper-parameters, and according to the change in AIC values, it is necessary to add seasonal parameters. The EMD algorithm decomposes the time series into a finite number of IMFs and uses them as input information for the LSTM model. Then, the time series were reconstructed to obtain the predicted values of the EMD-LSTM model. The predicted values of the

**Table 2:** Predicted Performance for Each Model

| Model | WF | | MAE | MSE | RMSE | MAPE(%) |
|---|---|---|---|---|---|---|
| | a | b | | | | |
| SARIMA | – | – | 7.589268 | 79.810659 | 8.933681 | 14.4404 |
| LSTM | – | – | 5.415297 | 46.756882 | 6.8379 | 10.2924 |
| EMD-LSTM | – | – | 4.006359 | 25.941904 | 5.09332 | 7.4304 |
| SARIMA-LSTM | 0.58358 | 0.41642 | 5.631832 | 46.778791 | 6.839502 | 10.7466 |
| SARIMA-EMD-LSTM | 0.65449 | 0.34551 | 5.33119 | 38.325154 | 6.190731 | 10.0702 |

Source: Composed by authors.

SARIMA-EMD-LSTM hybrid model are obtained by weighting factors. The EMD-LSTM model has the highest forecast accuracy of all models and can be applied to the forecasting of the Sino-Russia foreign trade turnover time series post-epidemic. However, the following limitations remain in this paper:

1. The EMD decomposition algorithm effectively solves the hysteresis phenomenon while improving the forecast accuracy of the time series, as a result in the paper has confirmed. Although during the research, the problem of information leakage did not arise, it is an undisputed fact that when using the EMD algorithm, the problem of information leakage is prone to arise, and no effective solution is proposed in this paper.
2. In this paper, the optimal hyper-parameters are selected after several validations in the construction of the LSTM model, with a strong subjectivity. The Bayesian optimization algorithm is an efficient global optimization algorithm (Torun et al., 2018). It is widely used for hyperparameter optimization in machine learning and deep learning. Using the Bayesian optimization algorithm to determine the optimal hyper-parameters of the LSTM model can improve the forecast accuracy, which is a further research direction.
3. The paper validates that the weight assignment is not the best-combined method for constructing the hybrid model. However, other combined methods, such as the linear-nonlinear hybrid model and correction hybrid model, have not been validated.
4. Foreign trade turnover is affected by various factors, such as politics, geography, war, and epidemics. Only trade turnover is used as the feature variable of the LSTM model in this paper, and other factor are not considered. Russian GDP, international oil prices, and related trade products such as timber, metals, chemicals, machinery, and electricity are the main factors that impact Sino-Russia foreign trade turnover (Zhou & Kang 2021). Therefore, in later research, more macro-political and macroeconomic factors will be taken into account to build a complete system of trade turnover indicators and to improve the forecast accuracy of the model.

## References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/tac.1974.1100705

Alzheev, A., & Kochkarov, R. (2020). Comparative Analysis of ARIMA and lSTM predictive models: Evidence from Russian stocks. *Finance: Theory and Practice*, *24*(1), 14–23. https://doi.org/10.26794/2587-5671-2020-24-1-14-23

Dave, E., Leonardo, A., Jeanice, M., & Hanafiah, N. (2021). Forecasting Indonesia exports using a hybrid model ARIMA-LSTM. *Procedia Computer Science*, *179*, 480–487. https://doi.org/10.1016/j.procs.2021.01.031

Ding, X. W. (2022). A time series-based statistical approach for trade turnover forecasting and assessing: Evidence from China and Russia. *The Journal of Asian Finance, Economics, and Business, 9*(4), 83–92. https://doi.org/10.13106/jafeb.2022.vol9.no4.0083

Farooqi, A. (2014). ARIMA model building and forecasting on imports and exports of Pakistan. *Pakistan Journal of Statistics and Operation Research*, *10*(2), 157. https://doi.org/10.18187/pjsor.v10i2.732

Gyamerah, S., & Korda, D. (2021). Prediction of stock market returns using the LSTM model and traditional statistical model. *International Journal of Computer Applications*, *183*(37), 57–61. https://doi.org/10.5120/ijca2021921773

Gers, F., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural Computation*, *12*(10), 2451–2471. https://doi.org/10.1162/089976600300015015

Haider, S., Naqvi, S., Akram, T., Umar, G., Shahzad, A., & Sial, M. (2019). LSTM neural network-based forecasting model for wheat production in Pakistan. *Agronomy*, *9*(2), 72. https://doi.org/10.3390/agronomy9020072

Huang, N., Shen, Z., Long, S., Wu, M., Shih, H., & Zheng, Q. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, *454*(1971), 903–995. https://doi.org/10.1098/rspa.1998.0193

John, E. E., & Patrick, U. U. 2016. Short-Term Forecasting of Nigeria inflation rates using seasonal ARIMA model. *Science Journal of Applied Mathematics and Statistics, 4*(3), 101–107. https://doi.org/10.11648/J.SJAMS.20160403.13

Krizhevsky, A., Sutskever, I., & Geoffrey, E. H. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*, 84–90. https://doi.org/10.1145/3065386

Kim, H. Y., & Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems Applications, 103*, 25–37. https://doi.org/10.1016/j.eswa.2018.03.002

Li, T., Wang, B., Zhou, M., & Watada, J. (2018). Short-term load forecasting using optimized LSTM networks based on EMD. *Communications, Circuits, and Systems, 69,* 84–88. https:\\doi.org\10.1109/ICCCAS.2018.8769255

Li, Y., Zhu, Z., Kong, D., Han, H., & Zhao, Y. (2019). EA-LSTM: Evolutionary attention-based LSTM for time series prediction. *Knowledge-Based Systems*, *181*, 104785. https://doi.org/10.1016/j.knosys.2019.05.028

Mladenovic, J., Lepojevi, V., & Jankovi-Milic, V. (2016). Modeling and prognosis of the export of the Republic of Serbia by using seasonal Holt-Winters and ARIMA method. *Economic Themes*, *54*, 233–260. https://doi.org/10.1515/ethemes-2016-0012

Nelson, D. M., Pereira, A. M., & Oliveira, R. A. (2017). Stock market's price movement prediction with LSTM neural networks. *Neural Networks, 796,* 1419–1426. https://doi.org/10.1109/IJCNN.2017.7966019

Ni, X., Huang, H., & Du, W. P. (2017). Relevance analysis and short-term prediction of PM2.5 concentrations in Beijing based on multi-source data. *Atmospheric Environment*, *150*, 146–161. https://doi.org/ 10.1016/j.atmosenv.2016.11.054

Qu, Q., Li, Z., Tang, J., Wu, S., & Wang, R. (2019). A trend forecast of import and export trade total volume based on LSTM. *IOP Conference Series: Materials Science and Engineering, 646*, 012002. https://doi.org/10.1088/1757-899X/646/1/012002

Qin, X., Sun, M. Z., Dong, X., & Zhang, Y. (2018). Forecasting of China consumer price index based on EEMD and SVR method. *Data Science and Business Analytics, 201,* 329–333. https://doi.org/10.1109/ICDSBA.2018.00069

RU-STAT. (2021). *Export and import of Russia by goods and countries*. https://ru-stat.com/date-M201301-202101/RU/trade/CN.

Shu, W., & Gao, Q. (2020). Forecasting stock price based on frequency components by EMD and neural networks. *IEEE Access*, *8,* 206388–206395. https://doi.org\10.1109/ACCESS.2020.3037681

Song, C. H., Li, Z. X., & Yu, H. X. (2021). An oil well fault identification method base on improved GoogleNet. *Journal of Jiangsu University of Science and Technology, 35*(2), 52–58. https://doi.org/10.11917/j.issn.1673-4807.2021.02.008

Song, W., Gao, C., Zhao, Y., & Zhao, Y. (2020). A time series data filling method based on LSTM—Taking the stem moisture as an example. *Sensors, 20*, 5045. https://doi.org/10.3390/s20185045

Solanki, S., Inumula, K. M., & Chitnis, A. (2020). Sectoral contribution to economic development in India: A time-series co-integration analysis. *The Journal of Asian Finance, Economics, and Business*, *7*(9), 191–200. https://doi.org/10.13106/jafeb.2020.vol7.no9.191

Suh, S. S., Park, J. W., Song, G., & Cho, S. G. (2014). A study of air freight forecasting using the ARIMA model. *The Journal of Asian Finance, Economics, and Business*, *12*(2), 59–71. https://doi.org/10.13106/jds.2014.vol12.no2.59

Torun, H. M., Swaminathan, M., Kavungal Davis, A., & Bellaredj, M. L. (2018). A global Bayesian optimization algorithm and its application to integrated system design. *IEEE Transactions on Very Large-Scale Integration (VLSI) Systems, 26,* 792–802. https://doi.org/10.1109/TVLSI.2017.2784783

Urrutia, J. D., Abdul, A. M., & Atienza, J. B. (2019). Forecasting Philippines imports and exports using Bayesian artificial neural network and autoregressive integrated moving average. *Mathematics Applications, 2192*, 090015. https://doi.org/10.1063/1.5139185

Yu, M., Liu, T., Guan, Z., Sun, Y., Guo, J., Chen, L., & He, Y. (2022). SALSTM: An improved LSTM algorithm for predicting the competitiveness of export products. *International Journal of Intelligent Systems, 22*, 1–16. https://doi.org/10.1002/int.22839

Yi, D., Ahn, J., & Ji, S. (2020). An effective optimization method for machine learning based on ADAM. *Applied Sciences*, *10*(3), 1073. https://doi.org/10.3390/app10031073

Zhang, Z., Liu, S., & Tian, L. (2013). The study of Sino-Russian trade forecasting is based on the improved grey prediction model. *Industrial Engineering and Engineering Management, 5*, 66. https://doi.org/10.1007/978-3-642-38391-5_66

Zhang, L., & Sun, D. S. (2019). Forecast of total import and export trade based on a combination model. *Jiangsu Commercial Forum, 2,* 57–59. https://doi.org/10.13395/j.cnki.issn.1009-0061.2019.02.011

Zhou, S., & Kang, C. (2021). Research on the influencing factors of Sino-Russia foreign trade based on R language regression analysis. *Mathematical Problems in Engineering, 3,* 1–11. https://doi.org/10.1155/2021/5638831

Zainuri, Z., Viphindrartin, S., & Wilantari, R. N. (2021). The impacts of the COVID-19 pandemic on the movement of the composite stock price index in Indonesia. *The Journal of Asian Finance, Economics, and Business*, *8*(3), 1113–1119. https://doi.org/10.13106/jafeb.2021.vol8.no3.1113