Price factors analysis and volatility prediction of Belt and Road Theme Index based on EMD application

Hongmin Zhang *, Jinming He

School of Finance, Harbin University of Commerce, Harbin, 150000, China

Abstract: This year marks the 10th anniversary of the Belt and Road Initiative, and relevant theme indexes have attracted market attention. As a characteristic policy of China’s foreign trade, it is of great significance to study the development situation of its related stocks for the progress and future development of the Belt and Road. In this paper, EMD method and ARMA model are adopted. Taking the SSE One Belt & One Road Index as the representative of the Belt and Road Theme Index, the paper analyzes the price influencing factors of the Belt and Road theme stocks and predicts the future trend of their volatility. It believes that the price of the Belt and Road theme index is mainly affected by three parts: market supply relationship, short-term common influencing factors and long-term major event impact. Making EMD-ARMA prediction on its volatility and finds it to be more accurate than using the ARMA model directly. Load forecasting is very important for power dispatching. Accurate load forecasting is of great significance for saving energy, reducing generating cost and improving social and economic benefits. In order to accurately predict the power load, based on BP neural network theory, combined with the advantages of Clementine in dealing with big data and preventing overfitting, a neural network prediction model for large data is constructed.

Keywords: The Belt and Road, EMD, EMD-ARMA, Shanghai Belt and Road Theme Index

1. Introduction

1.1 Preface
The Belt and Road Initiative was formally proposed in September and October 2013. By December 6, 2022, 150 countries had signed intergovernmental cooperation agreements on the Belt and Road Initiative, and more than 32 international organizations had signed more than 200 cooperation documents. The "Belt and Road" Initiative has benefited a series of listed companies in infrastructure, port, communication and power, tourism, etc. As of February 11, 2023, the "Belt and Road" concept stocks included 255 A and B shares. This paper takes SSE One Belt & One Road Index (abbr. SSE OBOR) index as the representative of the Belt and Road index. The fluctuation of financial time series is complicated, and the factors involved are intricate, so it is very difficult to analyze the impact of specific reasons on the index price. Therefore, the Empirical Mode Decomposition method is considered in this paper, which can separate the sequences of different feature scales in the signal, namely the Intrinsic Mode Function (IMF), so we can analyze the internal causes of financial price fluctuation. EMD can be applied not only to linear stationary data, but also to non-stationary nonlinear sequences. It adds an effective method to the study of complex financial sequences and solves the problem of extracting volatility periods on multiple time scales. The application of this method has been widely promoted in recent years.

Through model studies (e.g. Saiyan Lin (2022) [1], Huayong Niu et al. (2022) [2]), it is found that using ARMA model to predict the series obtained by EMD processing is more accurate and reliable than using ARMA model alone. Therefore, this paper will conduct EMD processing on the data, analyze the causes of the decomposed sequences with different characteristics, and then predict the fluctuation of the original sequence data after EMD decomposition.

1.2 Theoretical Mode
Empirical mode decomposition (EMD). Empirical mode decomposition (EMD) is a multi-time series signal processing algorithm created by HUANG N and others in 1998 [3], which was initially applied in the field of engineering. EMD can decompose the signal into components of different frequencies, the essence of which is the intensity of the oscillation, that is, the extreme value of the signal is numerous and dense. EMD can decompose the signal into several IMFs and monotone residuals by using the information of extreme value points.

\[ u(t) = \text{imf}_1(t) + \text{imf}_2(t) + \text{imf}_3(t) + \ldots + \text{imf}_N(t) + r_N(t) \]
The methods for finding IMF function are as follows:

For the given function $u(t)$, find the maximum and minimum points.

The maximum envelope $e_{\text{max}}(t)$ and the minimum envelope $e_{\text{min}}(t)$ are given by using spline curves for extreme points respectively.

By averaging the two envelope lines, the average $m(t) = \frac{e_{\text{max}}(t) + e_{\text{min}}(t)}{2}$ is obtained.

Subtract the average value to get a new function $h_{1}(t) = u(t) - m(t)$.

Replace the original $u(t)$ with a new function $h_{1}(t)$, and repeat steps (1) - (4) to obtain $h_{k}(t)$ ($k= 2, 3, \ldots$).

Imf$_{1}(t)$ is defined until the IMF function condition that the moving average is 0 is satisfied.

**Autoregressive moving average (ARMA).** ARMA model by the autoregressive model (AR) and moving average (MA) model, is a kind of commonly used stationary time series fitting model, the general form of the model are as follows:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \ldots + \alpha_p Y_{t-p} + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} + \mu_t$$

Where $(\alpha_0 + \alpha_1 Y_{t-1} + \ldots + \alpha_p Y_{t-p})$ is an autoregressive process, $(\theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q})$ is a moving average process.

**EMD-ARMA.** Model prediction steps:

1. Using EMD method, decompose the original financial time series into IMF components with different characteristics.
2. Constructing ARMA models for each IMF component and make forecasts, then obtain the corresponding component of the optimal lag order number prediction results.
3. To synthesize each component of the prediction results to build a new prediction results.

### 2. Empirical analysis

#### 2.1 Data

The research data of this paper is from MarkPoint, which selects the historical closing price (cp) of the SSE OBOR (000160) from January 4, 2013 to February 17, 2023 and two sets of data of 2, 459 trading days in total. Where, the horizontal axis is the time, 1 is taken as the basis point of January 4, 2013, and the subsequent dates are accumulated, and the ordinate is the closing price. The closing price data is as follows green line in Figure 2.

#### 2.2 Factor Analysis

Through the EMD decomposition we can get seven IMFs and a trend line (see Figure 1). It can be concluded that the first three sequences, IMF1-IMF3, belong to high frequency sequences, and the last four, IMF4-IMF7, belong to low frequency sequences. By referring to the practice of ZHANG X et al. (2009) [4], high frequency data and low frequency data were respectively synthesized, The results are shown in Figure 2.

**Trend item(residual)** is the dotted line, and the label is “residual”, indicating that when there is no shock event, the price fluctuation of the index will increase upward with the increase of supply and demand relationship, which is the long-term stable trend of the index. An increase in the number of shares and the amount of money flowing into the capital markets can explain this.

**High frequency sequence synthesis** is the blue solid line, and the label is ‘imf1- 3’, namely the synthesis of IMF1-IMF3, which represents the impact of short-term factors such as daily policy release, industry changes and large amount of financing on the Belt and Road Economic Belt. One of the more prominent reasons is the depreciation of RMB and related economic and political policy changes of the Belt and Road. Throughout the historical data of RMB exchange rate, it is found that the fluctuation trend is indeed more and more frequent. However, when we disassemble the fluctuation cycle of the short-term impact curve, we find that the fluctuation cycle of the short-term impact curve is shortening and the fluctuation range is increasing, both of which have the same trend: The People's Bank of China announced at the end of 2015 that...
The RMB would join the international Special Drawing Rights (SDR) currency basket on January 1, 2016; The AIIB was officially established in June 2015.

**Low frequency sequence synthesis** is the red solid line, and the label is “imf4-7”, namely, the synthesis of IMF4-IMF7, which represents the medium and long-term impact of mega events on the Belt and Road development. Its trend is generally consistent with the original series trend and the fluctuation trend of Shanghai composite index. It shows that the major event in the Belt and Road theme index is the systemic risk of the whole industry and the whole market. Among them, the central bank has lowered the deposit reserve ratio and benchmark interest rate for many times, resulting in a large amount of funds flowing to the stock market. Coupled with irrational leverage, the stock market bubble eventually resulted in a stock market crash. The stock market crash in 2014-2015 also had an impact on the Belt and Road theme index, with the price of SSE OBOR falling from 3,268.47 to 1,312.29, and the leading export index of Belt and Road foreign trade falling from 43.3 to 31.2 during the period. At the beginning of 2018, the United States began to impose tariffs on China, which caused a great impact on the Belt and Road development with foreign trade as the key. The outbreak of COVID-19 in 2020 brought a “cold winter” to tourism consumption, transportation, services and other industries with high human mobility, and brought the global economy to a standstill, and capital flows into the stock market slowed. Therefore, there are three factors affecting the price: relatively stable supply-demand relationship, short-term impact, and medium and long-term impact caused by major events.

### 2.3 Volatility prediction

Conduct the following operations on the rise and fall data of the SSE OBOR:

**Operation M1**: ARMA model prediction is made directly with the original data of rise and fall.

**Operation M2**: The original data of rise and fall are decomposed by EMD, and the results are as Figure 3. Each IMF sequence and trend items are predicted by ARMA, and the forecast results of each sequence are synthesized to obtain the final forecast results.

The ARMA fitting model are as follows: Original Sequence-ARMA(3, 2); IMF1- ARMA(3, 2); IMF2-ARMA(4, 4); IMF3-(3, 1, 1); IMF4-ARIMA(3, 2, 1); IMF5-ARIMA(3, 2, 1); IMF6-ARMA(2, 2); IMF7-ARIMA(1, 2, 1); IMF8-AR(8); residual-Logarithmic first-order difference AR(2).

All the prediction results of operation M1 and operation M2 were evaluated (see Figure 4). The evaluation indexes of prediction data used in this paper were mean square error (MSE), root mean square error (RMSE) and mean absolute error (MAE). 2459 data were evaluated. The results show that the errors predicted by the EMD-ARMA model are more accurate than those predicted by the original sequence.

### 3. Conclusion

In this paper, the original closing price sequence of the Belt and Road theme index was decomposed by EMD, and the decomposed sequence was reassembled according to different frequencies to obtain the trend item, high frequency and low frequency sequence. By comparing the three with the original series, it is found that the influencing factors of the Belt and Road are mainly the supply and demand relationship of stocks, the leverage bubble of the capital market and the trade conflict between China and the United States, and the common short-term market price influencing factors such as the exchange rate and the related policies of the Belt and Road. Then, EMD decomposition is carried out on the volatility of the Belt and Road theme index, i.e. the rise and fall data. Then, 8 IMF sequences and a trend item obtained from decomposition are combined with the results obtained by ARMA prediction respectively to obtain a set of forecast data. In addition, ARMA prediction is carried out on the unprocessed original sequence of rise and fall to obtain a set of forecast data. Comparing the two groups of prediction data with the original sequence, it is found that the results obtained by EMD-ARMA are more accurate.

### References

