Comparative Study of Short-Term Forecasting Methods for Soybean Oil Futures Based on LSTM, SVR, ES and Wavelet Transformation

Ganqiong Li¹², Wei Chen¹²*, Denghua Li¹², Dongjie Wang¹², Shiwei Xu¹²

¹. Agricultural Information Institute of Chinese Academy of Agricultural Sciences, Beijing, 100081
². Key Laboratory of Agricultural Information Service Technology, Ministry of Agriculture and Rural Affairs, Beijing, 100081

Corresponding author e-mail: chenwei@caas.cn

Abstract. Short-term forecasting of futures market is valuable and is also a technical challenge. In this paper, a hybrid approach for soybean oil futures price forecasting is proposed based on time-series analysis methods. The method combines wavelet transformation and exponential smoothing so that the characteristics of the time series can be captured at different time scales, and forecasting based on exponential smoothing is applied at each time scale. A comparative case study is then conducted that compares the proposed method with other three methods which are an RNN network with Long Short-Term Memory units, a Support-Vector Regression model, and an Exponential Smoothing model without wavelet decomposition to the time series. It could be concluded that the forecasting error performance of ES and Wavelet-ES was better than LSTM and SVR, and the Wavelet-ES achieved the best results for the direction forecasting. The case study provides valuable reference for application of short-term futures price forecasting.

1. Introduction

Soybean is an important kind of plant that has been widely grown in several countries. Soybean oil can be extracted from soybeans and is one of the mostly used cooking oil all over the world. In recent years, the prices of soybean oil fluctuate dramatically, which causes significant influence on the consumers’ daily lives. Soybean oil futures contracts are a kind of financial instruments that allow trading of soybean oil at a predetermined price in the future. The nature of the contracts makes them possible to be used as a potential indication of the ongoing trends of the product. Therefore, the research on the forecasting of soybean oil futures prices can be significantly beneficial to the market participants as well as policy makers.

The forecasting of commodity futures prices has long been studied, such as crude oil and gold. However, soybean oil futures prices forecasting has not been extensively covered by researchers. The forecasting can be achieved by investigating the relationship between the futures prices and other related factors. Wang and Liu analyzed the soybean futures prices according to its relation to the market spot prices using the VAR model [¹]. Wang also investigated Sino-US soybean oil futures based on the relation of the same futures contract from different countries’ futures exchanges [²]. On the other hand, soybean oil futures prices can also be analyzed using Time-Series (TS) analysis methods [³].

Conventional TS forecasting methods have been widely used in the forecasting of TS data, such as Autoregressive Integrated Moving Average (ARIMA) [¹][³], and Seasonal ARIMA (SARIMA) [⁶][⁷].
However, due to the linear nature of those methods, the TS data are assumed to be stationary. The limited capability of handling non-linear and non-stationary data makes it not suitable for financial instrument prices.

In order to deal with non-linear and non-stationary TS data, Artificial Intelligence (AI) and Machine Learning (ML) methods are widely adopted in financial instrument price forecasting, such as deep learning with Recurrent Neural Network (RNN) \cite{8} \cite{9} \cite{10}, and Support-Vector Regression (SVR) \cite{11} \cite{12}.

Exponential Smoothing (ES) is a method for smoothing TS data using exponential window function, in which exponential functions are used to assign exponentially decreasing weights over time. A number of variants can be applied that are capable of handling TS data that can be either stationary or non-stationary \cite{13} \cite{14}.

Financial instrument prices, especially soybean oil prices, can be influenced by multiple factors such as soybean production calendar, seasonal consumption, weather, trading policies, etc. Different factors change at different time scales. The Wavelet Transformation (WT) technique can be used to decompose the TS into sub-TSs of different frequencies so that forecasting techniques can be applied to each individual frequency \cite{15} \cite{16} \cite{17}. This can be beneficial to capture the characteristics of the TS in terms of its influence factors.

In this paper, a hybrid approach Wavelet-ES for soybean oil futures price forecasting is proposed based on TS analysis methods. The method combines WT and ES so that the characteristics of the TS can be captured at different time scales, and forecasting based on ES is applied at each time scale. A comparative study is then conducted that compares the proposed method with comparable methods, namely an RNN network with Long Short-Term Memory (LSTM) units, a Support-Vector Regression (SVR) model, and an ES model without WT decomposition to the TS.

The rest of the paper is organized as follow. Section 2 introduces the three comparing methods as well as the proposed method. Experiment results are presented in Section 3. Section 4 presents the concluding remarks.

2. Methods

In this section, the three comparing methods and the proposed method are introduced.

2.1. The LSTM Model

An RNN network with LSTM units is adopted for the forecasting of soybean oil futures prices. RNN is a kind of Artificial Neural Network (ANN) in which units contain recurrent connections. This kind of connections can be used to pass information from a time step to the next.

LSTM is an improved RNN architecture which can be used to solve the vanishing gradient and exploding gradient problems.

A LSTM network can be facilitated for the classification, processing, and prediction of TS data, since it is capable of capturing the lags of unknown duration between important events.

A typical LSTM unit consists a cell, an input gate, an output gate and a forget gate. The cell remembers values over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell \cite{18}, as shown in Fig. 1.

![Fig. 1. Structure of a typical LSTM unit](image-url)
The outputs of the cell and gates are calculated as follows:

\[ f_t = \sigma_g \left( W_f x_t + U_f c_{t-1} + b_f \right) \]  
(1)

\[ i_t = \sigma_g \left( W_i x_t + U_i c_{t-1} + b_i \right) \]  
(2)

\[ o_t = \sigma_g \left( W_o x_t + U_o c_{t-1} + b_o \right) \]  
(3)

\[ c_t = f_t \odot c_{t-1} + i_t \odot \sigma_c(W_c x_t + b_c) \]  
(4)

\[ h_t = o_t \odot \sigma_h(c_t) \]  
(5)

where matrices \( W \) and \( U \) contain the weights of the input and recurrent connections, respectively, the operator \( \odot \) denotes the Hadamard product (element-wise product), and the subscript \( t \) indexes the time step of the TS.

2.2. The SVR Model

The SVR model is based on the Support Vector Machine (SVM) model, which is a supervised learning method in the ML field. It has been widely used for data classification. A data point is considered as a vector. The algorithm aims to find a hyperplane that can separate data points of different classes with a maximized margin.

If the data points are not linear separable in its original space, they are mapped to a higher-dimensional space using a kernel function. This allows effective handling of complex problems with non-linear classification.

When solving the regression problem, instead of finding a maximum-margin hyperplane to separate the data points in the SVM model, the SVR model aims to find a hyperplane that minimize the distances to the data points.

Training the SVR model means solving

\[ \min \frac{1}{2} \| w \|^2 \]  
subject to \[ | y_i - \langle w, x_i \rangle - b | \leq \varepsilon \]  
(6)

where \( x_i \) is a training sample with target value \( y_i \), the inner product plus intercept \( \langle w, x_i \rangle + b \) is the prediction for that sample, and \( \varepsilon \) is a threshold.

The SVR model can also be used for TS forecasting if recent \( k \) data points \( \{ x_{t-k+1}, x_{t-k+2}, ..., x_t \} \) in the past are used as the input data to forecast the next coming data \( x_{t+1} \).

2.3. The ES Model

The ES model is a smoothing technique for TS data that smooths the current data point using past data points of a time window. Unlike the conventional Moving Average (MA) model that assigns equal weights to past data points, the ES model assigns more weights to more recent data points using an exponential function. This can be beneficial for handling TS data that are non-stationary and non-linear [17].

The ES model can be applied in a wide variety in terms of its trend and seasonal components. Both components can be classified as none, additive, and multiplicative according to the nature of the data. In this paper, an ES model with additive trend and additive multiplicative is adopted, namely the Holt-Winters’ additive method [19][20]. The forecasting is achieved recursively as follows:

\[ \hat{y}_{t+h|t} = l_t + h b_t + s_{t+h-m(k+1)} \]  
(7)

\[ l_t = \alpha (y_t - s_{t-m}) + (1-\alpha)(l_{t-1} + b_{t-1}) \]  
(8)

\[ b_t = \beta^* (l_t - l_{t-1}) + (1-\beta^*)b_{t-1} \]  
(9)

\[ s_t = \gamma (y_t - l_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m} \]  
(10)

where \( l_t \) is the level component of the TS which shows a weighted average between the seasonally adjusted observation and the non-seasonal forecast for time \( t \), \( b_t \) is the trend component which denotes an estimate of the trend, \( s_t \) is the seasonal component which shows a weighted average between the
current seasonal index and the seasonal index of the same season last year (i.e., \( m \) time periods ago), and \( k \) is the integer part of \( (h - 1)/m \) which ensures that the estimates of the seasonal indices used for forecasting come from the final year of the sample.

2.4. The Proposed Wavelet-ES Model

A hybrid approach is proposed that combines WT and ES models. The WT model can be used to decompose signals into different frequencies and is capable of capturing both frequency and location information while the conventional Fourier Transform (FT) model can only capture frequency information. When used in TS analysis, it is especially useful to handle non-stationary TSs and is more flexible than Short-Time Fourier Transform (STFT) in terms of handling the time window.

The time-scale WT of a continuous TS \( f(t) \) is denoted by \( W(\alpha, b) \):

\[
W(\alpha, b) = \int_{-\infty}^{\infty} \psi^*(\frac{t-b}{\alpha}) f(t) dt
\]

where \( \psi^* \) is the mother wavelet conjugate. The wavelet coefficients \( W(\alpha, b) \) are obtained by continuously varying the scale parameter and the position parameter in order to select the different portions of the signal and analyze the different scale variations.

In a discrete-time signal processing scenario, a dyadic grid is employed, where the mother wavelet is scaled by power two (\( \alpha = 2^j \)) and translated by an integer (\( b = k2^j \)). A Discrete Wavelet Transform (DWT) is expressed by the following equation.

\[
\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j}t - k)
\]

The DWT coefficients are obtained from the following expression.

\[
W_{j,k} = W(2^j, k2^j) = 2^{-j/2} \int_{-\infty}^{\infty} f(t) \psi(2^{-j}t - k) dt
\]

In order to reconstruct the original TS, an Inverse Discrete Wavelet Transform (IDWT) is calculated from the wavelet coefficients \( W_{j,k} \) using the following formula.

\[
f(t) = \sum_j \sum_k W_{j,k} \psi_{j,k}(t)
\]

Soybean oil futures price can be affected by a composition of different influencing factors, each of which can affect the price at a different time scale. Past data points in the TS are first decomposed using DWT into sub-TSs of different frequencies to capture different influencing factors at different time scales. A single-level DWT decomposes a TS into an approximation sub-TS and a detail sub-TS. Multi-level DWT can be achieved by further applying the decomposition to the approximation sub-TS at each subsequent level. A \( n \)-level DWT generates \( n+1 \) sub-TSs \( A_{n}, D_{n}, D_{n-1}, \ldots, D_{2}, D_{1} \), where \( A_{n} \) is the approximation sub-TS generated in the \( n \)-th level, and \( D_{k} \) (\( k \in \{1, 2, \ldots, n\} \)) is the detail sub-TS generated in the \( k \)-th level.

After the decomposition of the TS, forecasting techniques can be applied to each individual frequency sub-TS. In the proposed approach, a simple ES model without trend or seasonal components is applied to each sub-TS.

3. Experiment

Experiments are conducted for a comparative analysis that compares the proposed method with comparable methods.

3.1. Characteristics of the TS

The TS data are extracted for soybean oil futures prices of Dalian Commodity Exchange (DCE), China from January 9, 2006 to December 28, 2015. Totally 2,425 data points are collected. The price movement can be seen in Fig. 2.
The characteristics of the TS is measured. The overall mean of the TS is 7719.76, and the overall standard deviation is 1789.66. The rolling mean and standard deviation of the TS are shown in Fig. 3. From Fig. 3(a) we can see that the rolling standard deviation of the original TS is relevantly stable while the rolling mean is rather dynamic. The result of Dickey-Fuller (DF) test suggests that the original TS is not stationary. However, the original TS can be made stationary by applying first-order differencing. The results can be seen in Fig. 3(b).
Although the TS can be made stationary easily by applying first-order differencing, we can see from Fig. 3(b) that the statistical properties of the transformed TS is not always stable during some period of time, especially when the market volatility is rather high. Therefore, in order to achieve automatic online forecasting of soybean oil future prices, sophisticated methods that are capable of handling non-stationary data are better to be adopted for the forecasting task.

3.2. Evaluation Metric
The Root-Mean-Square Error (RMSE) metric is selected for the evaluation of the forecasting results. RMSE is a widely used metric that is used to measure the difference between predicted values computed by a model and the observed values. The RMSE of a predicted TS \( \{\hat{y}_1, \hat{y}_2, ..., \hat{y}_N\} \) and an observed TS \( \{y_1, y_2, ..., y_N\} \) is calculated as follows:

\[
RMSE = \sqrt{\frac{\sum_{t=1}^{N} (\hat{y}_t - y_t)^2}{N}}
\]

(15)

3.3. Experiment Results
A comparative study is conducted to the TS that compares the proposed method Wavelet-ES with comparable methods, namely LSTM, SVR, ES.

An out-of-sample walk-forward validation test is conducted to evaluate the four models. For each time step \( t \in \{1000, 1001, ..., N - 1\} \), its recent 1000 data points \( \{x_{t-999}, x_{t-998}, ..., x_t\} \) in the past are used as the training data to forecast the data point \( x_{t+1} \) that comes next. For SVR, ES, and Wavelet-ES algorithms, a forecasted value is acquired for every time step and compared with the observed value. However, due to lengthy training time, for the LSTM algorithm the test is run for every 10 time steps. The forecasted values are compared with the observed ones for the evaluation of the models.

For the proposed Wavelet-ES model, a 2-level DWT is applied to the original TS. The decomposition result for the earliest training data \( \{x_1, x_2, ..., x_{1000}\} \) is demonstrated in Fig. 4. From the figure we can see that the original TS is decomposed into three sub-TSSs with different frequencies.

Fig. 4. The wavelet decomposition result of Wavelet-ES
The evaluating results of the models are shown in Table 1 and the forecasted price movements are shown in Fig. 5. From the table it can be observed that ES and Wavelet-ES achieve better results than LSTM and SVR. This demonstrate the strong capability of ES for handling non-stationary data in soybean oil price forecasting. Especially, the RMSE is even lower when the WT is adopted for the decomposition of the TS into different frequencies. On the other hand, the LSTM and SVR models are less effective than the other two. This might be an indication that ML algorithms are less capable for soybean oil futures forecasting than TS smoothing models due to the highly dynamic nature of the movements of the futures prices.

| Model       | RMSE  |
|-------------|-------|
| LSTM        | 522.33|
| SVR         | 469.23|
| ES          | 100.38|
| Wavelet-ES  | 89.76 |

The forecasting effectiveness of the models in different time periods is evaluated. The absolute differences between the forecasted and the actual prices are illustrated over time in Fig. 6. It can be seen from the figure that the forecasted period from 2010 to 2015 can be roughly divided into two parts. Period 1 is from late 2010 to the middle of 2012, in which the price movement is relevantly stable. Period 2 is from the middle of 2012 to the middle of 2015, in which the prices go down dramatically with a strong movement trend. It can be found that all four models achieve good results when the prices are relevantly stable while ES and the proposed Wavelet-ES are even better in Period 1. However, during Period 2 in which the price movement is more dynamic, the absolute differences for LSTM and SVR are increased dramatically which indicates the low performances of them in this period. On the contrast, ES and Wavelet-ES achieve even lower differences in this period, which indicates their effectiveness to follow trends.
The capability of forecasting the price movement directions of the models is shown in Table 2. It can be seen in the table that all four models achieve correctly forecasted movement directions around 50% for the next day. This is due to the random walk nature of the short-term price movement. Among the four models, the proposed Wavelet-ES achieves the best results for the direction forecasting. This demonstrates its effectiveness to capture short-term price movements.

Table 2. Correctly forecasted movement directions of the models

| Model     | Correctly forecasted movement directions |
|-----------|-----------------------------------------|
| LSTM      | 47.6%                                   |
| SVR       | 50.2%                                   |
| ES        | 49.1%                                   |
| Wavelet-ES| 51.7%                                   |

The efficiency of the four methods to train a model for the TS forecasting is also evaluated and the result is shown in Table 3. From the table we can see that training an SVR model takes the shortest time. Training the proposed Wavelet-ES model achieves the second place and is also relevantly fast. On the other hand, the training time for LSTM is the longest among the four models.

Table 3. Training time of the models

| Model     | Time (ms) |
|-----------|-----------|
| LSTM      | 5230.77   |
| SVR       | 6.00      |
| ES        | 216.87    |
| Wavelet-ES| 20.99     |

4. Conclusion
This paper presents a hybrid approach Wavelet-ES for soybean oil futures price forecasting based on TS analysis methods. The method combines WT and ES so that the characteristics of the TS can be captured at different time scales, and forecasting based on ES is applied at each time scale. A comparative study is then conducted that compares the proposed method with comparable methods, namely an RNN network with Long Short-Term Memory (LSTM) units, a Support-Vector Regression (SVR) model, and an ES model without WT decomposition to the TS.

For our subsequent research, we plan to extend the proposed approach to futures price forecasting of other major agricultural products.
Acknowledgments
Project funded by Central Public-interest Scientific Institution Basal Research Fund (JBYW-2020-AII); National Key Research and Development Project (2018YFF0213506).

References
[1] Jun, W., and Yaqing, L. Research on the dynamic relationship between china's soybean oil futures and spot price: empirical analysis based on daily data. Journal of China Agricultural University, 12(6), 6-13, 2007.
[2] Jun, W. Research on the International Linkages of Sino-US Soybean Oil Futures & Spot Markets and Dynamic Forecast. Journal of Northwest A&F University, 2008.
[3] Wu, H., Wu, H., Zhu, M., Chen, W., and Chen, W. A new method of large-scale short-term forecasting of agricultural commodity prices: illustrated by the case of agricultural markets in Beijing. Journal of Big Data, 4(1), 1-22, 2017.
[4] Jiang, S., Yang, C., Guo, J., and Ding, Z. ARIMA forecasting of China’s coal consumption, price and investment by 2030. Energy Sources Part B-economics Planning and Policy, 13(3), 190-195, 2018.
[5] Ajadi, N. A., Ajadi, J. O., Damisa, A. S., Asiribo, O. E., and Dawodu, G. A. Modeling Monthly Average Temperature of Dharhan City of Saudi-Arabia Using Arima Models. Journal of data science, 3(5), 2017.
[6] Asamoahboaheng, M., and Polytechnic, K. Using SARIMA to Forecast Monthly Mean Surface Air Temperature in the Ashanti Region of Ghana. International journal of statistics and applications, 4(6), 292-298, 2014.
[7] Luo, C. S., Zhou, L. Y., and Wei, Q. F. Application of SARIMA Model in Cucumber Price Forecast. international conference on robotics and automation, 2013.
[8] Kong, W., Dong, Z. Y., Jia, Y., Hill, D. J., Xu, Y., and Zhang, Y. Short-Term Residential Load Forecasting Based on LSTM Recurrent Neural Network. IEEE Transactions on Smart Grid, 10(1), 841-851, 2019.
[9] Huang, C., and Kuo, P. A Deep CNN-LSTM Model for Particulate Matter (PM2.5) Forecasting in Smart Cities. Sensors, 18(7), 2018.
[10] Jiang Q., Tang C., Chen C., Wang X., and Huang Q. Stock Price Forecast Based on LSTM Neural Network. In: Xu J., Cooke F., Gen M., Ahmed S. (eds) Proceedings of the Twelfth International Conference on Management Science and Engineering Management. ICMSEM 2018. Lecture Notes on Multidisciplinary Industrial Engineering. Springer, Cham, 2019.
[11] Hong, W., Dong, Y., Chen, L., and Wei, S. SVR with hybrid chaotic genetic algorithms for tourism demand forecasting. Applied Soft Computing, 11(2), 1881-1890, 2011.
[12] Bican, B., and Yaslan, Y. A hybrid method for time series prediction using EMD and SVR. International symposium on communications control and signal processing, 2014.
[13] De Oliveira, E. M., and Oliveira, F. L. Forecasting mid-long term electric energy consumption through bagging ARIMA and exponential smoothing methods. Energy, 776-788, 2018.
[14] Yang, D., Sharma, V., Ye, Z., Lim, L. H., Zhao, L., and Aryaputera, A. W. Forecasting of global horizontal irradiance by exponential smoothing, using decompositions. Energy, 111-119, 2015.
[15] Risse, M. Combining wavelet decomposition with machine learning to forecast gold returns. International Journal of Forecasting, 35(2), 601-615, 2019.
[16] S. Li, P. Wang and L. Goel, A Novel Wavelet-Based Ensemble Method for Short-Term Load Forecasting with Hybrid Neural Networks and Feature Selection, in IEEE Transactions on Power Systems, 31(3), 1788-1798, 2016.
[17] Sharghi, E., Nourani, V., Najafi, H. et al. Wavelet-Exponential Smoothing: a New Hybrid Method for Suspended Sediment Load Modeling. Environ. Process. 6, 191–218, 2019.
[18] Long short-term memory – Wikipedia, https://en.wikipedia.org/wiki/Long_short-term_memory, last accessed 2020/7/12.
[19] Charles C. Holt, Forecasting seasonals and trends by exponentially weighted moving averages, International Journal of Forecasting, 20(1), Pages 5-10, 2004.

[20] Peter R. Winters, Forecasting Sales by Exponentially Weighted Moving Averages, Management Science, 6(3), 1960.

[21] Rhif, M., Abbes, A., Farah, I., Martinez, B., Sang, Y. Wavelet Transform Application for/in Non-Stationary Time-Series Analysis: A Review. Applied Sciences, 9(7), 2019.