Artificial Intelligence Evolution
http://ojs.wiserpub.com/index.php/AIE/

UNIVERSAL WISER PUBLISHER

Research Article

Enhanced Bollinger Band Stock Quantitative Trading Strategy Based on Random Forest

Keyue Yan[1], Yimeng Wang[2], Ying Li[3*]

[1]Department of Computer and Information Science, University of Macau, Macau, China
[2]Department of Mathematics, University of Macau, Macau, China
[3]College of Global Talents, Beijing Institute of Technology-Zhuhai Campus, Zhuhai, China
Email: yingleely@qq.com

Received: 2 October 2022; **Revised:** 5 December 2022; **Accepted:** 24 December 2022

Abstract: Constructing applicable automated stock trading strategies has become one of the best ways that people can earn profits from their underlying assets' investments now. Automated stock trading, also called quantitative trading, contains sets of human-defined rules, which are written in codes to make decisions to go long or short on stocks on a computer. Investment banks, brokerages, private equity funds, and other financial institutions around the world are keen on investigating and developing quantitative trading strategies with sustainable profitability to yield higher returns than the normal market. This research aims to observe the trading performance and profits of financial banking stocks in the Hong Kong stock market by building a quantitative trading strategy named Enhanced Bollinger Band Strategy based on Random Forest and Bollinger Bands. In experiments, the Random Forest algorithm is applied to predict the Weighted Moving Average the next day. Meanwhile, Bollinger Bands are the trading signals used to make decisions on going long or short positions based on the historical moving average lines and standard deviation. Performances of the Enhanced Bollinger Band Strategy are evaluated by test sets of ten financial banking stocks. We also compare the performance of the Enhanced Bollinger Band Strategy and Traditional Bollinger Band Strategy and find that the Enhanced Bollinger Band Strategy can earn 10-30% profits on a variety of stocks although these stocks are losing 10-50% original amount of investment in Traditional Bollinger Band Strategy and basic investment. Therefore, a combination of Random Forest and Bollinger Bands in the quantitative trading strategy generates higher returns than simply investing in stocks.

***Keywords*:** stock price prediction, weighted moving average, random forest, stock quantitative trading, bollinger band strategy

1. Introduction

The development of computer technology has made it possible to send instructions for buying and selling orders in stock exchanges by using different Apps on phones or laptops. In developed financial markets, quantitative trading, a combination of statistical models and investment strategies, is replacing traditional order placement gradually. Quantitative trading is a process of using models to invest in practice and implement trading strategies. During the investment process, quantitative trading strictly executes the strategy with codes made by the investors initially and never changes their decisions of buying and selling decisions due to investors' sentiments [1]. When constructing

DOI: https://doi.org/10.37256/aie.4120231991

quantitative trading strategies, we are able to simulate the trading environment and evaluate the performance of strategies by using historical data. For example, the kind of strategy used in this research is the Bollinger Band Strategy [2]. In this strategy, stock prices always move inside Lower Track and Upper Track and rarely break these tracks. Investors can observe how stock prices fluctuate with the two tracks and trade to earn profits. When stock prices are lower than the Lower Track of Bollinger Band, investors can buy the stock which is undervalued. When stock prices are higher than the Upper Track of Bollinger Band, investors can sell the holding stock immediately which is overvalued. To observe and predict the trend of stock prices' trends in Bollinger Band, we use machine learning to predict the stock prices trend of next day in our research. However, stock prices in the financial markets exist too many noises which make machine learning algorithms difficult to predict future price or trend accurately. Also, we find that traditional statistical forecasting models usually produce large errors, while deep learning algorithms not only cause overfitting but also require extra hardware resources. To solve problems of noises in data and overfitting, we use Weighted Moving Average (WMA) lines of stock prices to reduce noises in data and to predict the future trend. What's more, Random Forest, one of the most famous traditional machine learning algorithms, is applied to predict stock price trends in our research. The combination of predicted WMA prices and Bollinger Bands are used as signals for buying and selling stocks in the Enhanced Bollinger Band Strategy. To observe the gains or losses, we set up a simulation environment for each stock and evaluate them. Results show that investing with the Enhanced Bollinger Band Strategy owns good returns. The historical orders in the simulating trades are generated according to the rules of the quantitative trading strategy. After completing the simulations, we obtained detailed trading reports providing the performance of the strategy over the specified period. In the current tests, the Enhanced Bollinger Band Strategy can achieve returns of 4-6% and even 10% on some stocks.

The remainder of this research can be summarized as follows. Section 2 reviews related research work on machine learning or deep learning algorithms and quantitative trading strategies. Section 3 describes the technical methods and models in detail. In Section 4, we present the performance of the models' metrics and the performance of the Enhanced Bollinger Band Strategy after the experiments. Finally, we conclude the research in Section 5.

2. Related works

The stock market plays an important role in the financial sector and therefore, stock price forecasting has been a hot topic in academics. Current research show that machine learning algorithms are widely used in the field of finance, where clustering and classification algorithms are applied to select profitable stocks to generate extra returns [3]. The machine learning algorithm is applied to forecast the stock price trend of the next day in this research. The process of prediction involves building models to fit the structure of historical data and predicting the future value of different time series datasets [4]. And then the predicted results can be applied in the field of investment, which provides useful references for researchers, investors, and financial institutions. However, predicting prices is a difficult and challenging task, the gambling of major capital groups and the noise in the trading market such as politics, wars, and epidemics are always influencing price fluctuations and changes [5]. It has been extremely difficult to rely on traditional financial theories to gain a precise understanding of the market. Therefore, current research papers are preferable to use statistical models or machine learning. Traditional statistical models, such as Autoregressive Integrated Moving Average (ARIMA) and Generalized AutoRegressive Conditional Heteroskedasticity (GARCH), are widely used [6-8]. Khairalla et al. used a combination of ARIMA and other forecasting algorithms for different time series [9]. With the development of hardware, methods such as machine learning, deep learning, and reinforcement learning can be used in the prediction [10-12]. By constructing different technical indicator variables, Zhang et al. achieved high accuracy in classifying stock prices' increasing and decreasing for the coming day [13]. Yan et al. predicted future prices accurately by performing signal decomposition on the original time series [14]. Among the deep learning algorithms used in the prediction, Huang et al. combine the decomposed stock prices with a neural network to make a good contribution to fitting future trends of prices [15]. Sood et al. transformed different time series data into image datasets and successfully predicted the future trend of the stock using a Convolutional Neural Network [16]. Other recent studies use LSTM and reinforcement learning for the prediction, and Ge et al. [17-18] studied the predictive performance in a reinforcement learning setting with different neural network models.

For the research relating to quantitative trading, technical indicators and stock market trends are used in quantitative trading regularly at the beginning [19-20]. As algorithm trading became popular, machine learning algorithms begin to be involved in trading. Attanasio et al. [21] performed simple buying and selling operations with machine learning predictions such as ARIMA, Support Vector Machine (SVM), and Multi-Layer Perceptron (MLP) to achieve large returns on high-volatility digital currency investments. In statistical arbitrage research [22-23], Zhang et al. use Long Short Term Memory (LSTM) to predict the difference between two highly correlated stock prices for arbitrage operations and obtain stable returns during the simulation. Yue et al. [24] used a more complex CNN to predict the spreads of the Chinese stock index and built an arbitrage strategy based on the difference between the increment of the actual spread and the increment of the predicted spread. In other quantitative trading strategies such as the Bollinger Band Strategy, Seshu et al. [25] explored the backtesting returns of LSTM-based investment strategies and ordinary Bollinger Band Strategies and then make a more detailed comparison across different datasets and time periods.

In this research, we propose a novel idea to execute the Enhanced Bollinger Band Strategy combining Random Forest and Bollinger Bands. We aim to predict price trends of the next day through Random Forest. Then, we combine the future price trends and the Bollinger Bands to make investment decisions such as going long or short, position adjustments, and so on. This allows us to observe the profitability of the Enhanced Bollinger Band Strategy during the simulation process.

3. Methodology

3.1 *Data preparation*

Hong Kong owns a well-established market policy and a long history of trading financial products. For the selection of data, we select a sample of ten financial banking stocks from Hong Kong Exchanges and Clearing Limited, including HSBC Holdings plc (00005), Standard Chartered (02888), Hang Seng Bank (00011), Dah Sing Banking Group Limited (02356), The Bank of East Asia Limited (00023), Bank of China (Hong Kong) Limited (02388), China Construction Bank Corporation (00939), Bank of Communications Co., Ltd (03328), Agricultural Bank of China Limited (01288) and China CITIC Bank (00998). These stocks are pivotal in the Hong Kong banking industry, and three of the banks even have abilities to issue Hong Kong dollars. All datasets are downloaded from Yahoo Finance [26]. The daily prices of Open, High, Low, Close, and Volume for the period 1st January 2011 to 31st December 2021 are used as the experimental dataset. The form of the dataset is shown in Table 1.

Table 1. Form of the dataset

Date	Open	High	Low	Close	Volume
2011/1/3	79.75	80.45	79.45	80.05	14119580
2011/1/4	79.8	80.45	79.8	80.3	14060508
2011/1/5	80.8	82.35	80.55	82.25	44631532
2011/1/6	82.95	83.1	82.45	82.7	27712903
2011/1/7	82.8	83.6	82.55	82.7	28182115

Before the raw data is fed into the Random Forest model, we remove the missing values from datasets and construct the independent and dependent variables for Random Forest. Datasets from January 2011 to December 2018 are considered training sets and the remaining datasets from January 2019 to December 2021 are testing sets, which are used to evaluate the performance of Random Forest and strategy profitability of the strategy specifically.

3.2 *Random forest algorithm*

Random Forest is one of the well-known ensemble learning methods, which is a classification or regression algorithm consisting of multiple Decision Trees. During the training process, Random Forest randomly selects sample data with replacement. When building a Decision Tree inside the algorithm, Random Forest tries to find the most suitable variables from the sample data selected randomly to divide the sub-nodes of Decision Trees, instead of building Decision Trees based on all variables [27]. Therefore, each Decision Tree in the Random Forest is independent of each other and has different prediction results. Finally, the results of both regression and classification for Random Forest are based on the average of all constructed Decision Trees.

For the row stock dataset, we try to construct other independent variables. Except for the original variables in the original dataset, we construct the 3-day WMA variable additionally, which represents the weighted average value of *Close* in the past d days at time t. The formula of WMA is shown as follows.

$$WMA(d)_t = \frac{d \times Close_t + (d-1) \times Close_{t-1} + \ldots + 2 \times Close_{t-d+2} + 1 \times Close_{t-d+1}}{d + (d-1) + \ldots + 2 + 1} \tag{1}$$

When $d = 3$, the equation for WMA can be simplified as follows.

$$WMA(3)_t = \frac{3Close_t + 2Close_{t-1} + Close_{t-2}}{3+2+1} \tag{2}$$

Also, the WMA variables for the past five days, $WMA(3)_{t-1}$, $WMA(3)_{t-2}$, $WMA(3)_{t-3}$, $WMA(3)_{t-4}$, and $WMA(3)_{t-5}$ are also used as independent variables for the model. Such independent variables used as inputs for Random Forest make the algorithm identify stock prices' trends more easily. Combined with the original Open, High, Low, Close, and Volume variables, there are twelve independent variables in total. We calculate the difference between the 1-day future price $WMA(3)_{t+1}$ and the current price $WMA(3)_t$ to construct the dependent variable. The formula is calculated as follows.

$$WMA(3)_{t+1} - WMA(3)_t = \frac{3Close_{t+1} - Close_t - Close_{t-1} - Close_{t-2}}{6} \tag{3}$$

As mentioned, there are many noises in stock prices since the prices of stocks are always influenced by multiple investors and financial institutions during trading. To reduce them, some studies [14] use signal decomposition to separate a time series into several parts. However, a novel and efficient method is used in this research. We choose $WMA(3)_{t+1} - WMA(3)_t$ as a dependent variable. According to formula (3), $WMA(3)_{t+1}$ value is 50% in the numerator, and the remaining 50% is split between $Close_t$, $Close_{t-1}$, and $Close_{t-2}$. Using $Close_{t+1}$, $Close_t$, $Close_{t-1}$, and $Close_{t-2}$ as a combination to form the formula, which eliminates noises such as volatility and gap in the raw Close prices in the stock market. Hence, the difference of the WMA_{t+1} and $WMA(3)_t$ has less noise and a more stable trend compared to the direct difference such as $Close_{t+1} - Close_t$. Through the work of Khattak et al. [28], machine learning algorithms always have more accurate results in a trend prediction. So after getting the predicted results of $WMA(3)_{t+1} - WMA(3)_t$, we just need to add the current day $WMA(3)_t$ value to get the model's predicted $WMA(3)_{t+1}$ value for the day ahead, which can be applied in the Bollinger Band Strategy later.

3.3 *Enhanced Bollinger band strategy*

The Bollinger Bands are practical technical indicators designed by American financial analyst John Bollinger in around 1980 and widely used in the global stocks, futures and bond markets. Beginning with a Moving Average line (MA), the Upper Track and Lower Track can be calculated by MA values and the Standard deviation of the past d days. Their formulas are shown as follows.

$$MA(d)_t = \frac{Close_t + Close_{t-1} + \ldots + Close_{t-d+1}}{d} \tag{4}$$

$$Upper\ Track_t = MA(d)_t + k \times sd\left(Close_t,\ Close_{t-1},\ \ldots,\ Close_{t-d+1}\right) \tag{5}$$

$$Lower\ Track_t = MA(d)_t - k \times sd\left(Close_t,\ Close_{t-1},\ \ldots,\ Close_{t-d+1}\right) \tag{6}$$

The parameter d is always set to be 20 and k to be 3 in the Traditional Bollinger Band Strategy. Generally, the price operates in the area between the Upper and Lower Track, and the Track changes while the price and time change. The Traditional Bollinger Band Strategy constructs a trading strategy based only on the *MA* line, *Upper Track*, *Lower Track*, and the *Close* price. The specific algorithm for the Traditional Bollinger Band Strategy is shown in Algorithm 1.

Algorithm 1: Traditional Bollinger Band Strategy

Input: $Upper\ Track_t$, $Lover\ Track_t$, $Close_t$.

if no position opened and $Close_t \leq Lover\ Track_t$:

Open position (going long)

else if a position (going long) is opened and $Close_t \geq Upper\ Track_t$:

Close position

For the Enhanced Bollinger Band Strategy, we use the predicted $WMA(3)_{t+1}$ value combined with the Upper Track and Lower Track to open positions. The parameter d is set to be 20 and k to be 3 in the Enhanced Bollinger Band Strategy. When $WMA(3)_t$ is larger than the Upper Track, we define the stock is in an overvalued stage and we go short position. In contrast, when $WMA(3)_{t+1}$ is under the Lower Track, we define the stock is in an undervalued stage and we go a long position. Simultaneously, in order to solve the failure of the traditional Bollinger Band Strategy to open positions, the ATR indicator is set to stop losing money in the Enhanced Bollinger Band Strategy, and the formula of ATR is provided as follows.

$$TR_t = \text{Max}\left\{\left(High_t - Low_t\right),\ \left|Close_{t-1} - High_t\right|,\ \left|Close_{t-1} - Low_t\right|\right\} \tag{7}$$

$$ATR(n)_t = \frac{TR_t + TR_{t-1} + \ldots + TR_{t-n+1}}{n} \tag{8}$$

The parameter n is always set to 20 in practice. A long position is finished immediately if $WMA(3)_{t+1}$ is less than the Position price minus the triple ATR, and a short position is finished if $WMA(3)_{t+1}$ is larger than the position price plus the triple ATR. Backtrader in Python is applied to run the simulation. By inputting daily stock price data and predicted $WMA(3)_{t+1}$ from test sets, the simulation function in Backtrader is able to set up the Enhanced Bollinger Band Strategy and evaluate its performance. The specific algorithm for the Enhanced Bollinger Band Strategy is shown in Algorithm 2 below.

Algorithm 2: Enhanced Bollinger Band Strategy

Input: $Upper\ Track_t$, $Lover\ Track_t$, $Predicted\ WMA_{t+1}$, ATR_t.

if no position opened and $Predicted\ WMA_{t+1} \leq Lover\ Track_t$:

Open position (going long)

else if no position opened and $Predicted\ WMA_{t+1} \geq Upper\ Track_t$:

Open position (going short)

else if a position (going long) is opened and $Predicted\ WMA_{t+1} < going\ long\ price - 3 * ATR_t$:

Close position

else if a position (going long) is opened and $Predicted\ WMA_{t+1} > Upper\ Track_t$:

Close position

else if a position (going short) is opened and $Predicted\ WMA_{t+1} > going\ long\ price + 3 * ATR_t$:

Close position

else if a position (going short) is opened and $Predicted\ WMA_{t+1} < Lover\ Track_t$:

Close position

4. Experimental results

The whole framework is shown in Figure 1. After using Random Forest to predict the future trend, the predicted WMA is considered an important trading signal in the Enhanced Bollinger Band Strategy. In the following sections, we describe the metrics of Random Forest on ten stocks' datasets and evaluate the performances of the Enhanced Bollinger Band Strategy.

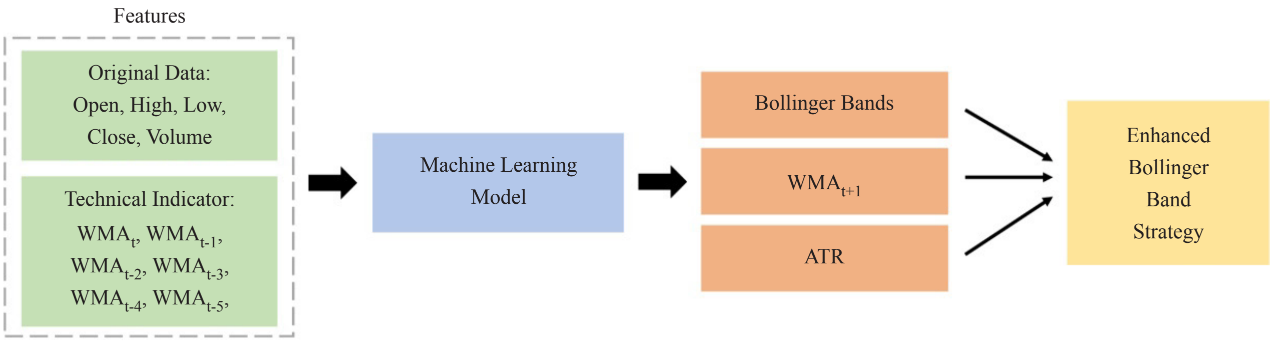

Figure 1. Experimental framework

4.1 *Regression results*

The predicted values $WMA(3)_{t+1}$ are compared with real $WMA(3)$ values in the day $t + 1$. In this research, we choose 3 indicators to evaluate the regression results. The R Squared value describes the comparison of errors between predicted values and mean values, which is usually in the range of 0 to 1. When R Squared approaches 1, it means that using predicted values can get less error. On the contrary, if R Squared approaches 0 or less than 0, the predicted results are extremely terrible. Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) are used to measure the average errors between predicted values and real values. The smaller the value of RMSE and MAE, the smaller the error between the predicted and real values, and the better the model fits. In the formulas below, y_i denotes the real value of the stock WMA price at the moment i, $\hat{y}_i$ denotes the predicted value of the stock WMA price at the moment i obtained by the Random Forest, and n is the total number of test sets. The detailed formulae for these three indicators are shown as follows.

$$R\ Squared = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{9}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{10}$$

$$MAE = \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n} \tag{11}$$

Table 2. Experiment results for ten stocks

Stock Code	R Squared	RMSE	MAE
00005	0.9968	0.6178	0.4896
02888	0.9944	0.8313	0.6508
00011	0.9964	1.4063	1.0606
02356	0.9940	0.2132	0.1649
00023	0.9948	0.2945	0.2425
02388	0.9953	0.2295	0.1726
00939	0.9888	0.0507	0.0389
03328	0.9960	0.0480	0.0363
01288	0.9913	0.0305	0.0216
00998	0.9934	0.0468	0.0370

Figure 2. Predicted WMA(3) price of 02888 stock

To implement Random Forest, we use Scikit-Learn in Python to predict the trend of each stock in the experiments first. Then we present the prediction performances of the ten stocks' $WMA(3)_{t+1}$ in Table 2, in which we find that the models predict perfectly for all ten stocks' price trends $WMA(3)_{t+1}$. The R Squared values of almost all models even reach 0.99. Moreover, in terms of the error evaluation indicators RMSE and MAE for real and predicted values, MAE in some stocks is basically less than 0.04. For the remaining stocks, the range is always from 0.2 to 1. For RMSE, most of the values are less than 1, except for the 00011 stock with relatively large errors. According to these results, we are successful in constructing independent and dependent variables for Random Forest. In Figure 2, we select the 02888 stock as an example and plot both the predicted $WMA(3)_{t+1}$ and real $WMA(3)_{t+1}$. The two different lines in this figure have almost the same trend. So the predicted $WMA(3)_{t+1}$ values manage to provide accurate trading signals in the Enhanced Bollinger Band Strategy precisely.

4.2 *Trading simulations*

The initial balance of HK$100,000 is prepared before trading. In addition, we use 4 technical indicators such as Model Return, Basic Return, Max Drawdown, and Sharpe Ratio to evaluate the performance of the Traditional Bollinger Band Strategy and Enhanced Bollinger Band Strategy. The model Return is obtained by calculating the final return as a percentage of the initial balance of HK$100,000. Similarly, the Basic Return is the return between all money that is bought into the stock by default at the beginning and sold at the end. For all investors, quantitative trading models perform better as returns are higher. In the Hong Kong stock market, a return of 10-20% or more over a 2-year test set is pretty good. The Max Drawdown describes the maximum loss that each investor may face during trading, which reflects the model's ability to control risks. For this indicator, all investors hope it to be lower. The Sharpe Ratio measures how many extra returns an investor can get when they take on each unit of the total risk. Similarly, the larger the indicator, the higher the return. We present all trading indicators' results for the ten stocks by using the Traditional Bollinger Band Strategy in Table 3, and the trading indicators' results by using the Enhanced Bollinger Band Strategy in Table 4. Meanwhile, 02888 stock is selected as an example, and plot its detailed Enhanced Bollinger Band Strategy trading situations in Figure 3.

Table 3. Returns of the Traditional Bollinger Band Strategy

Stock Code	Model Return (%)	Basic Return (%)	Max Drawdown (%)	Sharpe Ratio
00005	-43.1752	-25.9084	52.3963	-0.8611
02888	-10.4455	-22.1380	49.1928	-0.1685
00011	-14.8970	-17.5145	31.0313	-0.4612
02356	-25.4098	-51.3869	45.8429	-0.4784
00023	-37.2201	-54.4715	40.5196	-0.4687
02388	-10.7122	-10.1933	28.7999	-0.4061
00939	3.0356	-13.8756	16.6913	0.0404
03328	-4.9750	-20.9732	21.2508	-3.7281
01288	-18.0846	-20.7101	26.3050	-2.5081
00998	-30.5709	-27.4678	31.6509	-1.9588

Table 4. Returns of the Enhanced Bollinger Band Strategy

Stock Code	Model Return (%)	Basic Return (%)	Max Drawdown (%)	Sharpe Ratio
00005	12.9106	-25.9084	24.1647	0.2466
02888	19.8056	-22.1380	33.1721	0.5907
00011	14.8107	-17.5145	13.6669	0.3810
02356	25.0074	-51.3869	21.9033	0.6045
00023	58.5377	-54.4715	16.9019	1.4455
02388	10.4150	-10.1933	23.7936	0.2046
00939	30.2470	-13.8756	22.6699	0.7094
03328	1.5205	-20.9732	26.7088	0.0560
01288	39.8460	-20.7101	20.9400	1.0915
00998	15.6718	-27.4678	26.7116	0.3018

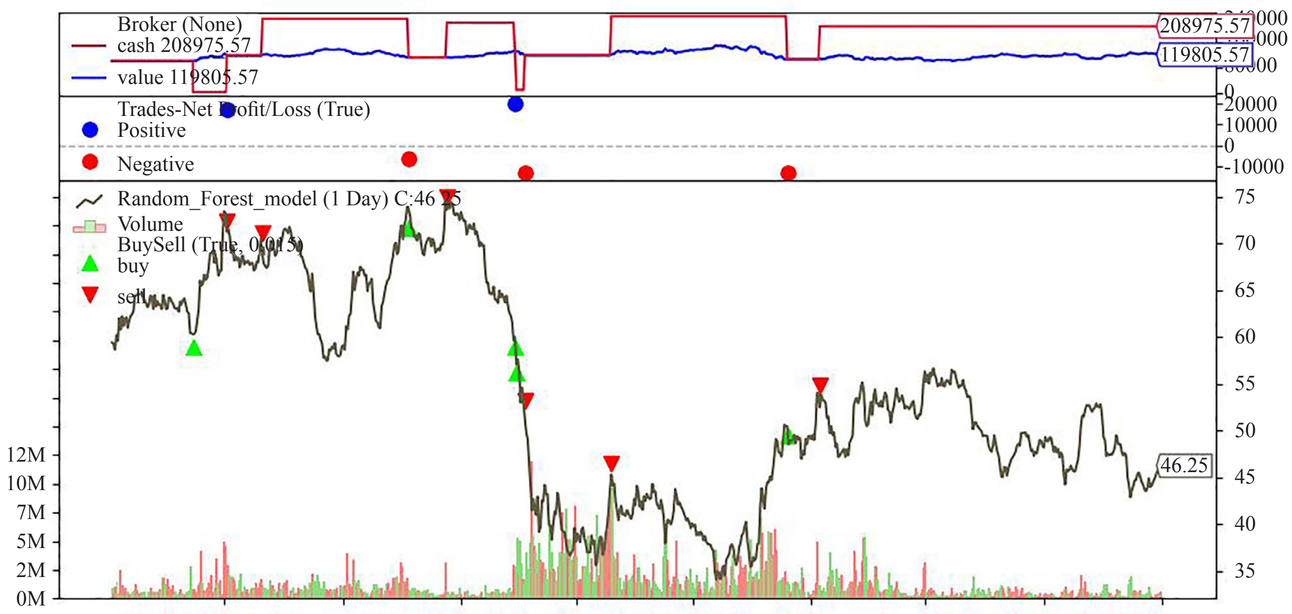

Figure 3. Trading simulation of 02888 stock

Due to the impression of the COVID-19 outbreak and the volatile world situation, the Hong Kong stocks market has been in an oscillating downward trend in the testing set and all of these financial banking stocks' prices decrease during this period. As the detailed trading situations of 02888 stock in Figure 3, the black line shows the trend of the Close price during the trading period. Green markers and red markers are the buying and selling points for the stocks in the Enhanced Bollinger Band Strategy, and the blue line indicates the stock's profits. Besides, 4 technical indicators for the Traditional Bollinger Band Strategy and Enhanced Bollinger Band Strategy on ten stocks are listed in Table 3 and Table 4. As the Basic Return shows in these two tables, when we buy the stocks with all the money at the beginning and sell them at the end, we always lose around 20-30% of our capital, in which losses reach 50% or above in 02356 and 00023. In Table 3, when we simulate trading with the Traditional Bollinger Band Strategy, almost all the stocks are losing money except for 00939 stock. However, compared with the Basic return, Traditional Bollinger Band Strategy's overall performance is still losing less money. When we use the Enhanced Bollinger Band Strategy, we not only get rid of losing money but also succeed in earning profits during the simulation period in Table 4. By capturing the trading signals to open positions through the predicted $WMA(3)_{t+1}$, as well as the stop-loss and stop-profit signals set by the ATR and Bollinger Bands, we manage to achieve a 10-20% return on most stocks, which is extremely better than the Traditional Bollinger Band Strategy and Base Return. To observe these three returns easily, the returns of all ten stocks are shown as a boxplot in Figure 4. What's more, this strategy is surprisingly able to achieve a 58% return on the 00023 stock of which the Base Return loses more money, which shows that the Enhanced Bollinger Band Strategy is specialized in capturing shorting signals for investors to refer to. In terms of the Max Drawdown, they show a decline of around 20%, which may be a little high for the average investors. However, the Max Drawdown of the Traditional Bollinger Band Strategy is much higher than Enhanced Bollinger Band Strategy, which brings more extra risk to investors. Also, the high Max Drawdown in this strategy may be caused by shocks in the stock market, and it is believed that the Max Drawdown in this strategy would be relatively low in a large market. For Sharpe Ratio, this strategy performs well in the 00023 and 01288 stocks but at a lower level in other stocks. Sharpe Ratio is related to the nature of stocks themselves. As defensive stocks in the stock market, financial banking stocks have lower levels of relative volatility and profitability compared to technology stocks and popular stocks. Therefore, the Sharpe Ratio for investing in such kinds of stocks is never high. At last, the application of the Enhanced Bollinger Band Strategy provided in this research achieves generating more profits and gives investors a reference for combining Random Forest with quantitative investing.

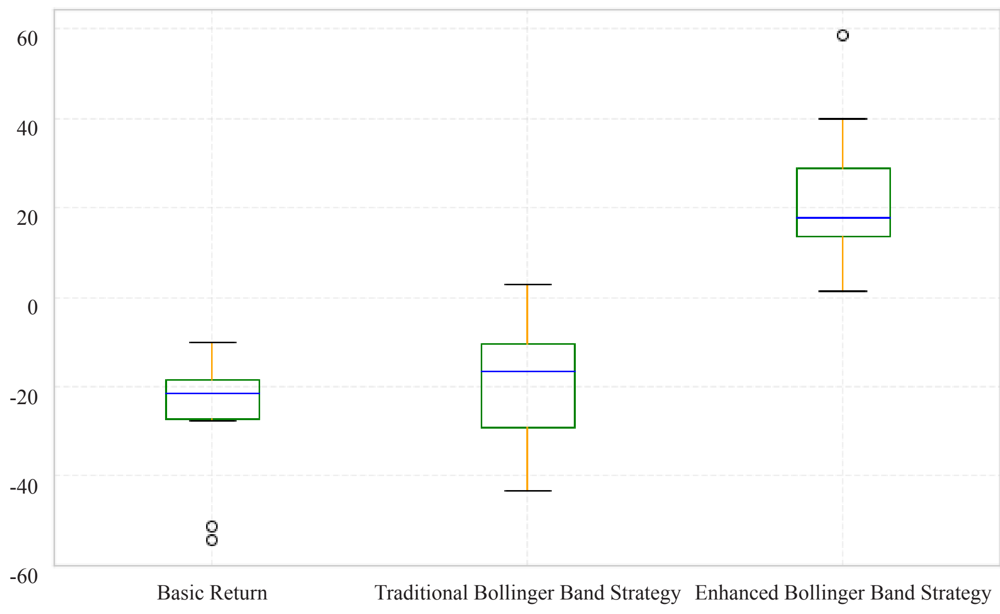

Figure 4. Comparison of three returns

5. Conclusion

Based on the prices of different kinds of stocks, this research proposes a Random Forest model to forecast the 1-day future price with the *WMA(3)* spread, which can be calculated by a simple additive transformation. The experimental results from the regression analysis show that the data of all ten stocks have satisfying prediction results under this modeling method, which also provides a reliable reference for the establishment of a quantitative strategy. In the next step, combined with the predicted *WMA*(3) and the Bollinger Band Strategy, the Enhanced Bollinger Band Strategy succeeds in getting positive returns for all the stocks. The predicted trend of stock and risk control makes the Enhanced Bollinger Band Strategy earn profits from ten financial banking stocks from Hong Kong Exchanges and Clearing Limited. The performance is much better than the Traditional Bollinger Band Strategy and Basic Return, which always loses money during simulation. Therefore, the Enhanced Bollinger Band strategy proposed in this research can provide investors a very useful inspiration and reference in practice. In future research, we can try to use some machine learning or deep learning algorithms to predict different trading signals such as Relative Strength Index signals and stocks' upside/downside signals. Moreover, combined with new predicted signals and other well-known trading strategies such as Moving Average Strategy, MACD Strategy, and Sentiment Trading Strategy, we can explore the impact of different combinations of quantitative trading strategies on stocks' return and risk control.

Conflict of interest

The authors declare no competing financial interest.

References

[1] Firza Y, Dirgantoro B, Setianingsih C. Bollinger bands and single exponential smoothing methods in the decision system of selling and buying stock. *2021 3rd International Conference on Electronics Representation and Algorithm*. Yogyakarta, Indonesia; 2021. p.173-178. Available from: https://doi.org/10.1109/ICERA53111.2021.9538632.

[2] Day MY, Cheng Y, Huang P, Ni Y. The profitability of bollinger bands trading bitcoin futures. *Applied Economics Letters*. 2022; 1-7. Available from: https://doi.org/10.1080/13504851.2022.2060494.

[3] Wu S. Application of cluster analysis in stock selection in united states stock market. *Proceedings of the 2020 11th International Conference on E-Education, E-Business, E-Management, and E-Learning*. Association for Computing Machinery, New York, United States; 2020. p.310-313. Available from: https://doi.org/10.1145/3377571.3377628.

[4] Jaiswal S, Jaiswal RT. Review on machine learning techniques for stock-market forecasting. *Artificial Intelligence Evolution*. 2020; 1(1): 34-47. Available from: https://doi.org/10.37256/aie.112020317.

[5] Xu S, Han C, Ran C. A time series combined forecasting model based on prophet-lgbm. *2021 2nd International Conference on Artificial Intelligence and Information Systems*. Association for Computing Machinery, New York, United States; 2021. p.1-6. Available from: https://doi.org/10.1145/3469213.3470280.

[6] Hu G, Cai C. A trade data predictive analytics approach for global digital trade. *2021 4th International Conference on Computing and Big Data*. Association for Computing Machinery, New York, United States; 2021. p.86-91. Available from: https://doi.org/10.1145/3507524.3507539.

[7] Zhu J, Huang J, Li H. Prediction of vehicle registration quantity based on arima model. *2020 9th International Conference on Computing and Pattern Recognition*. Association for Computing Machinery, New York, United States; 2020. p.508-512. Available from: https://doi.org/10.1145/3436369.3436451.

[8] Du J, Zhang JP. Research on stock risk based on quantile regression VaR model. *Statistics and Application.* 2018; 7(4): 407-421. Available from: https://doi.org/10.12677/sa.2018.74047.

[9] Khairalla MA, Ning X. Financial time series forecasting using hybridized support vector machines and arima models. *2017 International Conference on Wireless Communications, Networking and Applications*. Association for Computing Machinery, New York, United States; 2017. p.94-98. Available from: https://doi.org/10.1145/3180496.3180613.

[10] Li J, La L. Analysis of the rise and fall of international futures based on xgboost algorithm. *Finance.* 2018; 8(5): 211-220. Available from: https://doi.org/10.12677/fin.2018.85025.

[11] Liu J, Sun D. Stock price prediction based on attention mechanism and long short-term memory network. *Advances in Applied Mathematics*. 2021; 10(12): 4379-4385. Available from: https://doi.org/10.12677/aam.2021.1012466.

[12] Vicente ÓF, Rebollo FF, Polo FJG. Deep q-learning market makers in a multi-agent simulated stock market. *2021 Second ACM International Conference on AI in Finance*. Association for Computing Machinery, New York, United

States; 2021. p.1-9. Available from: https://doi.org/10.1145/3490354.3494448.
[13] Zhang C, Ji Z, Zhang J, Wang Y, Zhao X, Yang Y. Predicting chinese stock market price trend using machine learning approach. *2018 2nd International Conference on Computer Science and Application Engineering*. Association for Computing Machinery, New York, United States; 2018. p.1-5. Available from: https://doi.org/10.1145/3207677.3277966.
[14] Yan K, Li Y, Wang Y. Stock price prediction based on signal decomposition and machine learning models. *Computer Science and Application*. 2022; 12(4): 1080-1088. Available from: https://doi.org/10.12677/csa.2022.124111.
[15] Zhao W, Huang H, Shen J, Xing Y. Comparison between time series analysis and mode decomposition on the prediction of bank of china stock price. *2021 7th International Conference on Industrial and Business Engineering*. Association for Computing Machinery, New York, United States; 2021. p.168-173. Available from: https://doi.org/10.1145/3494583.3494644.
[16] Sood S, Zeng Z, Cohen N, Balch T, Veloso M. Visual time series forecasting: an image-driven approach. *2021 Second ACM International Conference on AI in Finance*. Association for Computing Machinery, New York, United States; 2021. p.1-9. Available from: https://doi.org/10.1145/3490354.3494387.
[17] Hou Y. Stock price prediction based on bi-lstm deep learning. *Statistics and Application*. 2021; 10(3): 538-546. Available from: https://doi.org/10.12677/sa.2021.103055.
[18] Ge J, Qin Y, Li Y, Huang Y, Hu H. Single stock trading with deep reinforcement learning: A comparative study. *2022 14th International Conference on Machine Learning and Computing*. Association for Computing Machinery, New York, United States; 2022. p.34-43. Available from: https://doi.org/10.1145/3529836.3529857.
[19] Li G. Futures trading strategy test based on computer programming. *2020 3rd International Conference on Data Science and Information Technology*. 2020. p.159-164.
[20] Abouloula K, Krit SD. Using a robot trader for automatic trading. *2018 4th International Conference on Engineering & MIS*. Association for Computing Machinery, New York, United States; 2018. p.1-9. Available from: https://doi.org/10.1145/3234698.3234701.
[21] Attanasio G, Cagliero L, Garza P, Baralis E. Quantitative cryptocurrency trading: Exploring the use of machine learning techniques. *2019 5th Workshop on Data Science for Macro-modeling with Financial and Economic Datasets*. Association for Computing Machinery, New York, United States; 2019. p.1-6. Available from: https://doi.org/10.1145/3336499.3338003.
[22] Xiao D, Deng X, Xia B. A statistical arbitrage strategy in forex market based on high-frequency data. *Finance*. 2017; 7(1): 38-46. Available from: http://doi.org/10.12677/fin.2017.71005.
[23] Zhang L. Pair trading with machine learning strategy in china stock market. *2021 2nd International Conference on Artificial Intelligence and Information Systems*. Association for Computing Machinery, New York, United States; 2021. p.1-6. Available from: https://doi.org/10.1145/3469213.3471353.
[24] Yue P, Li X. Research on stock index futures arbitrage strategy based on convolutional neural network. *Advances in Applied Mathematics*. 2021; 10(2): 557-567. Available from: https://doi.org/10.12677/aam.2021.102061.
[25] Seshu V, Shanbhag H, Rao SR, Venkatesh D, Agarwal P, Arya A. Performance analysis of bollinger bands and long short-term memory (lstm) models based strategies on nifty50 companies. *2022 12th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*. Noida, India; 2022. p.184-190. Available from: https://doi.org/10.1109/Confluence52989.2022.9734127.
[26] Yahoo Finance. 2022. Available from: https://finance.yahoo.com/.
[27] Géron A. *Hands-on machine learning with scikit-learn, keras, and tensorFlow: Concepts, tools, and techniques to build intelligent systems*. 2nd ed. O'Reilly Media, Inc; 2019.
[28] Khattak AM, Ullah H, Khalid HA, Habib A, Asghar MZ, Kundi FM. Stock market trend prediction using supervised learning. *2019 10th International Symposium on Information and Communication Technology*. Association for Computing Machinery, New York, United States; 2019. p.85-91. Available from: https://doi.org/10.1145/3368926.3369680.