International Review of Economics and Finance 71 (2021) 943–954

Contents lists available at ScienceDirect

International Review of Economics and Finance

journal homepage: www.elsevier.com/locate/iref

Implied volatility forecast and option trading strategy

Dehong Liu [a], Yucong Liang [a], Lili Zhang [a,*], Peter Lung [b], Rizwan Ullah [a]

[a] School of Economics and Management, Beijing Jiaotong University, China
[b] Reiman School of Finance, University of Denver, United States

ARTICLE INFO

JEL classification:
C6
D9
G1

Keywords:
ABC-BP neural network model
Implied volatility
Options trading strategy

ABSTRACT

We examine the implied volatility derived from an improved Artificial Bee Colony with Back Propagation (BP) neural network model that is Artificial Bee Colony-Back Propagation (ABC-BP) neural network model. We find that the improved model can better predict the implied volatility than basic BP neural network model and Monte Carlo simulation. Nevertheless, the option price derived from the Monte Carlo simulation is more efficient when we apply the simulation to the option straddle trading strategy. Additionally, in a robustness test we find that our proposed neural network model performs better than the traditional GARCH model in building up option trading strategies.

1. Introduction

Literature has well documented that the estimation of implied volatility (hereafter, IV) is crucial in risk management, derivatives pricing (i.e., Muzzioli, 2010). When all other option parameters are known, there is a one-to-one relationship between option prices and the underlying expected asset volatility. This yields the so-called implied volatility (Giot & Pierre, 2009). If option markets are informationally efficient and the option pricing model is correct, implied volatility should subsume the information contained in other variables in explaining future volatility (Jiang & Tian, 2005).

Implied volatility is often referred to as the market's volatility forecast and is said to be forward looking as opposed to historical methods which are by definition backward looking (Koopman, Jungbacker, & Hol, 2005). And studies on IV estimation have three major strands: backwardation derivation according to derivatives prices, time-series volatility models, and the artificial neural network models. For the backwardation method, Ser-Huang and Clive (2005) and Harvey and Whaley (1992) find IV contains the predictability of realized volatility in the future. Current literature also shows that model-free implied volatility (hereafter, MFIV) can yield the best forecasting performance both during normal and extreme market conditions (Markose et al., 2012). MFIV subsumes information contained in the Black-Scholes (hereafter, BS) and past realized volatility (Huang & Zheng, 2009; Jiang & Tian, 2005). But there has been a considerable debate on the use of the BS implied volatility as an unbiased forecast for future volatility of the underlying asset (Jiang & Tian, 2005).

For time-series volatility models, GARCH model is one of the most popular approaches to predict volatility. Zhang et al. (2009), for example, use GARCH family models to forecast the volatility of crude oil. Engle and Sokalska (2012) use similar time-series model to predict intraday volatility in the US equity market and Fang et al. (2017) apply GARCH to analyze the volatility in gold futures market.

* Corresponding author. School of Economics and Management, Beijing Jiaotong University No.3 Shangyuancun, Haidian District, Beijing, 100044, China.

E-mail addresses: dhliu@bjtu.edu.cn (D. Liu), 19120493@bjtu.edu.cn (Y. Liang), llzhang@bjtu.edu.cn (L. Zhang), Pei.lung@du.edu (P. Lung), rizwan.ullah.bjtu@gmail.com (R. Ullah).

https://doi.org/10.1016/j.iref.2020.10.023
Received 21 March 2020; Received in revised form 18 October 2020; Accepted 25 October 2020
Available online 27 October 2020

The impact of volatility demand on option prices is positive (Sophie et al., 2008). Trading volatility can pay off. In an artificial world without transaction costs both delta-neutral and straddle trading strategies lead to significant positive profits, regardless of which volatility prediction method is used, namely implied volatility and GARCH volatility (Guo, 2000).

According to the artificial neural network models, Andreou et al. (2006) show that neural network model with Huber function outperforms optimized least square models. Sermpinis et al. (2013) suggest that higher-order neural networks (HONNs) show a good performance in forecasting the future realized volatility of the FTSE 100 futures index. Jiang (2002) proposes a parsimonious GMM(Generalized method of moments) estimation for continuous time option pricing models with stochastic volatility, random jump and stochastic interest rate. Bollen and Whaley (2004) examined the relation between net buying pressure and the shape of the implied volatility function (hereafter, IVF) for index and individual stock options and verified that implied volatility of stock options is dominated by call option demand. More recently, Xu (2017) analyzes the key determinants of S&P 500 volatility at both monthly and daily frequencies. Therefore, there are many applications of neural models to solve financial problems, but there are few papers to solve the volatility problem.

In sum, IV estimation approaches have been widely applied to many option trading strategies for both speculation and risk management. However, current literature has not examined the performance of the Artificial Bee Colony-Back Propagation (hereafter, ABC-BP) neural network model in IV estimation. The model takes a part of the training value in the data range as the experimental object and uses another part of the measured value as the test object. In the current paper, we study the performance of this relatively new model in IV predictability. Our results suggest that ABC-BP neural network model can more efficiently and effectively measure IV. More importantly, the model is applicable to various option trading strategies.

The structure of this paper is as follows: Part II introduces the BP neural network model, ABC-BP neural network model, Monte Carlo simulation and traditional volatility calculation models, that is, BS and GARCH models. Part III outlines the data source and data error. The prediction results of implied volatility under the ABC-BP neural network model are also processed in Part III, and the option strategy is introduced. And finally, Part IV concludes the paper.

2. Models

2.1. Back propagation (BP) neural network model

The back propagation (hereafter, BP) neural network model is a multilayer feed-forward network trained according to the error backpropagation algorithm and is one of the most widely applied neural network models (Li et al., 2012). BP neural network model is a three-layer feed-forward network composed of the input layer, the hidden layer, and the output layer. Each layer contains several disconnected neurons nodes, and the adjacent nodes are connected according to certain weight values. The direction of the transfer of information is from the input layer to the hidden layer to the output layer. If the difference between the actual output and the expected output cannot meet the required error, the error value will be fed back layer by layer along the network path, and the connection weights and thresholds of each layer will be corrected.

In BP type models, the input vector is $I(I_1,\ldots,I_i,\ldots,I_p)$, and the output vector is $O(O_1,\ldots,O_j,\ldots,O_t)$. The weight of the input node i and the hidden node j is H_{ij}, and the weight of the output hidden node j and the output node k is E_{jk}, if the given training data is $D = \{(I(i), T(i)\}$, where $I(i)$, $T(i)$ are the i^{th} input value and the expected output value in the training set respectively, and the entire training is continuously adjusting the parameters of the neural network. The error is gradually reduced until the required conditions are satisfied. According to the principle of supervised learning with neural network, the training error can be described as follows:

$$E_d(W) = \frac{1}{2}\sum_{j=1}^{N}\sum_{i=1}^{Q}\left[T_i(j) - O_j(I^{(j)}|w)\right]^2 \tag{1}$$

N and Q respectively represent the number of data in the neural network training set and the number of neurons in the output layer. After training the neural network, arbitrarily input a data set $I_{(N+1)}$ to obtain the corresponding predicted data $O(I_{N+1}|w)$. The number of hidden units is calculated according to the following formula:

$$n_{input} = \varepsilon + \sqrt{n_{input} + n_{output}} \tag{2}$$

ε is a constant between 1 and 10.

2.2. Artificial bee colony algorithm optimization of BP neural network model

2.2.1. Principle of ABC algorithm

The Artificial Bee Colony (hereafter, ABC) algorithm was proposed in 2005 by Dervis Karaboga, inspired by the bees' intelligent behavior. As an optimization algorithm, it provides a population-based search process. Madeleine and Bin (2008) revealed that bees rely on waggle dance to convey information about the source of food. In the ABC algorithm, the position of the food represents a possible solution in the optimization problem, and the amount of nectar represents the quality or fitness of the corresponding solution.

First, the algorithm randomly generates the initial population, N_s solutions $(X_1,\ldots,X_N)$, it then sets the limit (*limit*) and the maximum cycles number (*MCN*). After initialization, the bees begin a cyclic search: employed bees carry out neighborhood search of old solutions using the greedy mechanism. If the fitness of the new solution is greater than the fitness of the old solution, the bees will forget the old solution and remember the new solution. And then calculate the possible values (P_i) of these new solutions, and onlookers start to search for new solutions near this possible value (P_i) and the solutions remembered by employed bees using the greedy mechanism. If the latest solution obtained can no longer be updated (the number of updates exceeded the limit), then the solutions will be abandoned by the scouts and replaced with a new solution. This cycle is up to the maximum cycle number. There is only one scout per cycle.

2.2.2. BP neural network model based on ABC algorithm (ABC-BP)

In this paper, we integrate the ABC algorithm with BP neural network model. We use this algorithm to find the optimal network weight and threshold. This paper also use this algorithm which combines the generalization performance of neural network, the global iteration, with local search ability of ABC algorithm.

We present the calculation method of the ABC-BP neural network model in the appendix.

2.3. Monte Carlo simulation

The principle of Monte Carlo simulation option pricing is: according to the given price movement process of the underlying asset, simulate the price change of the underlying asset, and when the simulation times reach a certain number, obtain the expectation by taking the mean value. According to the law of large Numbers, the Monte Carlo simulation results finally satisfy the convergence. Compared with BS model, this method has the following advantages: Firstly, it is flexible, easy to implement and improve; secondly, the error and convergence speed of the simulation estimate are independent of the dimension of the problem solved, which can solve the multi-asset option pricing and path dependent option pricing problems well. However, in order to achieve high accuracy, thousands of simulations are generally required, so the Monte Carlo method usually takes much more time than the BS model.

For comparison, this paper also examines the Monte Carlo method of option pricing whose theoretical basis is risk neutral pricing: under the risk neutral measure, the option price can be expressed as the expected value of the discount for its maturity return, that is:

$$P = E^{Q}[\exp^{(-rT)} f(S_1, S_2, \cdots, S_T)] \tag{3}$$

where E^{Q} represents the risk neutral expectation, r is the risk-free rate, and T is the expiration date, $f(S_1, S_2, \cdots, S_T)$ is the expected return on the underlying asset.

Calculating the option price is to calculate an expected value, and the Monte Carlo simulation method is used to estimate the expected value, that is Monte Carlo simulation method of option pricing. The Monte Carlo simulation method for option pricing includes the following four steps:

a) Divide the time interval into n sub-intervals, and the discrete form of calculating the underlying asset price are:

$$S^{j}(t_{i+1}) e^{\left(r - \frac{1}{2\sigma^2}\right)(t_{i+1} - t_i) + \sigma\sqrt{t_{i+1} - t_i} Z_i}, \ Z_i \sim N(0, 1) \tag{4}$$

$S^{j}(t_{i+1})$, $S^{j}(t_i)$ represent the price of underlying asset at the time of t_{i+1} and t_i receptively.

b) Calculate the return to maturity of option under this path and the discount of the return using the risk-free rate:

$$C^{j} = \exp^{(-rT)} \max\{0, S_K^{j} - K\} \tag{5}$$

c) Repeat the first two steps to obtain a large number of samples of the discounted value of option returns.
d) Calculate the mean of samples, and get the values of the option price using Monte Carlo simulation:

$$C_{MC} = \frac{1}{m} \exp^{(-rT)} \sum_{j=1}^{m} C^{j} = \frac{\exp^{(-rT)} \sum_{j=1}^{m} \max\{0, S_T^{j} - K\}}{m} \tag{6}$$

Where C_{MC} is one of the values calculated by Monte Carlo simulation. This paper carries out Monte Carlo simulation of the predicted implied volatility and calculates the logarithmic rate of return of the S&P500 index from January 2017 to May 2018, and then uses it as the expected rate of return and finally predicts closing price of the S&P500 option.

2.4. Estimation of implied volatility based on traditional model

Considering many time series models, the family of GARCH models is still the most popular, and the most commonly used in a financial modeling. Jiang and Tian (2005) suggested that the information content of implied volatility is not necessarily higher than that of standard time series models such as GARCH/EGARCH. And (Pilbeam & Langeland, 2015) indicated that a GARCH (1,1) model is especially shown to be a good parameterization of the process. Moreover, the complex model is not necessarily better than the simple model, like the simple GARCH model volatility which is the representative of the time series model. What we use in estimating equations

Table 1
Descriptive analysis.

	Maximum	Minimum	Mean	Median	Std.Dev
PRICE	208.5900	0.050000	88.92787	91.50000	46.02796
RATE	232.3000	7.600000	62.06735	32.10000	62.48160
STRIKE	2875.000	1450.000	2109.478	2075.000	344.6443
TIME	1.936986	0.000000	0.716983	0.720548	0.429604
UNDERLYING PRICE	2872.820	1457.150	2104.769	2071.260	334.6420
VIX	0.321800	0.090100	0.145826	0.137800	0.036855
IV	0.393900	0.048700	0.134003	0.133100	0.026361

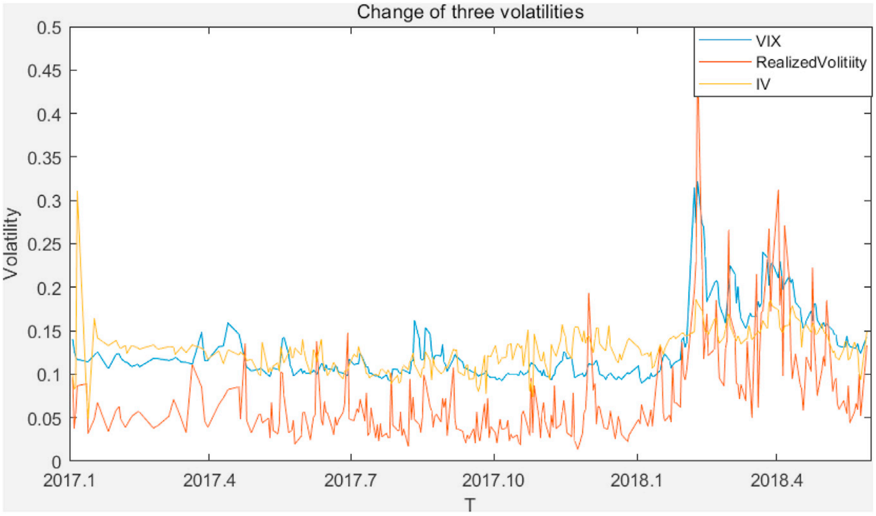

Fig. 1. Volatility comparison.

is rolling estimate method, namely for forecasting volatility of week (month), respectively, to select 120 daily returns before the week (month) to establish a GARCH model to forecast individually, rather than building a model for all week (month) volatility forecast. And the rolling estimation method can further improve the model's predictive ability, because the prediction of the variance is dynamic in this approach. We use GARCH (1,1) model for daily data:

$$h_{t+1} = \omega + \alpha \times a_t^2 + \beta h_t \tag{7}$$

where h_{t+1}, h_t is the predicted variance at time $t+1$ and t respectively, a_t is the predicted residual at time t, and α, β are both random parameters.

It is assumed that the first day to be predicted from the $t+1^{th}$ day, the predicted residual a_t and the predicted variance h_t obtained from the t^{th} day can predict the conditional variance h_{t+1} of the $t+1^{th}$ day, but there is a problem when predicting the conditional variance of the second day, that is, starting from the t^{th} day, the prediction residual of the $t+1^{th}$ day is not known. At this time, the predicted rate of return on the $t+1^{th}$ day is used to replace the true rate of return on the $t+1^{th}$ day, thereby obtaining the residual a_{t+1} on the $t+1^{th}$ day, and placing it into the conditional variance model.

In the same way, the predicted variance h_{t+1} of the $t+2^{th}$ day is obtained. By analogy, the variance of the rate of return for the day can be predicted, and this prediction method is the dynamic forecasting method. It is conceivable that since the predicted value of the mean equation return is substituted for its true value, each prediction error is added into the subsequent prediction, and the longer is the prediction period, the worse is the prediction result.

3. Prediction results and Option Trading Strategy

3.1. Data and statistics

This study is based on the S&P 500 index options accessed through the Bloomberg terminal, and selecting the daily trading data from January 8th, 2013 to May 29th, 2018.

In order to ensure the accuracy and efficiency, the steps taken to filter the data are as follows:

Table 2
Correlation among three volatility.

	IV	Realized Volatility	VIX
IV	1	0.5028	0.7220
Realized Volatility	0.5028	1	0.6988
VIX	0.7220	0.6988	1

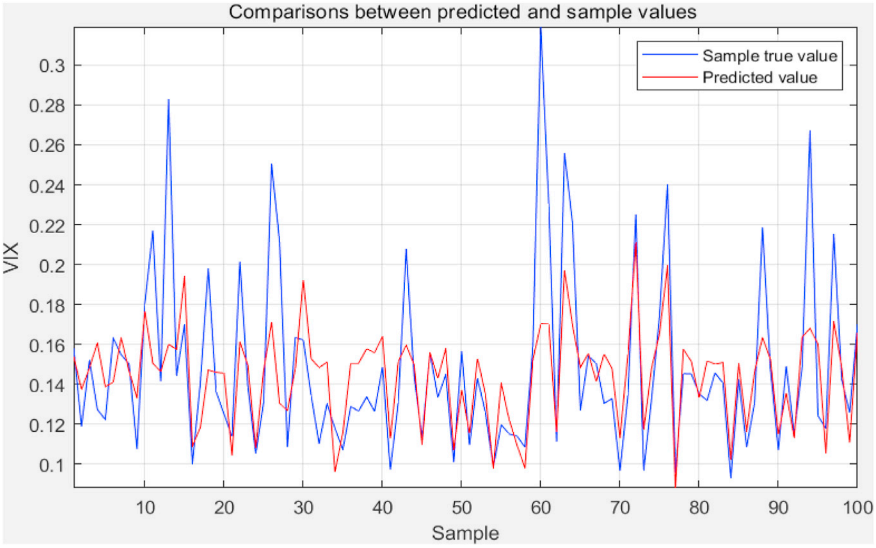

Fig. 2. Comparisons between predicted and sample true values based on ABC-BP neural network model.

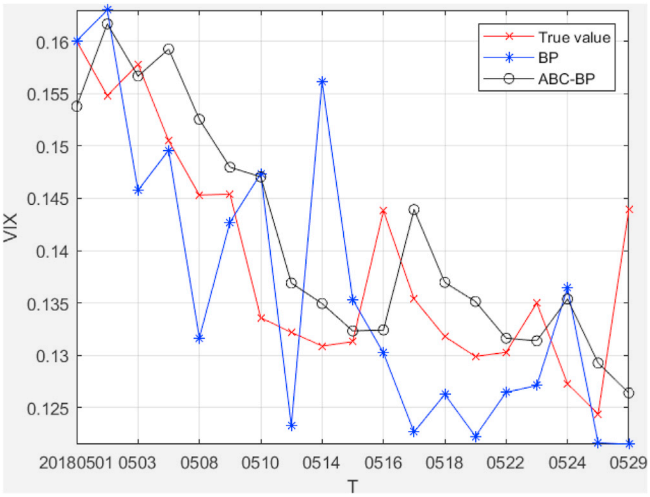

Fig. 3. Comparisons based on BP and ABC-BP neural network model optimal and true value.

a) The transaction volume is not 0;
b) The price of the option cannot be lower than \$0.05;
c) Select at-the-money option, $0.45 < delta < 0.55$, and the delta ratio is the percentage change in the option premium for each dollar change in the underlying, and in-the-money option, $delta < 0.45$, and out-the-money option, $delta > 0.55$. And descriptive statistics of the data are shown in Table 1.

The volatilities include realized volatility, implied volatility, and VIX (Volatility Index). This paper selects the VIX that best substitute for implied volatility as the training set which will be used in ABC-BP neural network model (Fig. 1).

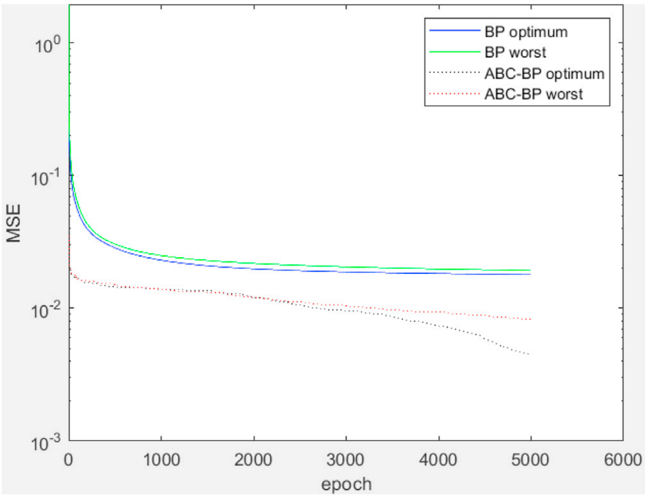

Fig. 4. Comparison of MSE based on BP neural network model and ABC-BP neural network model. Note: Fig. 4 shows that MSE decreases with the increase of iterations, and the MSE of ABC-BP neural network model is obviously smaller than that of BP neural network model under the optimal fitting condition.

Table 3
Error comparison under three models.

	MSE	MAE	MBE
BS	0.001	0.024	0.001
BP	0.000**	0.010**	−0.000**
ABC-BP	0.000***	0.010***	0.000***

The VIX represents market's expectations for volatility over the coming 30 days. It is composed of eight groups of options for the call option and put option that are closest to the at-the-money option of the S&P 500 index option in the near month and the next month, whose implied volatility is worked out respectively, and the VIX is obtained using weighted average method. It provides market participants with an indicator that is more reflective of the overall market trend. The correlation among three kinds of volatilities can be seen in Table 2.

3.2. Out-of-sample testing

We calculated the correlation between historical volatility, implied volatility and VIX, and found that the correlation coefficient of implied volatility and VIX was higher than that of implied volatility and historical volatility, so the subsequent volatility fitting was fitted and tested using VIX as the standard. And Fig. 2 suggests that the model have predictability which can be shown in the comparisons between predicted and sample true values based on ABC-BP neural network model.

Comparing the BP neural network model in Fig. 3 with the findings in Fig. 2, we find that ABC-BP neural network model is closer to the VIX, which reveals that the volatility prediction under this method is more optimal. In addition, we compare the volatility predicted by the above two models using the following indicators, mean square error (*MSE*), mean absolute error (*MAE*), mean bias error (*MBE*). The mean square error (*MSE*) of the two models decreases as the number of iterations increases, and it is shown in Fig. 4 more intuitively:

In Fig. 3, we find that MSE decreases with the increase of iterations, and the MSE of ABC-BP neural network model is clearly smaller than that of BP neural network model under an optimum fit. So, the validity of model prediction is illustrated.

To further illustrate the advantages of the ABC-BP neural network model, we use three predictive indicators of MSE, MAE and MBE in Table 3. The calculation of these three indicators using formulas (8-10), are listed in the following. Starting with the mean square error (*MSE*), such as:

$$MSE = \sum_{i=1}^{N} \frac{(I_i - O_i)^2}{N} \times 100\% \tag{8}$$

In this formula, $i = 1, 2, \cdots N$, N indicates the number of predicted data in the experiment, I_i and O_i denote the i^{th} input and output data(the predicted value and the real value) respectively.

The formula for the mean absolute error is as follows:

Table 4
Descriptive Analysis of option prices in Monte Carlo simulation.

	Maximum	Minimum	Mean	Median	Std.Dev
Underlying price	2872.820	2257.830	2530.086	2648.980	234.174
Call	1611.265	108.0463	1088.706	1334.268	395.1518
Put	1121.573	−110.7903	633.1015	593.3589	292.2569
delta	0.970963	0.536306	0.757982	0.812611	0.126825
vega	1234.165	171.0937	864.9403	876.3806	199.771

Table 5
Annual performance of options volatility trading strategies.

Y	Rate of return	Sharp Ratio
The first half year of 2017	−4.422%	−1.086
The second half year of 2017	−2.082	−0.731
The first half year of 2018	5.022%***	0.259

Note: *, **, *** indicate significance at the 10%, 5%, 1%.

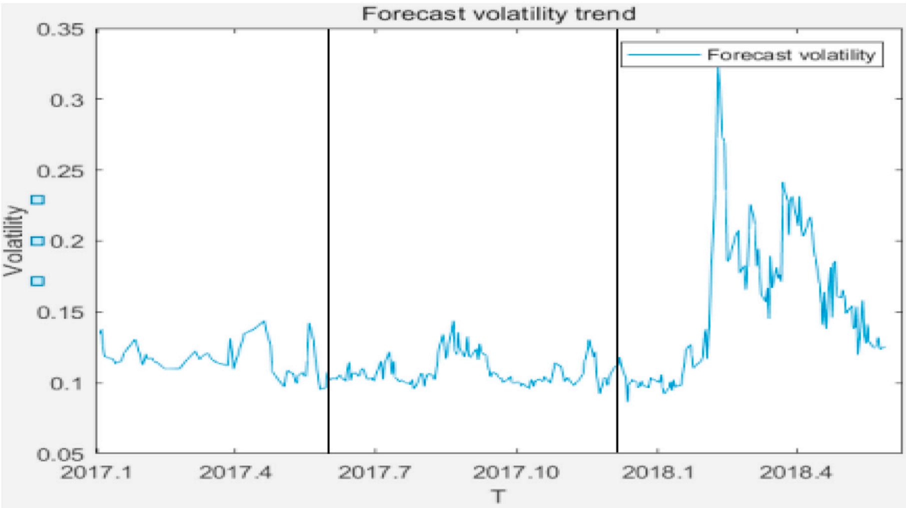

Fig. 5. 2017–2018.5 volatility trend.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|I_i - Q_i| \tag{9}$$

The mean bias error is expressed in the form of MBE, and its formula is given below:

$$MBE = \frac{1}{N}\sum_{i=1}^{N}(I_i - Q_i) \tag{10}$$

Table 3 exhibits the results. In this table, we show mean square error (*MSE*), mean absolute error (*MAE*), mean bias error (*MBE*) of the BP and ABC-BP neural network model and the calculated results of the BS model from January 2017 to May 2018. According to Table 3, the MSE, MAE and MBE of ABC-BP neural network model is the smallest among the three models, hence the overall performance of the ABC-BP neural network model proposed in this paper is superior to that of the BP neural network model and the BS model.

Finally, we bring the volatility data into Monte Carlo simulation to calculate the predicted option price and Greeks (including delta and vega). The predicted results are shown in Table 4.

3.3. Application of predicted option price in three option strategies

In this section, we apply the findings of the predictability of our model to major options trading strategies: straddle, butterfly, and calendar spread to verify the validity of model prediction. In addition, we also apply our model to volatility momentum and mean reversion. The momentum refers to the short-term volatility behavior while the mean-reversion to long-term behavior. We choose three-day period for the volatility momentum and use the benchmark of 10% and 90% of volatility movement for the mean-reversion.

Table 6
The impact of forecast accuracy, maturity and execution price matching on return using straddle strategy.

Panel A: The influence of the accuracy of volatility forecast on the return											
Filter	Volatility Forecast	Effect	ITM			ATM			OTM		
			Annum Return (%)	Sharp	vega	Annum Return (%)	Sharp	vega	Annum Return (%)	Sharp	vega
0–2%	ABC-BP	1	−0.116	−0.711	419.543	−0.852	−5.371	936.970	4.750***	0.370	255.361
		2	−0.128	−1.037	390.801	−0.816	−5.886	922.648	0.083**	0.561	267.825
2–5%	ABC-BP	1	−0.070****	−1.992	92.227	−0.8481	−8.201	592.329	−0.034	−0.389	286.661
		2	−0.081***	−0.459	384.305	−0.813	−6.247	914.902	−0.061	−0.589	264.606
	GARCH	2	−0.209	−2.770	489.562	−0.893	−17.462	887.755	0.028	0.476	189.461
5–10%	ABC-BP	1	−0.217*	−4.002	431.809	−0.8555	−10.633	973.604	0.015**	0.055	130.080
		2	−0.143	−1.557	82.743	−0.820	−3.969	884.684	0.082**	0.764	57.773
	GARCH	1	−0.442	−0.686	161.418	−0.860	−12.452	969.302	−0.0839	−0.994	210.752
		2	−0.135	−1.169	403.818	−0.937	−9.751	194.078	−0.054	−0.618	260.417
Panel B: The effect of the time to maturity on the return											
0-30d	ABC-BP	1	−0.337	−7.793	85.290	−0.933	−21.321	223.099	0.082*	1.411	14.047
		2	−0.201	−3.649	77.904	−0.920	−16.473	208.995	0.097**	1.369	51.015
	GARCH	1	−0.045	−0.549	76.622	−0.925	−10.568	207.341	0.026	0.120	44.684
		2	−0.039	−0.366	53.833	−0.930	−5.820	199.676	0.002	−0.126	21.955
30-60d	ABC-BP	1	−0.215	−2.774	250.907	−0.922	−12.429	382.640	0.017**	−0.024	66.676
		2	0.041*	0.510	116.406	−0.919	−15.994	346.170	0.051**	0.634	79.361
	GARCH	1	−0.025	−0.356	166.943	−0.925	−8.852	363.670	0.001	−0.152	97.343
		2	−0.067	−0.525	187.829	−0.921	−5.942	373.426	−0.053	−0.497	83.439
60-90d	ABC-BP	1	−0.321*	−5.141	316.449	−0.926	−14.091	480.924	0.066**	0.874	112.489
		2	0.034**	0.377	98.374	−0.932	−16.361	459.499	0.073**	1.029	120.475
	GARCH	1	−0.109	−0.900	264.813	−0.922	−7.107	476.368	−0.050	−0.677	165.845
		2	−0.171	−1.096	308.106	−0.915	−5.905	476.339	−0.132	−1.102	180.182
90-120d	ABC-BP	1	−0.136	−1.146	279.842	−0.917	−6.999	550.288	−0.034	−0.440	147.632
		2	−0.172	−1.207	316.049	−0.913	−5.570	557.628	−0.017	−0.276	104.863
	GARCH	1	−0.052	−1.107	209.567	−0.927	−20.789	581.079	−0.164	−4.006	241.430
		2	−0.093	−1.773	241.373	−0.942	−15.355	550.838	−0.052	−1.107	209.567

In this table: In the “effect” column, “1” represents momentum effect, and “2” represents inversion effect. In Panel A with filters, options are only traded when the predicted price deviation is larger than the filter value. The filter value means the difference between the predicted option price and the actual option price. When the value of the filter increases, the number of trades decreases, and the agent is allowed to invest in the risk-free asset on no trading days. This article selects three parts, 0–2%, 2–5%, 5–10% in panel A. And panel B is divided into 1, 2, 3 and 4 months to the maturity. "-" means the data is missing. Missing data of the whole row has been deleted. And *, **, *** indicate significance at the 10%, 5%, 1%.

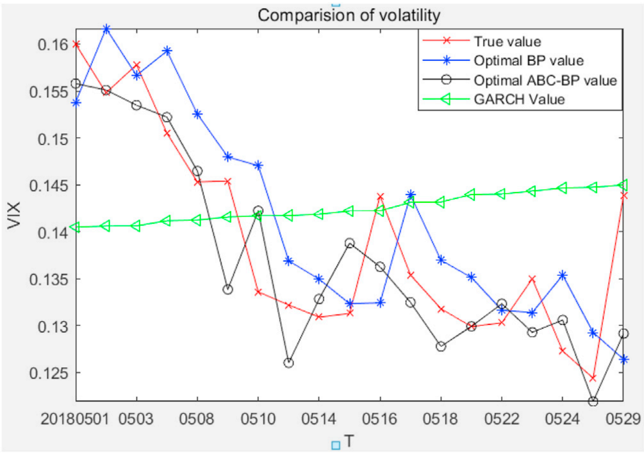

Fig. 6. Comparison of volatility based on different four models (Robustness Test). Note: The BP, ABC-BP neural network model here depicts the best and worst average level.

Table 7
The impact of the distance between the price of the option in the middle and the execution price on return using short butterfly strategy.

Filter	Volatility Forecast	Effect	Strategy 1		Strategy 2	
			Annum Return	Sharp	Annum Return	Sharp
0-182d	ABC-BP	1	0.421*	3.662	0.153**	−117.776
		2	0.480**	2.508	−0.102*	−98.871
	GARCH	1	0.289	4.174	−0.815	−12.870
		2	−0.502	−9.134	−0.401	−7.355
183-364d	ABC-BP	1	0.388*	3.431*	0.666*	12.775
		2	0.419	6.068**	0.427	5.985
	GARCH	1	0.158	1.833	0.115	1.272
		2	0.514	2.198	0.688	2.973
365-546d	ABC-BP	1	0.291	1.672	0.003*	−0.099
		2	0.377*	1.848	–	–
	GARCH	1	0.646	4.311	−0.217	−2.792
		2	0.239	3.193	0.997	16.520

Note: In the “effect” column, “1” represents momentum effect, and “2” represents inversion effect.
Strategy 1 is to sell a call option with lower execution price and higher execution price, and then buy two call options with intermediate execution price; strategy 2 is to sell a put option with lower execution price and higher execution price, and then buy two call options with intermediate execution price. We divide the data in "-" indicates that the data is missing. Missing data of the whole row has been deleted. Where *, **, *** indicate significance at the 10%, 5%, 1%.

We calculate the trading return as $min(S,X) - C_l - C_2$, where S is the forecasted price of underlying asset using Monte Carlo simulation, X is the strike price, C_1 is the price of the call option using predicted price based on Monte Carlo simulation, and C_2 is put option which also uses predicted price based on Monte Carlo simulation. We also compute Sharpe ratio for these strategies. Note that the transaction cost is set at $0.14 per contract.

Table 5 describes the bottom straddle strategy back testing of from January 2017 to May 2018, which is based on the prediction of implied volatility. And the results of one and a half years’ exploration show that the strategy has a good effect in 2018 when the market volatility changes greatly, and a poor performance in 2017 when the volatility is relatively stable.

In Fig. 5, we find the overall volatility mean value in 2017 including the first and second half of the year is not as high as that in 2018’s first half of the year. Therefore, we can infer that the average return of options traded in 2017 is less than that in 2018. It is reasonable because the larger the volatility of the market, the more it will benefit from the higher volatility. And we will continue to explore.

Furthermore, we use a filter to improve the profitability of the strategy. In Table 6, we discuss the following three aspects. First, we classify the data including in-the-value, at-the-value and out-the-value options. In the data classification, we consider the prediction accuracy and the time to maturity. The following data are classified: one is ABC-BP model, the other is GARCH model. Fahlenbrach and Sandås (2010) found evidence that the order flow in volatility strategies with vega exposure contains information about future realized volatility, but there is no evidence that the order flow in directional strategies with delta exposure contains information about future returns. So according to the momentum and reversal effects mentioned in preceding text, the average return rate, sharp ratio, vega value and the significance of the return results are calculated. The calculation results are analyzed in Table 6.

Table 8
The influence of time and volatility on the return of calendar spread trading strategy.

Filter	Volatility Forecast	Effect	Strategy 1			Strategy 2		
			Annum Return (%)	Sharp	vega	Annum Return (%)	Sharp	vega
0-30d	ABC-BP	1	0.000	0.418	56.816	0.279	2.950	48.650
		2	−0.300	−2.188	70.488	0.130	0.833	47.233
	GARCH	1	0.000	−0.199	23.693	–	–	–
		2	0.002	−0.173	45.764	0.259	6.142	31.128
30-60d	ABC-BP	1	0.002*	−0.162	127.907	0.180*	1.660	109.436
		2	−0.297	−2.122	159.083	0.099	0.557	116.754
	GARCH	2	0.001	−0.213	101.164	0.677	16.366	67.346
60-90d	ABC-BP	1	0.003*	−0.136	183.577	0.161	1.182	167.494
		2	−0.216	−1.480	274.586	0.100	0.542	177.857
	GARCH	1	0.002	−0.185	96.235	–	–	–
		2	0.003	−0.170	164.372	0.747	18.247	113.425
The influence of the change range of volatility on the strategic return								
−30~-15%	ABC-BP	1	0.003	−0.137	137.320	0.299**	4.392	104.779
		2	−0.275	−3.074	186.310	0.824***	8.645	135.281
	GARCH	1	0.950	18.916	22.513	0.098	19.638	28.716
		2	0.932	16.437	17.765	0.095	16.903	26.487
−15~-0%	ABC-BP	1	0.002	−0.152	91.272	0.390*	6.285	82.303
		2	–	–	–	0.214**	1.675	34.766
	GARCH	1	0.152	2.024	7.948	0.239	3.746	8.923
		2	0.094	0.960	22.513	0.090	17.948	24.1232
0~+15%	ABC-BP	1	0.002	−0.164	110.092	0.247***	3.355	121.915
		2	–	–	–	0.189***	1.427	75.364
	GARCH	1	0.037	−0.463	0.100	0.000	−1.000	14.265
		2	0.992	6.702	71.770	−0.097	−8.597	−1.025
+15–30%	ABC-BP	1	0.002	−0.167	142.171	0.252***	4.298	144.056
		2	−0.243	−3.788	249.808	0.357	4.742	151.759
	GARCH	1	0.992	6.702	71.770	−0.098	−8.599	−1.041

Note: In the "effect" column, "1" represents momentum effect, and "2" represents inversion effect. Strategy 1 is the calendar spread strategy of reverse call option and strategy 2 is the calendar spread strategy of reverse put option. So the buy calendar spread is equal to the long implied volatility, while the sell is the opposite. In Table 8, the filter of panel a is the distance between near and far months, mainly divided into 0–30 days, 30–60 days and 60–90 days. "-" indicates that the data is missing. Missing data of the whole row has been deleted. Where *, **, *** indicate significance at the 10%, 5%, 1%.

For the two models, we consider three conditions in panel A, 0–2%, 2–5%, and 5–10%. We find that the rate of return of out-the-money options is higher than that of at-the-money options and in-the-money on the whole, whatever the model is and the performance of at-the-money options are worst. According to these three conditions, we find that the strategic return increases and prediction accuracy improves, particularly in out-the-money category. In combination with the momentum effect and reversal effect of volatility, we find that they have the same results and the results are in line with the bottom straddle strategy's reality. In addition, vega value and the strategy return change reversely. The larger the return, the smaller the vega value. Moreover, we find that the return of ABC-BP neural network model is generally higher than that of GARCH model.

To examine the impact of time-to-maturity, we separate far month's volatility from near month's volatility, and recalculate the return in panel B of Table 6. We select the months of 1, 2, 3 and 4 respectively. We find that the return decreases with time-to-maturity, especially out -the-money option. The results also indicate that the return decreases because the value of volatility prediction error is getting bigger, especially for GARCH model (see Fig. 6), but not for ABC-BP neural network model. And the performance of out-the-money options are still the best. These findings suggest that ABC-BP neural network model performs better than GARCH model.

Table 7 presents the test results of butterfly strategies. As we all know, we can buy one call option, sell two call (typically at-the-money) options and buy another call option to build a butterfly strategy. And their exercise prices are X_1, X_2, and X_3. The premiums of these three options are C_1, C_2, and C_3 which are predicted based on Monte Carlo simulation. So, the return of butterfly strategies can be divided into four intervals. We mainly calculate the trading return as $min(S, X_1(X_3)) - C_1 - C_2$, where S is the forecasted price of underlying asset using Monte Carlo simulation.

We find that the closer to the maturity, the higher the return under both reversal and momentum effect. In this table, Strategy 1 is shorting a call butterfly and Strategy 2 is shorting a put butterfly. We find that the call butterfly's performance surpasses that of put butterfly, indicating ABC-BP neural network model works better for call options.

We examine the time-to-maturity impact by dividing the options into three maturity categories: 0–182 days, 183–364 days and 365–546 days. The results show that the performance decreases as getting closer to the maturity, suggesting that the model works better for short-term butterfly.

In order to consider the impact of volatility changes, we adopt shorting calendar spread strategy. We construct an options spread by simultaneously entering a long and a short option positions with the same strike price but different maturity. Our quantitative signal model is $X = 2V_1 - V_2 - V_3$, where V_1, V_2, and V_3 are the implied volatility for short-term, mid-term, and long-term maturities, respectively. When $X < 0$, the calendar spread portfolio is constructed by a long call option for front month and a short call option for back month.

Table 9
Comparison of volatility errors in GARCH and ABC-BP neural network model.

	MSE	MAE	MBE
GARCH	0.264	0.510	0.022
BS	0.001**	0.024*	0.001**
ABC-BP	0.000***	0.010***	0.000***

Note: Where *, **, *** indicate significance at the 10%, 5%, 1%.

For credit calendar spread strategy in panel A of Table 8, we divide the time to maturity into 0–30 days, 30–60 days, and 60–90 days of maturity, respectively. We find the returns gradually increase in maturity. This finding suggests that the ABC-BP neural network model works well in this strategy (see Table 9).

In Table 8, Panel B, we find the return range of option strategy is also related to the change of volatility. As volatility rises, the option return increases, and vice versa. When we classify our sample into four parts according to the change of volatility, 30-15% drop, 0–15% drop, 0–15% rise, and 15–30% rise, respectively. Again, we find ABC-BP neural network model is more effective. With these impacts from maturity and volatility, we suggest that momentum effect is stronger than reversal effect.

4. Conclusion

In this paper, we examine the artificial bee colony improved model (ABC-BP neural network model) in IV predictability and apply the model to three popular option trading strategies. We document two major findings.

First, the inherent convergence speed of BP neural network model is slow. This model is easy to fall into local optimum and easy to be overfit. It also prolongs the training time while has lower accuracy of implied volatility. Our experimental results show that ABC-BP neural network model performs better than the BP neural network model in terms of speed and predictability, and also performs better than the traditional GARCH model. Secondly, we document that ABC-BP neural network model is applicable to option trading strategies such as straddle, butterfly, and calendar spreads. The performance of the model is better than traditional GARCH models. And we believe that this conclusion is helpful for option traders to select trading strategies and specific trading products.

Acknowledgements

This study was supported by the National Social Science Fund of China (19BJY242). The authors would like to thank Carl Chen (the Editor) and two anonymous Referees for their insightful comments.

Appendix A

The computational process of ABC-BP neural network model.

1) Create a BP neural network model.
2) Initialize the parameters of the ABC algorithm, comprising the bee colony's size (N_c), the number of employed bees (N_e), the number of onlookers (N_o), the number of solutions (N_s), the limit (*limit*), the maximum number of cycles (*MCN*), and the initial solution of the D-dimension X_i ($i = 1, \ldots, N_s$), and N_c, N_e, N_o and N_s satisfy the following relation:

$$N_c = 2N_s = N_e + N_o, \ N_e + N_o \tag{A1}$$

The D-dimensional vector X_i ($i = 1, \ldots, N_s$) represents the connection weights and thresholds of the network created in (1), and the dimension D of each solution satisfies the following equation:

$$D = N_{input} \times N_{hidden} + N_{hidden} + N_{hidden} \times N_{output} + N_{output} \tag{A2}$$

N_{input}, N_{hidden} and N_{output} are the number of neurons in the input layer, the hidden layer, and the output layer respectively. The value of the initial solution is a randomly generated number between (−1, 1).

3) Calculate the fitness of each solution according to equation (A1).

$$f(X_i) = \begin{cases} 1 & MSE_i = 0 \\ \dfrac{1}{MSE_i} & MSE_i > 0 \end{cases} \tag{A3}$$

In this equation, $i = 1, \ldots, N_s$. MSE_i represents the BP neural network model mean squared error of the i^{th} solution. When the fitness reaches 1, this would be the most ideal situation clearly.

4) The employed bees search for a new solution based on the current solution memoried before.

$$V_{ij} = X_{ij} + rand(-1,1)\left(X_{ij} - X_{kj}\right) \quad \text{(A4)}$$

i is the number of the solution, $j \in \{1,\ldots\ldots,D\}$, $k \in \{1,\ldots\ldots,N_s\}$ are randomly generated, and $k \neq i$. The employed bees adopt the greedy mechanism. If the fitness of the new solution is larger than the fitness of the old solution, write down the new solution, otherwise the number of update failures of the old solution is increased by one.*5)* Calculate the possible values (P_i) of each solution.

$$P_i = \frac{f(X_i)}{\sum_{n=1}^{N_s} f(X_n)} \quad \text{(A5)}$$

$f(X_i)$ is the fitness of the i^{th} solution. And $\sum_{n=1}^{N_s} f(X_n)$ is the sum of the fitness of solutions. The onlooker searches for a new solution (Formula 4) from the neighborhood of the existing solution based on these possible values.

6) If the failures number of the solution X_i updates exceeds the preset limit value (*limit*), it means that the solution can no longer be optimized, and it must be discarded, and replaced by a new solution generated by the following formula:

$$X_i = X_{\min} + rand(0,1)(X_{\max} - X_{\min}) \quad \text{(A6)}$$

X_{min} is the smallest solution, and X_{max} is the largest one. Save the optimal solution.

7) If the number of iterations is greater than the maximum number of cycles (*MCN*), the training ends. Otherwise, return to step (4).
8) The obtained optimal solution is transformed into the connection weights and thresholds of the BP neural network model, and the neural network is simulated and tested with data.

References

Andreou, P. C., Charalambous, C., & Martzoukos, S. H. (2006). Robust artificial neural networks for pricing of European options. *Computational Economics, 27*(2–3), 329–351.

Bollen, N. P. B., & Whaley, R. E. (2004). Does net buying pressure affect the shape of implied volatility functions? *The Journal of Finance, 59*(2), 711–753.

Engle, R. F., & Sokalska, M. E. (2012). Forecasting intraday volatility in the us equity market. multiplicative component GARCH. *Journal of Financial Econometrics, 10*(1), 54–83.

Fahlenbrach, R., & Sandås, P. (2010). Does information drive trading in option strategies? *Journal of Banking & Finance, 34*(10), 0-2385.

Fang, L., Chen, B., Yu, H., & Qian, Y. (2017). The importance of global economic policy uncertainty in predicting gold futures market volatility: A GARCH-MIDAS approach. *Journal of Futures Markets, 38*(3), 413–422.

Guo, D. (2000). Dynamic volatility trading strategies in the currency option market. *Review of Derivatives Research, 4*(2), 133–154.

Giot, & Pierre. (2009). Implied volatility indexes and daily value at risk models. *Journal of Derivatives, 12*(4), 54–64.

Harvey, C. R., & Whaley, R. E. (1992). Market volatility prediction and the efficiency of the S&P 100 index option market. *Journal of Financial Economics, 31*(1), 43–73.

Huang, Y. Z., & Zheng, Z. L. (2009). Model-free implied volatility and its information content: Evidence from Hang Seng index options. *Systems Engineering-Theory & Practice, 29*(11), 46–59.

Jiang, G. J. (2002). Testing option pricing models with stochastic volatility, random jumps and stochastic interest rates. *International Review of Finance, 3*(3–4), 233–272.

Jiang, G. J., & Tian, Y. S. (2005). The model-free implied volatility and its information content. *Review of Financial Studies, 18*(4), 1305–1342.

Koopman, S. J., Jungbacker, B., & Hol, E. (2005). Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements. *Journal of Empirical Finance, 12*(3), 445–475.

Li, J., Cheng, J. H., Shi, J. Y., & Huang, F. (2012). Brief introduction of back propagation (BP) neural network algorithm and its improvement. *Advances in Computer Science and Information Engineering*. Springer Berlin Heidelberg.

Madeleine, B., & Bin, L. J. (2008). Foraging in honeybees—when does it pay to dance? *Behavioral Ecology, 19*(2), 255–261.

Markose, S. M., Peng, Y., & Alentorn, A. (2012). Forecasting extreme volatility of ftse-100 with model free VFTSE, Carr-Wu and generalized extreme value (GEV) option implied volatility indices. *Economics Discussion Papers, 34*(4), 871–881.

Muzzioli, S. (2010). Option-based forecasts of volatility: An empirical study in the dax-index options market. *The European Journal of Finance, 16*(6), 561–586.

Pilbeam, K., & Langeland, K. N. (2015). Forecasting exchange rate volatility: garch models versus implied volatility forecasts. *International Economics & Economic Policy, 12*(1), 127–142.

Ser-Huang, P., & Clive, G. (2005). Practical issues in forecasting volatility. *Financial Analysts Journal, 61*(1), 45–56.

Sermpinis, G., Laws, J., & Dunis, C. L. (2013). Modelling and trading the realised volatility of the FTSE100 futures with higher order neural networks. *The European Journal of Finance, 19*(3), 165–179.

Sophie, X. N. I., Pan, J., & Poteshman, A. M. (2008). Volatility information trading in the option market. *The Journal of Finance, 63*(3), 1059–1091.

Xu, L., & Yin, X. (2017). Exchange traded funds and stock market volatility. *International Review of Finance, 17*(4), 525–560.

Zhang, Y., Yao, T., He, L., & Ripple, R. (2018). Volatility forecasting of crude oil market: Can the regime switching GARCH model beat the single-regime GARCH models? *International Review of Economics & Finance, 59*, 302–317.