Physica A 482 (2017) 444–457

Contents lists available at ScienceDirect

Physica A

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

Quantified moving average strategy of crude oil futures market based on fuzzy logic rules and genetic algorithms

Xiaojia Liu [a,b], Haizhong An [a,b,c,*], Lijun Wang [d], Qing Guan [a,b]

[a] School of Humanities and Economic Management, China University of Geosciences, Beijing 100083, China
[b] Key Laboratory of Carrying Capacity Assessment for Resource and Environment, Ministry of Land and Resources, Beijing 100083, China
[c] Key Laboratory of Strategic Studies, Ministry of Land and Resources, Beijing 100812, China
[d] Zhongyuan Bank, Zhengzhou 450018, China

HIGHLIGHTS

- We proposed a trading rule called Fuzzy Moving Average Strategy (FMAS).
- Fuzzy logic rule and Moving Average Strategy (MAS) are combined for trading.
- Genetic algorithms are improved to a 3D version for better parameter optimization.
- The FMAS generate more stable rate of return than other three trading strategies.
- The holding amount is highly sensitive to price series.

ARTICLE INFO

Article history:
Received 16 September 2016
Received in revised form 13 March 2017
Available online 29 April 2017

Keywords:
Moving average strategy
Fuzzy logic rules
Genetic algorithms
Technical analysis
Trading rule

ABSTRACT

The moving average strategy is a technical indicator that can generate trading signals to assist investment. While the trading signals tell the traders timing to buy or sell, the moving average cannot tell the trading volume, which is a crucial factor for investment. This paper proposes a fuzzy moving average strategy, in which the fuzzy logic rule is used to determine the strength of trading signals, i.e., the trading volume. To compose one fuzzy logic rule, we use four types of moving averages, the length of the moving average period, the fuzzy extent, and the recommend value. Ten fuzzy logic rules form a fuzzy set, which generates a rating level that decides the trading volume. In this process, we apply genetic algorithms to identify an optimal fuzzy logic rule set and utilize crude oil futures prices from the New York Mercantile Exchange (NYMEX) as the experiment data. Each experiment is repeated for 20 times. The results show that firstly the fuzzy moving average strategy can obtain a more stable rate of return than the moving average strategies. Secondly, holding amounts series is highly sensitive to price series. Thirdly, simple moving average methods are more efficient. Lastly, the fuzzy extents of extremely low, high, and very high are more popular. These results are helpful in investment decisions.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

The Moving Average Strategy (MAS) is one of the most popular technical indicators [1–4]. This strategy informs traders of the existing trend and helps identify the upcoming reversal trend. The MAS generates a buy signal or a sell signal by judging

* Corresponding author at: School of Humanities and Economic Management, China University of Geosciences, Beijing 100083, China.
E-mail addresses: jazzyliu@live.cn (X. Liu), ahz369@163.com (H. An), swarm@live.cn (L. Wang), guanqing35@126.com (Q. Guan).

http://dx.doi.org/10.1016/j.physa.2017.04.082
0378-4371/© 2017 Elsevier B.V. All rights reserved.

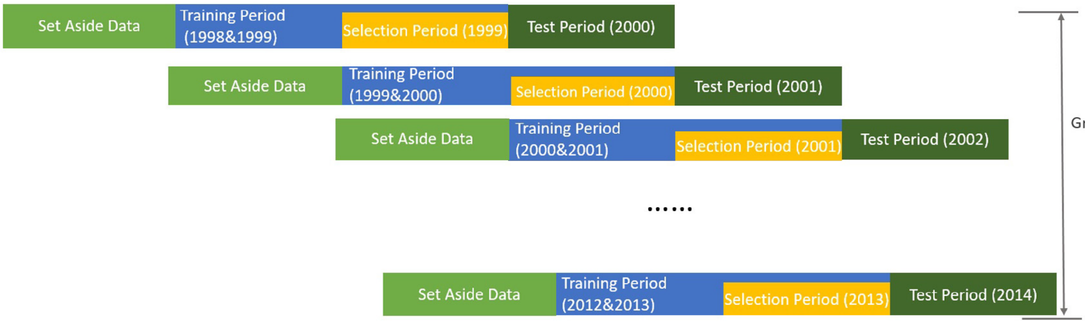


Fig. 1. Experimental groups.

whether the difference between a long-term average and a short-term average is positive or negative. Glabadanidis (2016) combined simple moving average with several window lengths to improve its predictability on portfolio [5]. Liu (2017) used Particle Swarm Optimization to investigate the distribution of the weights of different moving averages [6]. However, the previous MAS used unit 1 as the trading volume and did not consider the change of trading volume [7–9]. Using of unit 1 is a good way to simplify the research question, but it ignores such questions as the cost of the handling charge and the trading volume requirement for one transaction. In the financial market, one transaction usually requires a trading volume larger than one unit. Therefore, it is reasonable to clear the trading volume and it helps to calculate the rate of return accurately. The investigation of the change of trading volume thus is important and fulfill the research gap in the literature.

Choosing appropriate tool to investigate the trading volume issue is critical. The fuzzy logic rule is an imitation of human decision mechanism under uncertainty circumstances, which breaks the limitation of binary logic and shows more accuracy for complex situation [10]. It is an effective method to find trading rule in the financial market [11–14]. As for its applications in previous moving average research in financial field, Ghandar (2009) used the fuzzy logic rule to integrate the moving average and other investment strategies to form an adaptive computational intelligence system for learning trading rules [15]. Gradojevic (2013) took the moving average as the input of fuzzy logic rules to design an uncertainty reduction approach [16]. Because of its effectiveness, we use the fuzzy logic rule to design the strategy for trading volume decision.

In this paper, we use the moving average strategy and the fuzzy logic rule to compose the trading strategy. We apply the moving average strategy to generate the trading signal, and use the fuzzy logic rule to describe the strength of the moving average signals, followed by the integration of the strength, as trading volume, into the MAS to form a quantified moving average strategy. Throughout this paper, we call this integrated strategy as the Fuzzy Moving Average Strategy (FMAS). Through this way, the FMAS can provide both trading signals and trading volumes. To search for the optimized trading strategy, we apply genetic algorithms and utilize crude oil futures prices as the test data because of the variety of trends.

The main contribution of this paper to the literature is that we propose a quantified moving average strategy for the crude oil futures market. This strategy not only considers the trading signals but more importantly the trading volume. Furthermore, we use genetic algorithms to present the optimized trading strategy. The results from this study could provide additional aid to those who want a quantified trading strategy and those who use moving average strategy in investment decisions.

The rest of this paper is outlined as follows: Section 2 describes the data and methods that we use here to integrated the FMAS in the crude oil futures market. The experimental results in Section 3 present the performance and the profitability of the FMAS. We discuss the experimental results and parameters selection in Section 4. Section 5 concludes the paper with suggestions for traders in the crude oil futures market.

2. Data and methods

2.1. Data

The data used in this paper are crude oil futures daily prices from the New York Mercantile Exchange (NYMEX). We downloaded it from the Energy Information Administration (EIA) website (http://www.eia.gov/dnav/pet/pet_pri_fut_s1_d.htm) on May 7th, 2015. We choose contract 1 prices from the EIA crude oil futures data, which is the most active contract from 12 crude oil futures contracts. The test period in this paper is from 2000 to 2014.

The FMAS needs previous data for the training period and the selection period. The fuzzy logic rule requires background data. Therefore, we choose data from Jan 09, 1995 to Dec 12, 2014, which is 5000 trading days, as the experimental data. The experimental data are divided into 15 groups as shown in Fig. 1. There are 750 trading days being set aside. The training period contains 500 trading days. The selection period and the test period both contain 250 trading days, which is approximately one year.

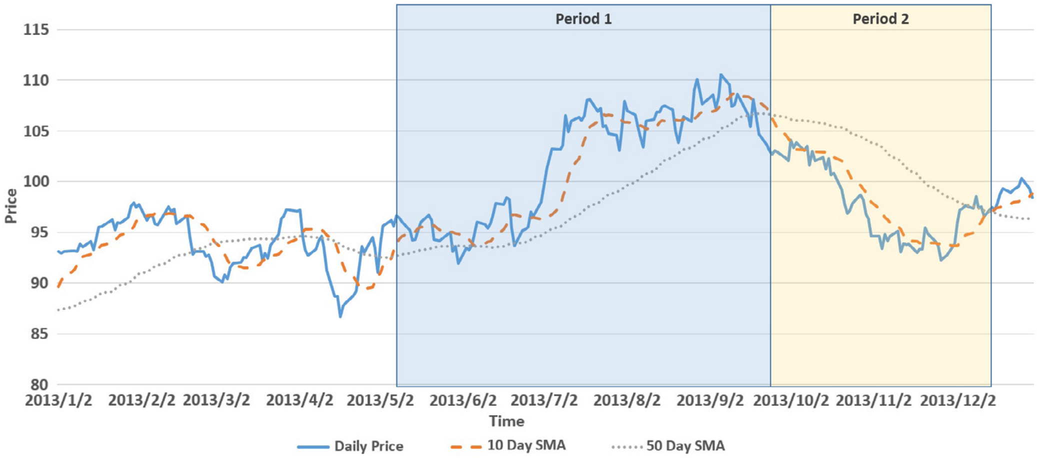


Fig. 2. Schematic diagram of moving average strategy.

2.2. Moving average strategy

The principle of the moving average strategy is to use the relative position of a long-term moving average line and a short-term moving average line to estimate the market trend. When the short-term moving average line is above the long-term moving average line, the strategy believes that the market price is increasing. Conversely, when the long-term moving average line surpasses the short-term line, the market price is believed to be decreasing. Fig. 2 shows the two situations. In period 1, the 10-day line is above the 50-day line. Thus, in this period, the moving average strategy believes the market price is increasing. Conversely, the market price in period 2 is decreasing.

There are six types of moving averages: Simple Moving Average (SMA), Weighted Moving Average (WMA), Exponential Moving Average (EMA), Adaptive Moving Average (AMA), Typical Price Moving Average (TPMA), and Triangular Moving Average (TMA). These moving averages have different calculation methods and have different performances. In this paper, we choose four of them based on the previous study [17,18]. They are SMA, AMA, TPMA, and TMA.

SMA calculates the mean prices of a period. AMA calculates based on weights, and the weights change along with the prices trend. TPMA solely calculates the mean of the maximum price, minimum price and the closest price in a period. TMA is the mean value of SMAs in a period. The calculation methods are shown from formula (1) to formula (4).

$$SMA(k) = \frac{1}{n}\sum_{i=1}^{n} c_{k-i}. \tag{1}$$

Variable n indicates the length of period. Variable c indicates the daily price of crude oil futures market, and k is the day we need to calculate.

$$\begin{aligned} &AMA(k) = AMA(k-1) + SSC_k^2(c_{k-1} - AMA(k-1)) \\ &SSC_k = ER_k(fastSC - slowSC) + slowSC \\ &fastSC = 2/(1+2), \qquad slowSC = 2/(1+30) \\ &ER_k = |c_{k-1} - c_{k-n}| \Big/ \sum_{i=k-n}^{k-1} |c_i - c_{i-1}| \,. \end{aligned} \tag{2}$$

SSC means Structured Smoothing Constant, which is determined by the efficiency ratio (ER), fast smoothing constant ($fastSC$) and slow smoothing constant ($slowSC$). Constant $fastSC$ indicates the smoothing constant when the moving average period is the smallest, which is 2 days. The $slowSC$ indicates the smoothing constant when the moving average period is a month, which is approximately 30 days. ER is the signal to the noise. The signal indicates the absolute value of the differences between the current close price and the close price n days before. Noise means the summation of absolute differences of n cycles between the current close price and the close price one day before.

$$\begin{aligned} &TPMA(k) = (high + low + close)/3 \\ &high = \max(c_{k-1}, c_{k-2}, \ldots, c_{k-n}), \qquad low = \min(c_{k-1}, c_{k-2}, \ldots, c_{k-n}), \qquad close = c_{k-1}. \end{aligned} \tag{3}$$

TPMA calculates the maximum value, the minimum value and the closest value. Variable $high$ indicates the maximum price. Variable low indicates the minimum price and variable $close$ indicates the price before the trading day.

$$TMA(k) = \frac{1}{n}\sum_{i}^{n} SMA(k-i). \tag{4}$$

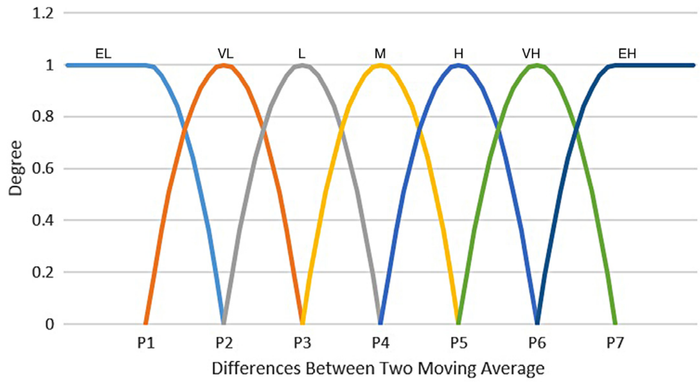


Fig. 3. Membership function.

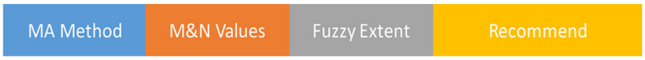


Fig. 4. Structure of fuzzy logic rule.

TMA is the SMAs averaged again. We need to calculate the SMA of every day in the period. Then, TMA calculates the mean values of the SMAs.

The moving average strategy calculates the moving averages of two periods, a short-term period and a long-term period. In this paper, we use n and m to represent the two periods. If the moving average of n (the short period) exceeds the moving average of m (the long period), the price series can be regarded as having a rising trend and vice versa. In this paper, we select 6 n values and 6 m values, as shown in formula (5). Therefore, there are 32 different combinations of m and n since value m must be greater than value n.

$$\begin{aligned} &m \in \{10, 20, 50, 100, 150, 200\}, \\ &n \in \{1, 3, 5, 10, 15, 20\}. \end{aligned} \tag{5}$$

2.3. Fuzzy logic rule

The moving average provides the buy or sell signals. If we only use the moving average strategy, we must make decisions immediately when the two moving average lines cross over. This method has large risks and may cause extra costs and losses. Some researchers set a range parameter to make decisions until the differences of the two moving averages exceed a certain range. By using the fuzzy logical rule, we cannot only set a range parameter, but also make the parameter change along with different situations. The fuzzy logic rule is essential to precisely describe and process the fuzziness object. It extends the ordinary set, which take two values (0 and 1). The fuzzy set can take infinite values of [0, 1]. In addition, the concept "membership" is used to accurately portray the relation between elements and the fuzzy set.

The difference between two moving averages is divided to 7 extents: Extremely Low (EL), Very Low (VL), Low (L), Middle (M), High (H), Very High (VH), and Extremely High (EH). The extent division is based on the previous data. For example, if the experimental data are in 2004, we need to calculate the moving average differences in 2003 and sort the entire year's differences in order from low to high. The lowest number is P1, and the highest number is P7, as shown in Fig. 3. To distinguish the upward and downward trend, we set the P4 value nearest to 0. The values less than 0 are divided into three parts of the same amount. The values greater than 0 are also treated in the same manner. Thus, the P1, P2 and P3 are less than or all equal to zero. In contrast, the P5, P6 and P7 are more than or all equal to zero. With these points, we can obtain our membership function, which is shown in Fig. 3. In this paper, we choose a quadratic function because it performs better than a linear function, which can improve the recommended level. Thereafter, we shall calculate the moving average differences of every trading day in 2004, and decide to which extent and on which range it falls.

One fuzzy logic rule contains an *if* part and a *then* part. Here we set the moving average method, the length of two periods, the fuzzy extent as the *if* part, and the recommend value as *then* part, as shown in Fig. 4. For example, one rule is *if the AMA (15, 50) is L, then recommend value is 0.5*. This rule means, if the difference between 15 days AMA and 50 days AMA falls to the Low interval, we shall calculate a degree value according to the membership function shown in Fig. 3. Multiply the degree value by the recommendation value. The result is the rating level that this fuzzy logic rule gives. It is worthy of attention that we set the recommend values of the first three fuzzy extents as positive numbers and the recommend values of the last three fuzzy extents as negative numbers to obtain better performance. Appendix provides two examples that are more specific.

Because sometimes the difference between the two moving averages does not fall on the fuzzy extent according to the membership function, we use a rule set to decide an overall rating level in this paper. Each set contains 10 fuzzy logic rules as shown in Fig. 5. Each rule provides its rating degree or not. Then, we use the average number as the overall rating level.

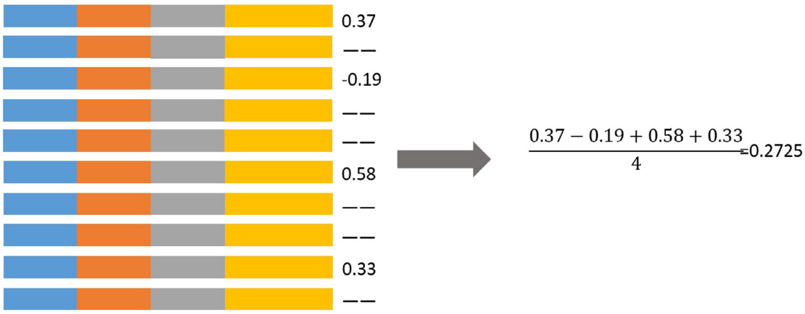


Fig. 5. Fuzzy rule set.

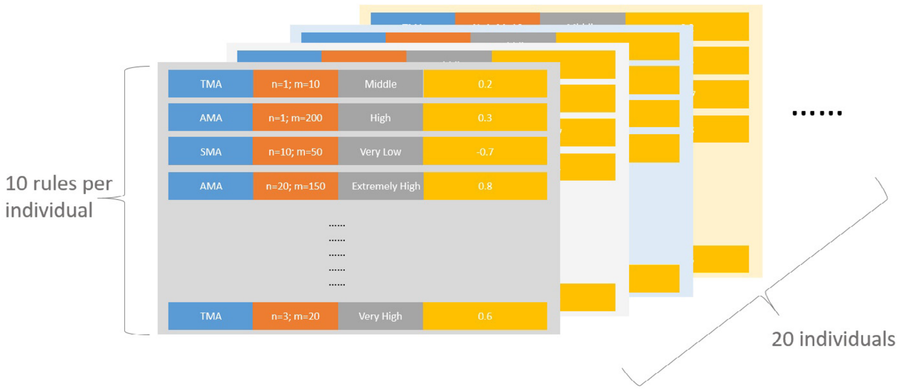


Fig. 6. Population in genetic algorithms.

2.4. Genetic algorithms

Based on the above, the moving average method has four options. The values of n and m have 32 options. The fuzzy extent has seven options and the recommend value has 21 options (range from −1 to 1, and the interval is 0.1). Since the first three fuzzy extents solely cause negative recommend values, and the last three solely cause positive values, we will finally have 11 136 types of rules, as shown in formula (6). We choose 10 of these as a set. There are approximately $8.05 * 10^{33}$ sets as shown in formula (7).

$$RN = 4 \times 32 \times 3 \times 11 + 4 \times 32 \times 1 \times 21 + 4 \times 32 \times 3 \times 11 = 11136, \tag{6}$$

$$SN = C_{RN}^{10} \approx 8.05 \times 10^{33}. \tag{7}$$

It is impossible to try to use every set to find out which one is the best set. In the search for the optimal fuzzy rule set, genetic algorithms method is used. Genetic algorithms is one method used to simulate biological evolution and natural selection [19–23]. Genetic algorithms can search the optimal solutions by using the processes of selection, crossover and mutation. Compared with other optimization algorithms such as neural network or partial swarm optimization, genetic algorithms method has a transparent mechanism and a lower possibility of local convergence [18]. Since we need the estimates of all parameters in the intermediate process, the genetic algorithms method is more suitable and adopted for this paper. In the genetic algorithms, the fuzzy sets are called individuals. The collection of individuals is called a population. In this paper, every population contains 20 individuals, as shown in Fig. 6.

After one experiment of one generation, the population needs to evolve. We take the rate of return as the fitness value and use Roulette Wheel Selection (RWS) to choose 10% of the 20 individuals from the old population. They are most likely the two best individuals. We retain the two individuals as the first part of the new population. The second part of the new population is two individuals that are randomly generated. Then we continue use RWS to select 80% of the 20 individuals from the old population which is mostly like the 16 best individuals. These 16 individuals need to go through crossover and mutation to become the third part of the new population.

The crossover here indicates the exchange rules between two individuals. The crossover probability is 0.7, which is the most commonly used. We choose two individuals with a 0.7 probability and randomly generate a spot in the individuals. The rules after the spot in the two individuals exchange from each other. In this way, two new individuals are generated, as shown in Fig. 7(a).

The mutation means a change within one rule and the exchange between two rules. The latter is also a type of crossover, but it occurs within an individual. Therefore, we would call it mutation in this paper. For the former, the mutation probability is 0.01, which means a rule choice with a 0.01 probability and a spot randomly generated in the rule. Next, the binary string on the point changes from 0 to 1, or 1 to 0, as shown in Fig. 7(b). The exchange between the two rules is similar to the crossover process. The algorithm generates a spot and the part after the spot exchange, as shown in Fig. 7(c).

2.5. Rate of return calculation

Each individual (10 rules) provides a final overall rating level, which is a percentage to the largest possible holding. The largest possible holding is based on the initial capital and the current price as shown in formula (8). The final earning contains the profit that is gained by buying and selling crude oil futures contracts, the handling charge, and the risk free return of the remaining capital, as shown in formula (12). The profit is the money earned minus the money spent. When the holding today should be more than yesterday, the trader needs to spend money to buy a contract. If the opposite occurs, the trader sells a contract and earns money, as shown in formula (9). The crude oil futures trades are charged certain with handling fees per thousand contracts. These fees are the costs in the trade, and the calculation method is shown in formula (10). The risk free return is the growth of the money, which is not in the market. The calculation method is shown in formula (11). Finally, the whole rate of return is the calculation. The rate of return function is the objective function in the Genetic Algorithm in this paper.

$$holding_i = capital/(price_i \times deposit) \times rlevel_i. \tag{8}$$

Variable *holding* indicates the amount that the trader holds. Because the trading volume in the crude oil futures market is counted by thousand, the digits below kilobit all are 0 and the numbers less than one thousand is also directly treated as 0. Constant *capital* means the initial capital that the trader has in the beginning. In this paper, the initial capital is 1 million dollars. Variable *price* indicates the crude oil futures price. $Pirce_i$ means the price when the trader buys futures oil contract. Constant *deposit* is the rate that a trader needs to pay for buying contracts. Usually in the crude oil futures market, the traders do not need to pay all the money denoted in the contract in advance. They usually use a deposit instead. In this paper, the deposit rate is 0.15, which is the average rate for the crude oil futures market. Variable *rlevel* indicates the rating level each individual provides.

$$profit = \sum_{i=2}(price_i \times (holding_{i-1} - holding_i)) - price_1 \times holding_1. \tag{9}$$

Variable *profit* indicates the money earned when trading crude oil futures contracts.

$$cost = \left(\sum_{i=2}(holding_i - holding_{i-1}) + holding_1\right)/1000 \times charge. \tag{10}$$

Variable *cost* indicates the cost of the handling charge. Constant *charge* is the cost when trading a thousand of contracts. Usually, this cost is between 15 and 20. In this paper, we set the *charge* to 18 dollars.

$$riskfree = \sum(capital - |holding_i| \times price_i \times deposit) \times rfrate. \tag{11}$$

Variable *riskfree* means risk free return. Constant *capital* means the initial capital the trader has in the beginning. Variable *rfrate* indicates the risk free rate. In this paper, it is 0.02, which is the return of national debt.

$$rreturn = (profit + riskfree - cost)/capital. \tag{12}$$

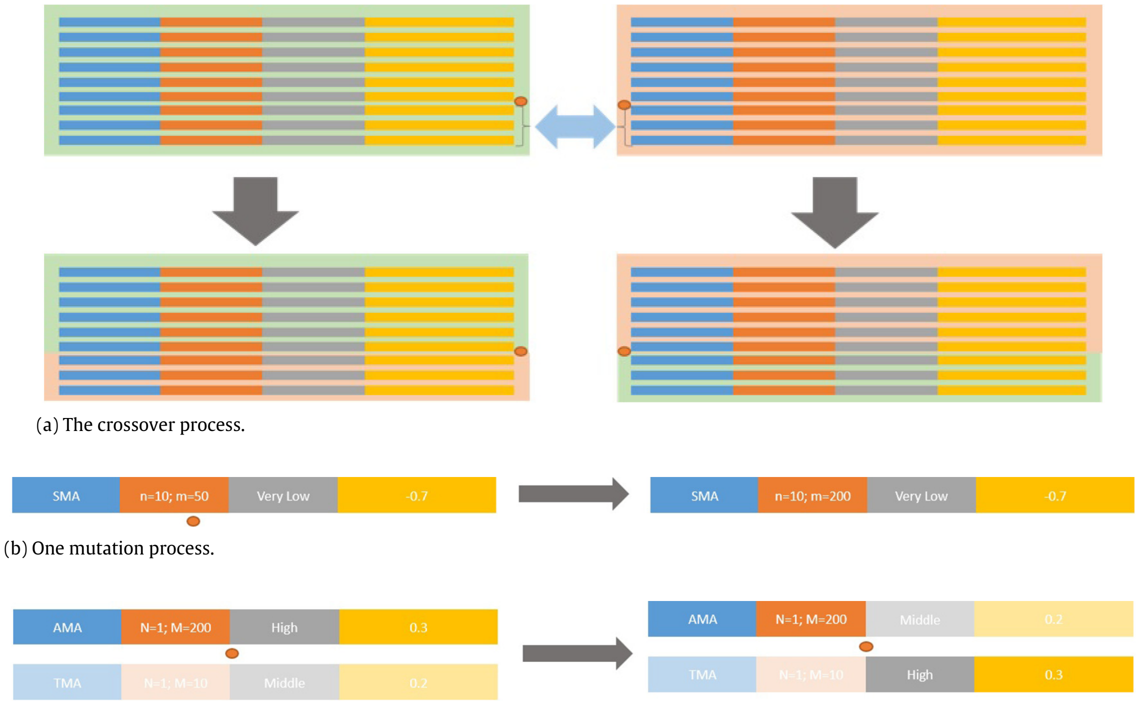
Variable *rreturn* means rate of return. Variable *profit* indicates the money earned from the trade of crude oil futures contracts. Variable *riskfree* means risk free return. Variable *cost* indicates the cost of handling charge. Constant *capital* means the initial capital the trader has in beginning.

2.6. Experimental process

The framework of the methods is shown in Fig. 8. The moving average strategy provides the signal and difference between two moving averages. The fuzzy logic rule uses the difference between the two moving averages to decide the rating level. The rating level and the initial capital joint provide the trading volumes. The signals and trading volumes together form the FMAS. During this process, there are certain variables that need to be optimized, including the moving average method, the length of two periods, the fuzzy extent, and the recommendation procedure. We use genetic algorithms to address these.

The detailed process for the experiment is as follows:

(1) Generate the initial population. Randomly generate 20 individuals. Each individual contains 10 rules in accordance with the rules set.
(2) Calculate the fuzzy moving averages in the training period. Calculate the differences between two moving averages in the training period and the previous training period using the moving average method and the length of two moving averages that the rule provides. Next, obtain the membership function according to the moving average differences of the previous training period. Calculate the rating level according to the moving average differences of this training



(a) The crossover process.

(b) One mutation process.

(c) The other mutation process.

Fig. 7. Evolution process of genetic algorithms.

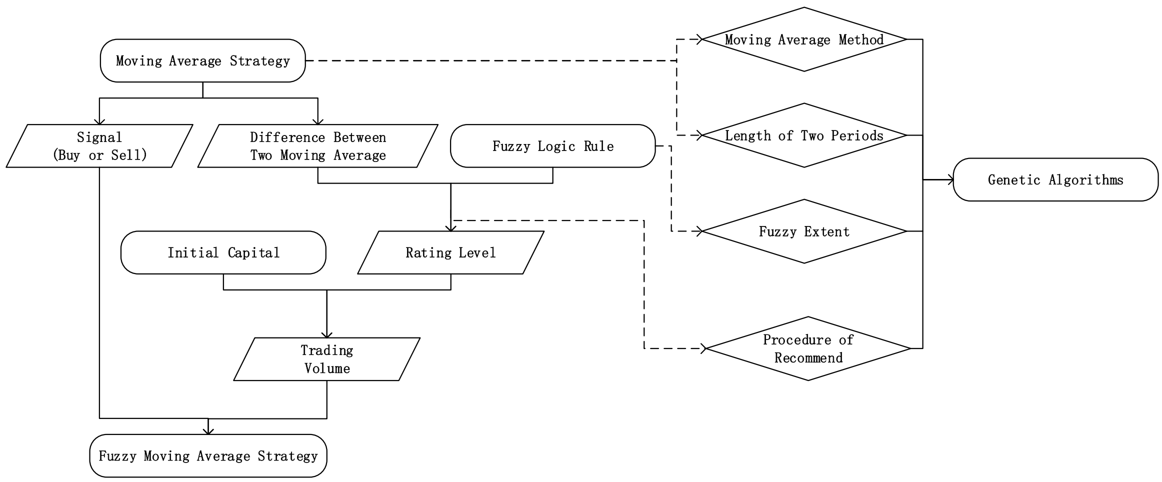


Fig. 8. Structure of the methods.

period, the membership function, and the recommend values that the individual provides. Because every individual provides a set of recommend values, there are ultimately 20 sets of rating levels.

(3) Select the best individual in the training period. Calculate the trading volume using the initial capital and the rating degree. Then, calculate the rate of return according to the rate of return calculation method provided above. Every individual has a rate of return. The one that has the highest rate of return is the best individual in the training period.
(4) Calculate the rate of return in the selection period using the best individual in the training period. Calculate the differences between the two moving averages of this selection period and the previous selection period. Then, obtain the membership function and calculate the rating level and the rate of return.
(5) Mark the best individual. In the first cycle, mark the rate of return as the best rate of return and the best individual in the training period as the overall best individual. From the second cycle, compare the rate of return in this loop and the best rate of return. If the rate of return from this loop is higher than the best individual in the last loop, and the gap is over 0.05, replace the best rate of return and the overall best individual.

Table 1
The rate of return generated by FMAS.

Year	2000	2001	2002	2003	2004	2005	2006	2007
Rate of return	0.055	−1.511	−0.555	−0.913	0.776	0.394	0.169	0.304
Year	2008	2009	2010	2011	2012	2013	2014	
Rate of return	2.225	0.045	0.005	−0.194	−0.284	−0.274	1.153	

Table 2
The rate of return generated by MAS.

Year	2000	2001	2002	2003	2004	2005	2006	2007
Rate of return	−4.559	2.970	−2.737	−2.506	−4.148	5.698	0.490	1.089
Year	2008	2009	2010	2011	2012	2013	2014	
Rate of return	13.852	2.545	−3.145	1.631	−0.749	0.056	7.358	

Table 3
The rate of return generated by SMA.

Year	2000	2001	2002	2003	2004	2005	2006	2007
Rate of return	−1.247	−5.955	−2.875	−2.268	−1.972	−2.868	−0.708	−2.113
Year	2008	2009	2010	2011	2012	2013	2014	
Rate of return	−0.256	−3.703	−1.665	−3.285	0.517	−1.448	−0.127	

Table 4
The rate of return generated by BH.

Year	2000	2001	2002	2003	2004	2005	2006	2007
Rate of return	0.045	−0.258	0.330	0.075	0.254	0.276	−0.145	0.431
Year	2008	2009	2010	2011	2012	2013	2014	
Rate of return	−1.131	0.412	0.129	0.073	−0.132	0.082	−0.681	

(6) Conduct a population evolution. Use genetic algorithms to choose 10% of the individuals from the old population in terms of their rate of return. Next, choose 80% of the individuals from the old population together with 10% randomly generated new individuals to experience the crossover and mutation. Finally, these two parts combined to form the new population.
(7) Repeat Steps (4)–(6) for 50 repetitions.
(8) Calculate the rate of return of the best individual in the test period using the overall best individual selected in the previous steps. Calculate the two moving averages in this test period and the previous test period. Then, obtain the membership function and calculate the rating level and the rate of return.
(9) Move forward to the next experiment period and return to step (1).

3. Results

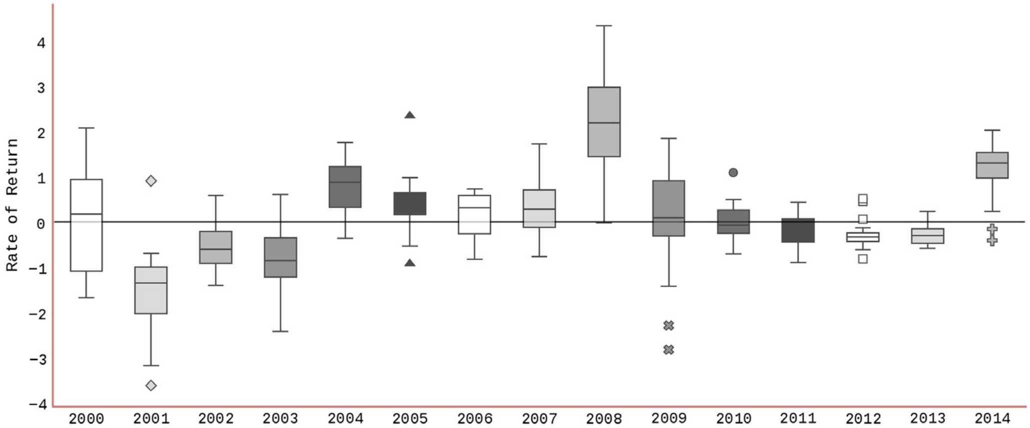
3.1. Rate of return

The test period is from 2000 to 2014. It contains 15 experiment groups. Each group is repeated 20 times, and the mean rate of return of each test period is calculated. The result is shown in Table 1. The average rate of return of all 15 groups are 0.093. The smallest rate of return is −1.511 (in 2001), and the largest one is 2.225 (in 2008). The range is 3.736, and the standard deviation is 0.872. The FMAS gains profit in 9 groups. The percentage of positive return is 60%. The rate of returns is relatively small and stable compared with our previous study, which solely used the moving average strategy [17]. In consideration of the different group composition of the training period, the selection period, and the test period, we conduct another experiment that use the same group composition as we use in this paper. Table 2 shows the result of MAS. The average rate of return is 1.910, the range is 18.411, and the standard deviation is 4.920. In addition, we conduct another two experiment of SMA and Buy-and-Hold (BH) strategy as the benchmark strategies. Table 3 shows the result of SMA, we choose 10 days as the long moving average period and 1 day as the short moving average period. The average rate of return is −1.868, the range is 6.472, and the standard deviation is 0.125. Table 4 shows the result of BH strategy. The average rate of return is −0.049, the range is 1.562, and the standard deviation is 0.625. Overall, the rate of return generated by FMAS is higher than SMA and Buy-and-Hold strategy, and the standard deviation of FMAS is lower than MAS and SMA. In general, the rate of return generated by FMAS is relatively low but stable.

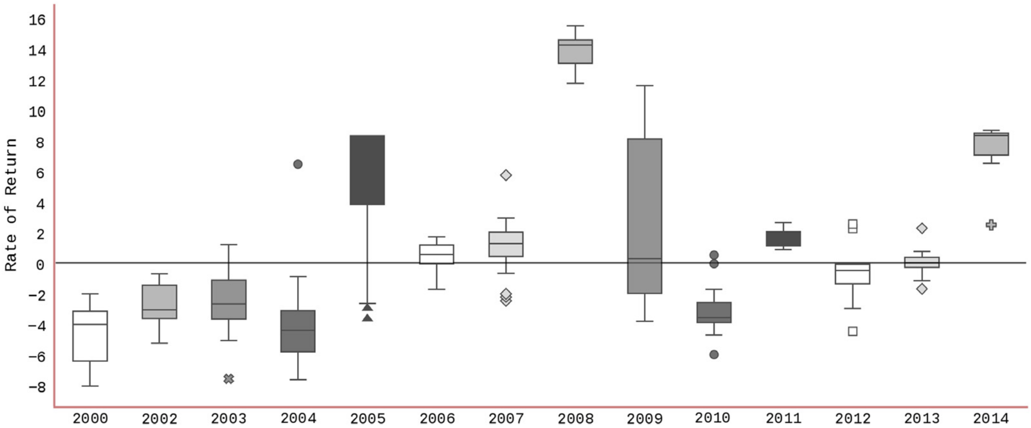
Table 5 gives the standard deviation, range, inter-quartile ranges, and sharp ratio (excess returns per unit standard deviation) of the four methods. From the perspective of sharp ratio, the performance of FMAS is no better than MAS, but we can still find the advantages of FMAS. Compared the results of FMAS with that of MAS (Tables 1 and 2), we can see that they have made the same judgment at the most of the time, that is, when they both get positive or negative returns. While there

Table 5
The summary statistics of the four methods.

	FMAS	MAS	SMA	BH
Standard deviation	0.843	4.753	1.578	0.407
Range	3.736	18.411	6.472	1.562
Inter-quartile range	0.628	5.379	1.893	0.403
Sharp ratio	0.087	0.246	−1.279	−0.088



(a) Return of return by FMAS.



(b) Rate of return by MAS.

Fig. 9. Boxplot of the two experiment.

are three years (2000, 2004, and 2010), the FMAS gains positive returns and MAS have negative returns. This situation may be related to the price trend of the three years.

The boxplot of the two experiments (Fig. 9) describes the entire 300 (15 groups multiply 20 times) results. It is important to note that the Y-axis scales of the two boxplots are different because the ranges of the results are different. The boxplot not only shows the average but also the maximum, upper quartile, lower quartile, minimum, and outliers. Compared with the results generated by the MAS, the results generated by FMAS are more concentrated with less outliers. The ranges of each group by FMAS are small, which means the FMAS method provides more stable results.

3.2. Holding amount

In addition to the rate of return, we also calculate the holding amount, which is generated by FMAS. Because the experiment is repeated 20 times, and the results are not greatly different, we calculate the average holding amount and conduct a comparison with the price series as shown in Fig. 10. The trend of the holding amount series and the price series show great consistency, not only in the general trend but also in the small fluctuations. When the prices increases, the holding amount rises, and when the price decreases, the holding amount declines.

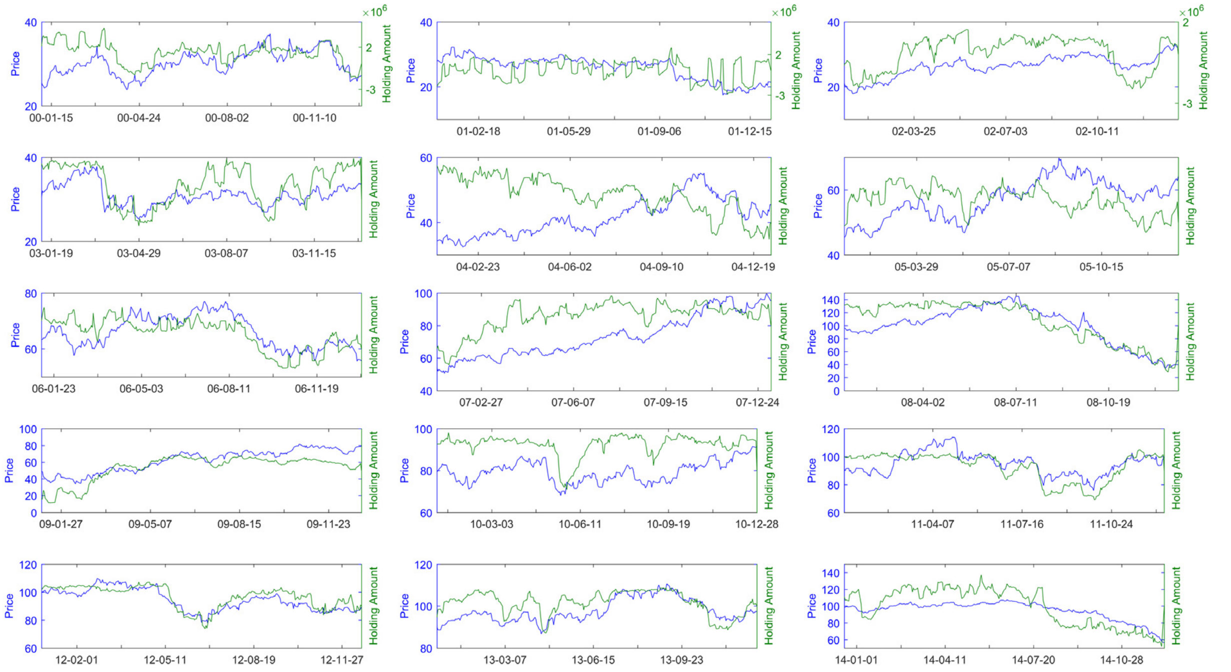


Fig. 10. Holding amount of 15 groups. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The increasing holding amount indicates capital inflows to the futures market. In contrast, the decreasing holding amount indicates that the capital is flowing out of the futures market. Generally, the FMAS tends to a large trading amount when it fairly sure of its own judgment. If the FMAS has doubts of the present market situation, it will tend to not hold.

In Fig. 10 the blue line is the price series, and the green line is the holding amount series. Each group shows one year, which is the test period. The holding amount is the average of 20 times, and the scales of Y-axis are different from each group and the two series.

3.3. Fuzzy logic rule

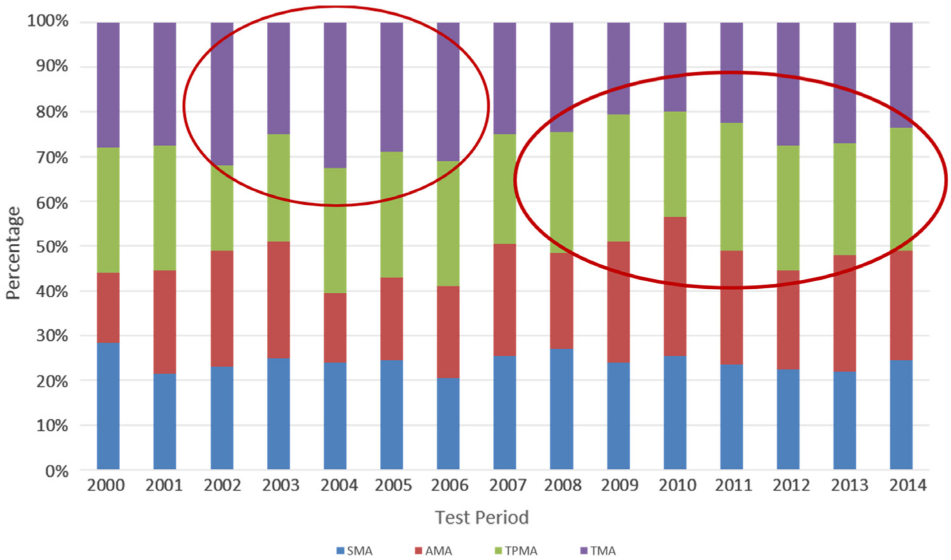
We conduct a statistics of the fuzzy logic rules generated by the FMAS. Each individual contains 10 rules. We repeat the experiment 20 times, and there are 15 test periods. Therefore, there are 3000 ($10 * 20 * 15$) best rules since there is one best individual in every test period. The result is as following.

Firstly, the moving average method situation is different from the previous study. Fig. 11(a) shows the distribution of the moving average methods. Generally, the four methods have a similar share of the entire amount. TPMA and TMA have slight superiority in certain periods, which we mark in red circles on Fig. 11(a). In contrast, AMA, which is the most frequently used in MAS in our previous study [17,18,24], has the least share when using FMAS. This situation may be caused by the different decision mechanism. The FMAS is relatively complicated. Therefore, simple methods such as TPMA and TMA are more popular. AMA concerns many parameters, so it is more suitable in MAS.

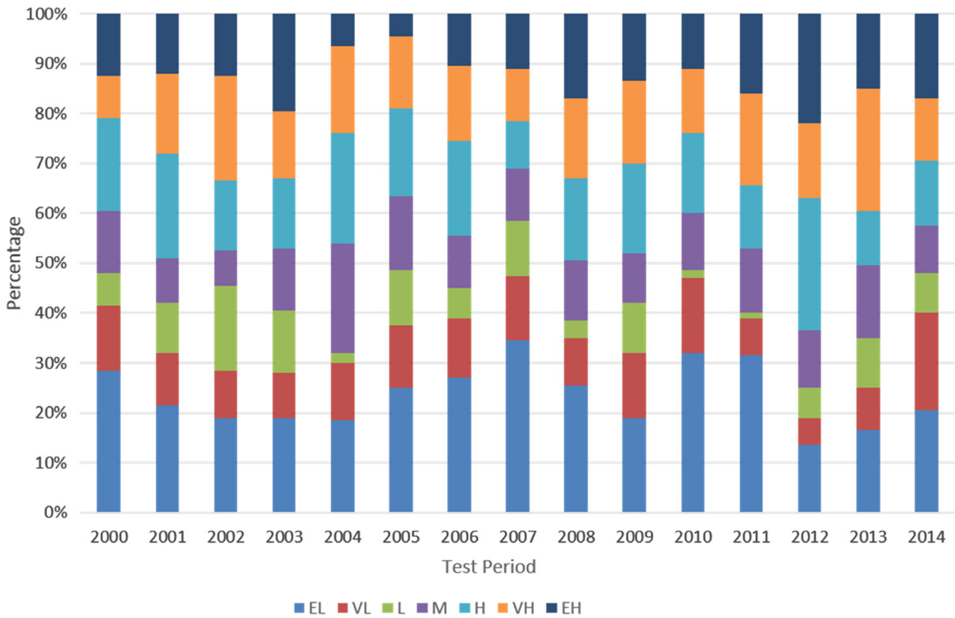
Secondly, the distribution of the 7 fuzzy extents is shown in Fig. 11(b). It is obvious that the EL (Extremely Low) is the most frequently showed fuzzy extent in the test periods. This situation is suitable for the crude oil futures price decline, when the difference of the short period moving average and the long period moving average is larger than the previous test period. The fuzzy extents H (High) and VH (Very High) also take a large part. This combination is the best selected by the genetic algorithms. It may be because this combination contains both a low extent and a high extent so that it considers many situations.

4. Discussion

In the rate of return results (Tables 1 and 2), it is worthy of attention that in most years, the FMAS and MAS both gain profit or lose money, except for the different amounts. The FMAS is the improvement of the MAS. They have similar results. Although we use different moving average methods dynamically, it is impossible to guarantee profits all the time. In nine of the fifteen groups, they both gain or lose. In three test periods (2000, 2004, and 2010), the FMAS gains profit, whereas MAS losses. It is may be caused by the market situation. Perhaps each type of investment strategy can only reflect one aspect of the market.



(a) Moving average methods.

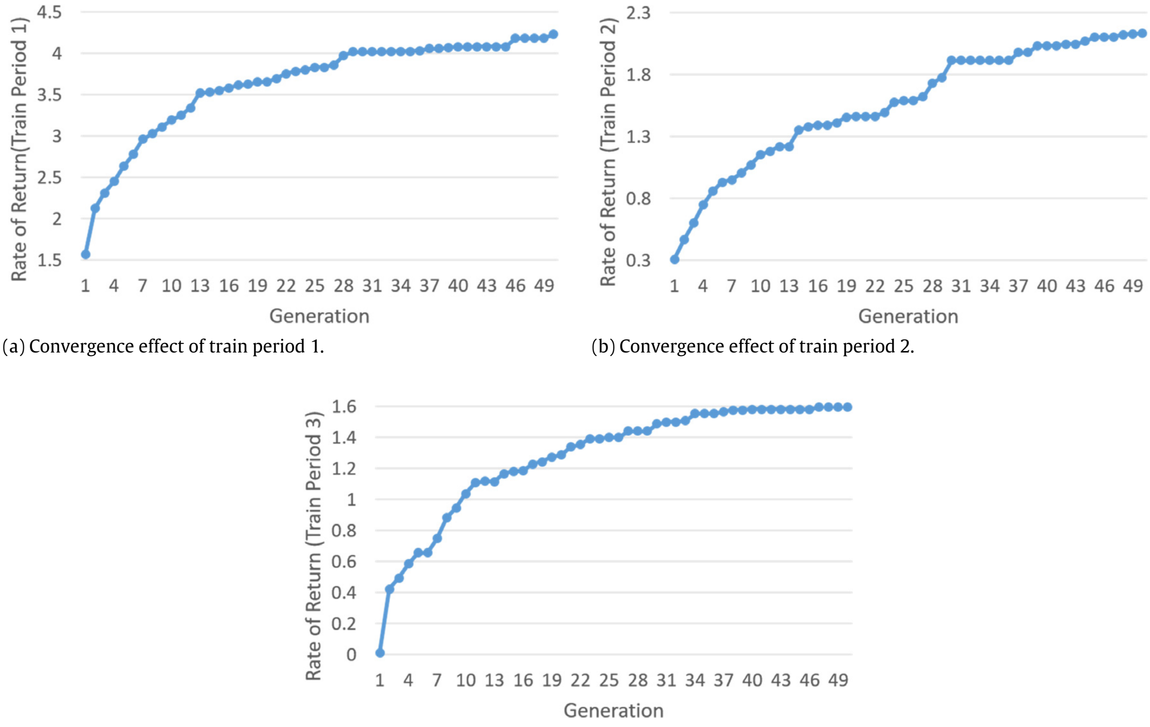


(b) Fuzzy extents.

Fig. 11. Statistics about fuzzy logic rules. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The method used in this paper is a typical technical analysis method. We only use the historical data as the basis for decision making. Therefore, the result highly depends on the price trend. From the previous study, we know that the moving average rules performs better when the market price falls, such as the year of 2008. The FMAS also inherits this property. It gains profit in the year of 2006, 2008, and 2014, which all have significant price fall. While in the year of 2004, when the price experiences a significant raise, the FMAS also gains profit. It proves that the FMAS also has some profitability in the upward market.

The result of the holding amount is interesting. Generally, it fluctuates synchronously with the price series. Although we set a fixed initial capital, if the rating level is the same, the holding amount would increase when the price decreases and would decrease when the price increases. However, here the opposite situation appears. It proves that the holding amount is very sensitive to the price series. In reality, this situation also exists. Usually the price and trading volume have a lead–lag relation. In this paper, the price cannot be affected by other parameters, so the price should be the lead. It is the cause of



(a) Convergence effect of train period 1.

(b) Convergence effect of train period 2.

(c) Convergence effect of train period 3.

Fig. 12. Convergence effect along with the generations.

Table 6
The rate of return under different initial capitals.

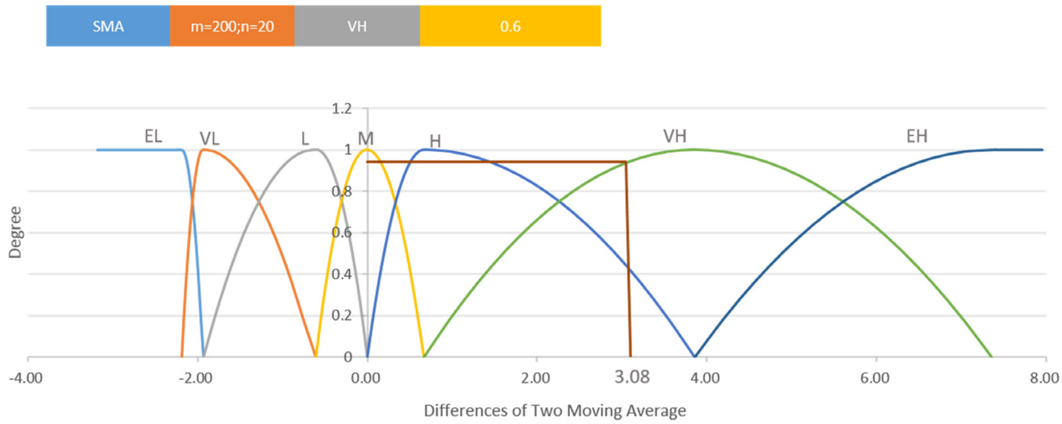
Initial capital	100,000	1,000,000	10,000,000
2000	−0.040	0.055	0.050
2001	−1.271	−1.511	−1.638
2002	−0.568	−0.555	−0.460

other parameters. However, in Fig. 11 the holding amount is obviously not lag. Perhaps this result proves FMAS has a certain prediction ability. We need to perform a further analyze to explore the relation.

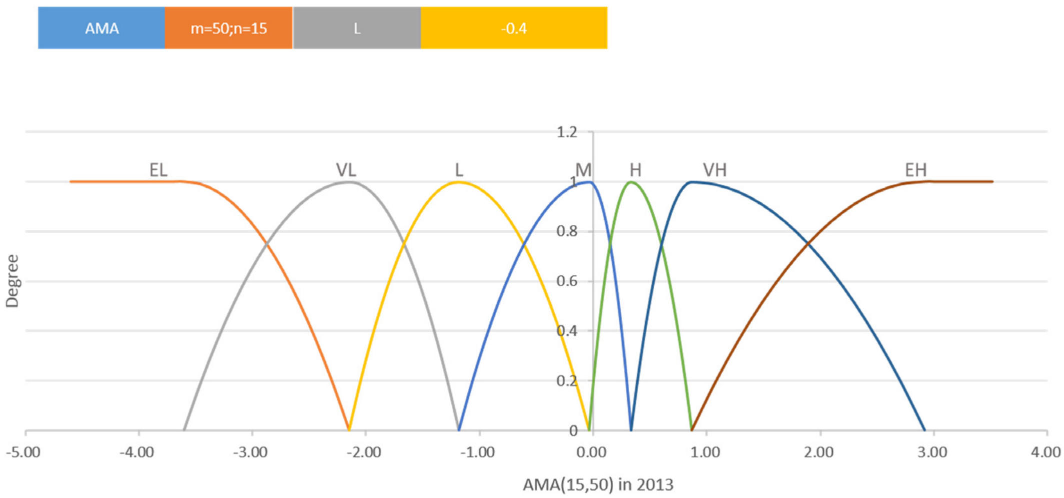
In the process of experiment, the genetic algorithm is not totally random. When using the Fuzzy logic rule, we specify the direction. That is, we set negative recommend values when the fuzzy extent is EL, VL, or L, and positive recommend values when EH, VH, or H. This improvement reduces the burden of the genetic algorithms. Furthermore, the basic principle of the moving average is affirmative. When the short-term moving average line is higher than the long-term moving average line, it is time to buy, and when the opposite occurs, it is time to sell. Therefore, we make the improvement to confirm this principle.

When choosing initial parameters, we set the initial capital as 1 million dollars. It bases on two reasons. Firstly, in crude oil futures market, the trading volume of one transaction must be over 1000. The oil price during the experiment period is between 20 and 140. Therefore, the money that one transaction may cost is between 20,000 and 140,000. Therefore, we set 1,000,000 as the initial capital to ensure the transaction keep going. Secondly, lower initial capital will cause small trading amount. In this paper, the trading amount is based on the initial capital and the rating level. When the rating level is relatively low, we shall suggest trading in small amount. But when the suggested trading amount is less than 1000, the transaction cannot be achieved, which will affect the rate of return. We can see from Table 6 that when initial capital is 1 million or 10 million, the rate of return has little difference. However, when initial capital is 100,000, the rate of return is significantly different from the other two especially in 2000. It is because in 2000 when initial capital is 100,000 there are many trading days that the trading amount is 0. In general, higher capital amount will not affect the result. But individual investors are not all billionaires. Therefore, we choose 1 million as the initial capital.

The choice of generations is based on the experiments. The speed and effect of the convergence are highly related to the crossover probability and the mutation probability. In this paper, we choose 0.7 as the crossover probability and 0.01 as the mutation probability, which are the most commonly used probabilities. The number of generation that the convergence need is determined by observing the experiment. Fig. 12 shows the convergence effect of the first 3 training periods. We can see from the figures that no matter in which period, it is a fast convergence stage before 30 generations. After 30 generations, the trend tends to be smooth and stable. In order to ensure both speed and effect, we choose 50 as the number of generations.



(a) The first example.



(b) The second example.

Fig. 13. Two examples of fuzzy logic rule.

In this paper, we combine genetic algorithms with fuzzy logic rule and the moving average. We compare its performance on crude oil futures with another three methods. For the computational complexity, the BH strategy and SMA are same. For the MAS, the complexity increases due to the use of genetic algorithms. For the FMAS, the computational complexity is higher than the other three in the program we wrote in MATLAB. Higher computation complexity makes the FMAS program much slower. If the transaction happens frequently within a very short time, the program will be ineffective. However, in this paper, we use the daily price of the crude oil futures and the transaction happens once per trading day. At this level, it is acceptable. In future works, we will continue to improve the algorithm to make it more efficient and more suitable for high frequency transactions.

5. Conclusions

In this paper, we combine the fuzzy logic rule and the moving average strategy to form a fuzzy moving average strategy. This FMAS can generate both trading signal and trading volume. The MAS is used to generate the trading signal, and the fuzzy logic rule dictates the trading volume decision. Genetic algorithms method is used for better trading strategy optimization. We conclude that, compared with the MAS, the FMAS generates a low but stable rate of return, and the results of FMAS experiment are more concentrated with less outliers. The holding amount is highly sensitive to price series. When the price increases, the holding amount rises, and when the price decreases, the holding amount declines. The simple moving average such as TPMA and TMA are more efficient. EL, H, and VH are the most frequently showed fuzzy extents.

Authors' contribution

Model design was done by Haizhong An, Xiaojia Liu and Lijun Wang. Program development and experiments performance were done by Xiaojia Liu, Lijun Wang, and Qing Guan. Paper composition was done by Xiaojia Liu, Lijun Wang and Qing Guan.

Acknowledgments

This research was partially supported by grants from the National Natural Science Foundation of China (Grant No. 71173199), the Fundamental Research Funds for the Central Universities (Grant No. 2652016164) and the scholarship from China Scholarship Council (CSC) (Grant No. 201606400042). We also would like to acknowledge the valuable help from Xiaoqi Sun.

Appendix. Two examples of fuzzy logic rule

When the experiment period is 2001 and our program generate a rule which is, *if the SMA (20, 200) is VH, rate* $= 0.6$ as shown in Fig. 13(a). Firstly, we calculate the SMA of 20 days and 200 days and their differences in the previous year (2000). We sort the differences and divide them into 7 parts with equal amounts. Thus, our membership function is explicit as shown in Fig. 13(a). Then, calculate the SMA (20, 200) of January 14th 2001. It is 3.08 and falls into the Very High interval in the 0.94° based on our membership function. The final rating level is $0.94 * 0.6 = 0.564$.

If the moving average value cannot meet the *if* part of a rule, the rule would not provide a rating degree. For example, in Fig. 13(b) *if the AMA (15, 50) is L, then rate* $= -0.4$. However, the AMA (15, 50) on January 20th 2004 is 1.101, which not fall to the Low interval. Thus, this rule does not provide a rating level on that day.

References

[1] A.C. Szakmary, Q. Shen, S.C. Sharma, Trend-following trading strategies in commodity futures: A re-examination, J. Bank. Finance 34 (2) (2010) 409–426.
[2] W.E. Shambora, R. Rossiter, Are there exploitable inefficiencies in the futures market for oil? Energy Econ. 29 (1) (2007) 18–27.
[3] P. Glabadanidis, The market timing power of moving averages: evidence from US REITs and REIT Indexes, Int. Rev. Finance 14 (2) (2014) 161–202.
[4] L.X. Cui, W. Long, Trading strategy based on dynamic mode decomposition: Tested in Chinese stock market, Physica A 461 (2016) 498–508.
[5] P. Glabadanidis, Timing the market with a combination of moving averages, Int. Rev. Finance (2016) http://dx.doi.org/10.1111/irfi.12107, in press.
[6] X. Liu, et al., An integrated approach to optimize moving average rules in the EUA futures market based on particle swarm optimization and genetic algorithms, Appl. Energy 185 (Part 2) (2017) 1778–1787.
[7] S. Asadi, A. Tavakoli, S.R. Hejazi, A new hybrid for improvement of auto-regressive integrated moving average models applying particle swarm optimization, Expert Syst. Appl. 39 (5) (2012) 5332–5337.
[8] C.J. Huang, S.H. Tai, S.L. Lu, Measuring the performance improvement of a double generally weighted moving average control chart, Expert Syst. Appl. 41 (7) (2014) 3313–3322.
[9] Y.S. Jeong, B. Kim, Y.D. Ko, Exponentially weighted moving average-based procedure with adaptive thresholding for monitoring nonlinear profiles: Monitoring of plasma etch process in semiconductor manufacturing, Expert Syst. Appl. 40 (14) (2013) 5688–5693.
[10] L.A. Zadeh, Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems 90 (2) (1997) 111–127.
[11] M. Khashei, S.R. Hejazi, M. Bijari, A new hybrid artificial neural networks and fuzzy regression model for time series forecasting, Fuzzy Sets and Systems 159 (7) (2008) 769–786.
[12] K.Y. Shen, G.H. Tzeng, Fuzzy inference-enhanced VC-DRSA model for technical analysis: Investment decision aid, Int. J. Fuzzy Syst. 17 (3) (2015) 375–389.
[13] R. Dash, P.K. Dash, R. Bisoi, A differential harmony search based hybrid interval type2 fuzzy EGARCH model for stock market volatility prediction, Internat. J. Approx. Reason. 59 (2015) 81–104.
[14] P.C. Chang, C.Y. Fan, J.L. Lin, Trend discovery in financial time series data using a case based fuzzy decision tree, Expert Syst. Appl. 38 (5) (2011) 6070–6080.
[15] A. Ghandar, et al., Computational intelligence for evolving trading rules, IEEE Trans. Evol. Comput. 13 (1) (2009) 71–86.
[16] N. Gradojevic, R. Gencay, Fuzzy logic, trading uncertainty and technical trading, J. Bank. Finance 37 (2) (2013) 578–586.
[17] L.J. Wang, et al., Generating moving average trading rules on the oil futures market with genetic algorithms, Math. Probl. Eng. (2014) 10. http://dx.doi.org/10.1155/2014/101808, Article ID 101808.
[18] X. Liu, H. An, L. Wang, Performance of generated moving average strategies in natural gas futures prices at different time scales, J. Nat. Gas Sci. Eng. 24 (0) (2015) 337–345.
[19] J. How, M. Ling, P. Verhoeven, Does size matter? A genetic programming approach to technical trading, Quant. Finance 10 (2) (2010) 131–140.
[20] H. Dong, W. Huang, Prediction of geothermal resources by means of wavelet neural network optimized by genetic algorithm, Res. Ind. 16 (3) (2014) 101–106.
[21] A. Behroozsarand, H. Soltani, Hydrogen plant heat exchanger networks synthesis using coupled Genetic Algorithm-LP method, J. Nat. Gas Sci. Eng. 19 (2014) 62–73.
[22] X.F. Li, Z.M. Lu, Optimizing the controllability of arbitrary networks with genetic algorithm, Physica A 447 (2016) 422–433.
[23] Z.T. Li, J. Liu, A multi-agent genetic algorithm for community detection in complex networks, Physica A 449 (2016) 336–347.
[24] L. Wang, et al., Selecting dynamic moving average trading rules in the crude oil futures market using a genetic approach, Appl. Energy 162 (2016) 1608–1618.