ROBUST OPTIMIZATION TO AVOID CURVE-FITTING IN ALGORITHMIC TRADING

Farid Helalat, Department of Business Administration, Istanbul Aydin University, Istanbul, Turkey
farid.helalat@gmail.com

ABSTRACT

Curve-fitting is the biggest obstacle on the way of success in algorithmic trading. Anoptimization approach of financial markets trading system development is evaluated in this study which is called "Scientific Approach". The heart of this approach is known as Walk Forward Analysis and if a system can pass it, the development would be followed by Monte Carlo Simulation. In this study I coded four different trading strategies for trading on EURUSD and Crude Oil. Although they all showed profitability during the in-sample test, but only two of them, (Bollinger Reversion with 143.6 % and RSI Reversion with 161.33 % annualized return during the out-of-sample test) have passed the Walk Forward Analysis. The most important finding of this study is that the scientific approach is a reliable, practical and useful method of trading strategy development. Unprofitable trading systems cannot pass scientific approach tests. This approach avoids curve-fitting.

KEYWORDS-Optimization; Curve-fitting, Walk Forward Analysis

I. INTRODUCTION

Due to some psychological causes of trading failure[3]among financial markets traders like arrogance, greediness, revenge trading, fear of failure and Also the power and speed of computers, many traders tend to become an algorithmic trader and the most common reason of their failure is known as curve-fitting (over-fitting or over-optimization) which caused by insufficient degree of freedom during an optimization. Using many rules and variables during an optimization along with inadequate historical data, create a situation that best trading set found during the optimization is fitted to the data to the extent that it cannot produce acceptable amount of profit in out-of-sample testing [2] [6]. There are many traders that perform an optimization with many variables involved with very tight scanning step and also, they do not validate in-sample results by out-of-sample testing (performing at least one cross-validation). This leads to curve-fitting and losing trading systems.

II. METHODOLOGY

In this study, we test a method of optimization and development of trading systems known as "***Scientific Approach***"[6] that many successful algorithmic traders are using and teaching. It is claimed that this approach is designed to avoid curve-fitting during a trading systemoptimization.Inthisresearch, four trading systems are coded by researcher to study whether the scientific approach can really avoid curve-fitting.

Most of the trading systems demonstrate a kind of profitability during the in-sample testing (optimization) which is called the "***Edge Window***" of the trading system. In other word, the edge window is the period that trading system shows a profitable equity curve and it will be closed when the drawdown period (losing capital) starts due to market changes. Curve-fitting creates a situation in which, the edge window does not last long enough so that it does not continue from the in-sample period into the out-of-sample with acceptable positive correlation. Figure 1shows the curve-fitted trading system that does not demonstrate profitability during the out-of-sample testing; while the Figure 2 illustrates the trading system with the edge window during out-of-sample. It means that the edge window lasts long enough so that it continues from in-sample into the out-of-sample period and the trade settings (parameters) found during in-sample testing works with positive and acceptable correlation during the out-of-sample but soon or later, the edge window will be closed and the trading system needs re-optimization to adapt itself to the new market situation. That is why the Walk Forward Analysis [6] (Pardo, 2011, p. 237) is used in Scientific Approach.

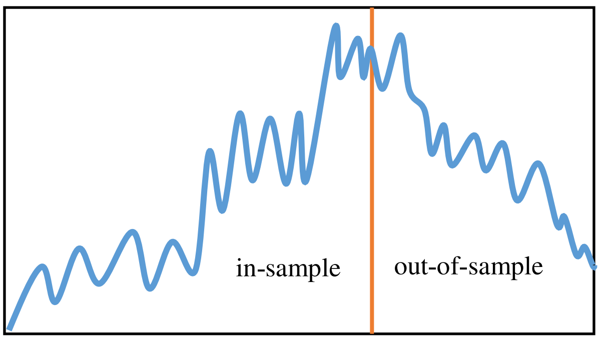


Figure 1. Illustrates the curve-fitting

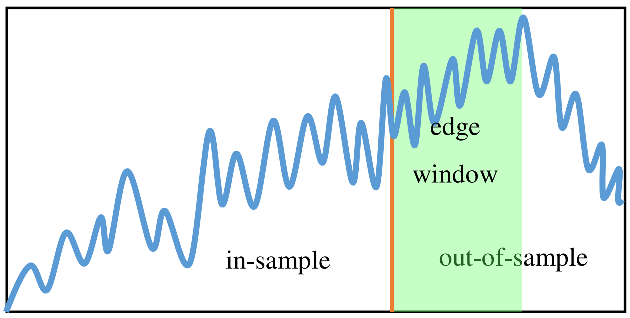
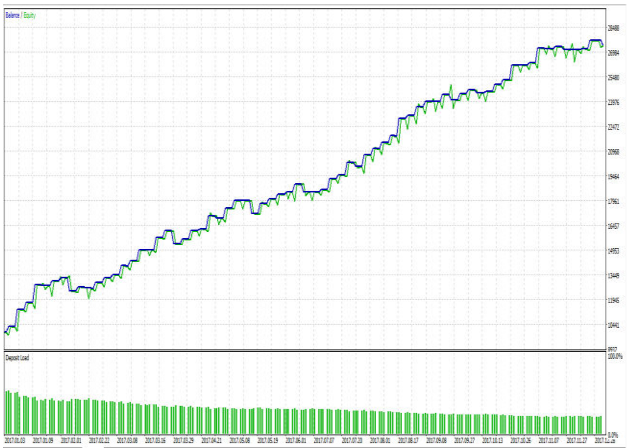


Figure 2. Illustrates the out-of-sample edge window

Variable	Value	Start	Step	Stop	Steps
☑ Reverse	false	false		true	2
☑ MA_Period	5	5	1	50	46
☑ MA_Method	Simple	Simple		Linear wei...	4
☑ MA_Price	Close price	Close price		Weighted ...	7
☑ Env_Period	24	24	1	240	217
☑ Env_Method	Simple	Simple		Linear wei...	4
☑ Env_Price	Close price	Close price		Weighted ...	7
☑ Deviation	0.05	0.05	0.005	0.1	11
☑ Cross_Shift	5	1	1	10	10
☐ Lot	0.5	0.1	0.01	1.0	
☑ SL	100	10	1	200	191
					328843813760

Figure 3shows an optimization space which has many variables involved in an optimization with too small scanning steps. It leads to curve-fitting due to insufficient degree of freedom. Also, inadequate degree of freedom can be created by too short optimization window (insufficient historical data for optimization) which cannot support variables and rules involved in an optimization.



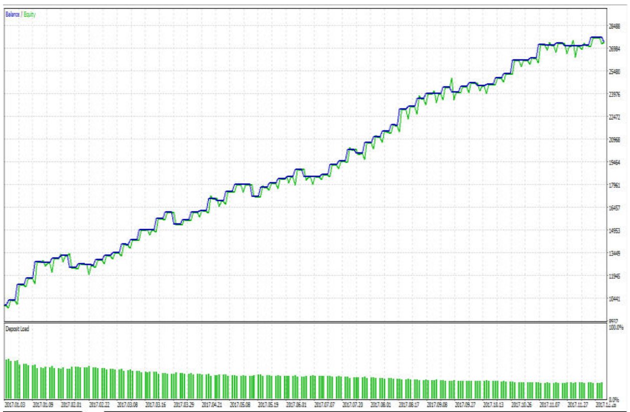


Figure 4 shows the equity curve of best trading set found during the optimization. As you can see, the equity curve has very good shape but the out-of-sample result shown in Figure 5 illustrates an over-fitted trading system.

Variable	Value	Start	Step	Stop	Steps
☑ Reverse	false	false		true	2
☑ MA_Period	5	5	1	50	46
☑ MA_Method	Simple	Simple		Linear wei...	4
☑ MA_Price	Close price	Close price		Weighted ...	7
☑ Env_Period	24	24	1	240	217
☑ Env_Method	Simple	Simple		Linear wei...	4
☑ Env_Price	Close price	Close price		Weighted ...	7
☑ Deviation	0.05	0.05	0.005	0.1	11
☑ Cross_Shift	5	1	1	10	10
☐ Lot	0.5	0.1	0.01	1.0	
☑ SL	100	10	1	200	191
					328843813760

Figure 3. An optimization space with too many variables involved

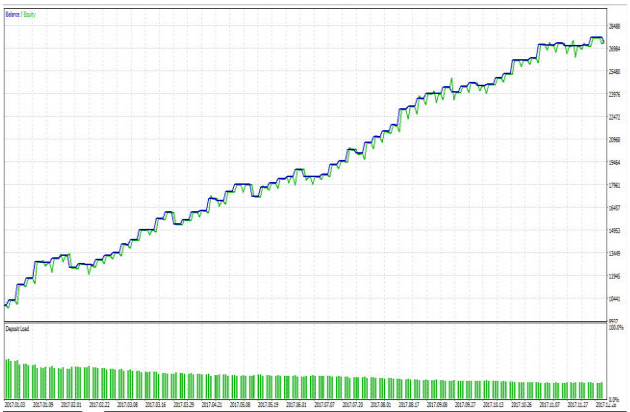


Figure 4. Shows the most profitable set with 17,499 $ profit during 2017

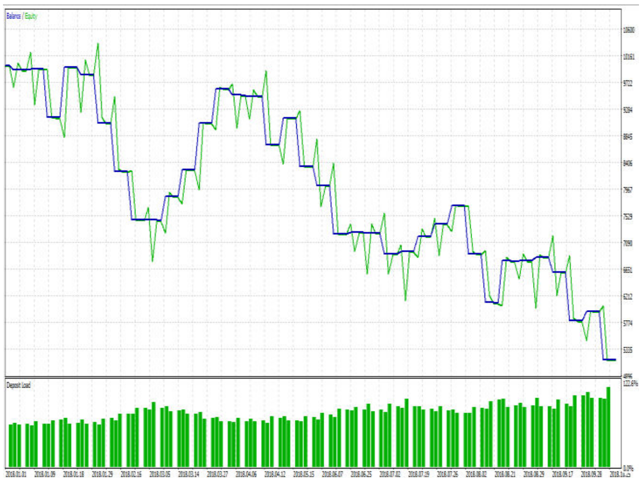


Figure 5. Equity curve of selected set during 2018 with 4,842 $ loss

A. *Broker and Platform*

Determined broker in this study is "***Info Broker***" which is a Forex broker and regulated under the SPK rules. SPK is the capital markets board of the Turkey. All the trade costs during the optimizations and simulations are automatically applied by platform of the Info broker in the trades. In other words, the spread, commission and swap are considered in net profit calculation during the study. ***Meta-Trader-5 (MT5)*** is used as the platform in this study. It has two main parts. In "***Trade Terminal"***, one has access to price charts and data which are downloaded automatically from the broker server. Also, we can optimize and run the trading systems on a certain account. The second part of MT5 is known as "***Meta Editor***" and it provides a programming possibility based on the MQL5 (MetaQuotes Language 5) which is similar to C++.

B. *Sample data*

EURUSD and Crude Oil are two symbols involved in the study. They are tested and optimized on H1 timeframe. For EURUSD the sample data is from 2016.01.01 to 2019.04.01 with 20,160 H1 bars. Also, the lower timeframes (M30, M15, M5, M1) are available for the H1 period to support the accurate tick simulation by MT5. Furthermore, on the Crude Oil the sample data is from 2017.11.01 to 2019.04.01 (includes 8,265 H1 bars) and also the lower timeframes (M30, M15, M5, M1) are available for the given period.

C. *Measures*

The fitness function is net profit and the robustness function is "***Walk-Forward Efficiency***" (WFE) [6] (Pardo, 2011, p. 239)which is calculated as follow:

$$WFE = \frac{O \times R}{I} \times 100$$

O: Out-of-sample profit
I: In-sample profit
R: Time-length Ratio between the in-sample to the out-of-sample.

For example, if the in-sample result is 5000$ profit during the three months and the out-of-sample have shown 1000$ in one month the WFE is (1000 × 3) / 5000 × 100 = 60%. In other words, we can expect to experience 60% of profitability of the in-sample testing during live trading or out-of-sample testing.

D. *Approach details*

Scientific Approach has three main parts as follow:

- Coding and preliminary testing
- Walk Forward Analysis
- Monte Carlo Simulation

After coding a trading idea, some preliminary testing would be performed to make sure there is no error in the code and the system is trading as it is supposed to trade. Then preliminary optimization would be done to evaluate whether the variables involved in the optimization are really effective on the fitness function. If a variable has no important effect on the fitness function it is better to keep it as a constant value and improve the degree of freedom to avoid curve-fitting as much as possible. In next step Walk Forward Analysis [6](Pardo, 2011, p. 237)is performed and results of the test will be transferred to Microsoft Excel for further evaluation.

As it is shown inFigure 6, the Walk Forward Analysis[2] [6]is a series of cross validations. In-sample is called ***"Optimization Window"*** and the out-of-sample is ***"Walk-Forward Test"*** (WFT) [6]. During Walk Forward Analysis the best trading parameters found during optimization #1 will be applied in WFT #1 and then the performance (profit, drawdown, trade number and etc.) would be recorded for further evaluation. Process continues with moving optimization window forward in time as much as one WFT. Now we optimize during optimization #2 and best trade setting would be applied for WFT #2. This process goes on until the end of the historical data (sample). The results of All WFTs creates ***"Cumulative WFT"*** that illustrates complete out-of-sample performance. For each WFT the WFE would be recorder and if system is profitable enough with average WFE above 50%, the systemhas passed the Walk Forward Analysis. Also, the R2 of cumulative WFT equity curve helps to have a better understanding about the consistency of profitability and drawdown period. In this study the R2 >= 0.7 would be considered as a desirable shape of the equity curve.

Jan	Feb	Mar	Apr	May	Jun
optimization #1			WFT		
	optimization #2			WFT	
		optimization #3			WFT
			optimization #4		
			Cumulative WFT		

Figure 6. Walk Forward Analysis

If a system passes the Walk Forward Analysis, the Monte Carlo Simulation [2] would be performed toselect best initial account balance to trade so that system can survive during drawdown periods and avoid margin failure. Margin failure is a situation in which the system cannot trade one standard contract due to shortage of capital. Before performing the simulation, trade sample and few candidate-account-balances must be transferred into Microsoft Excel. For each hypothetical initial account balance, simple Monte Carlo Simulation would be done for 2500 times by shuffling the sequence of trade results randomly to evaluate to what extent the drawdown and risk of margin failure can ger worse. For each hypothetical initial account balance, the following results will be transferred to the evaluation table:

- Risk of margin failure
- Median drawdown
- Median profit
- Median annualized return (AR%)
- Probability of being profitable

Lowest initial account balance which has risk of margin failure lower that 10% would be accepted in Monte Carlo Simulation.

E. *Trading strategies logic and settings*

1) *MA-Candle Cross*

It is a trend following trading system. Trade logic is based on crossover between the simple moving average (SMA) [5] and candles on the H1 chart. Two SMA with same period are applied on the chart; One based on high price and the other one applied on low price. This creates a band and if the last bar closes above the band, then price shows a big enough pullback followed by an enough upward (bullish) movement, the system would be ready to open a buy position. This is a symmetrical trading system and sell entry terms are just opposite the buy entry terms. Also, opening a new position leads to stop-and-reverse situation. In other words, a buy entry will close the last sell position.

Also, we put the stop-loss (SL) equals to high (for sell) or low (for buy) of the last closed bar add/minus the SMA (200) of ATR (14) [5]. Furthermore, Take-profit (TP) is the entry price add/minus a coefficient of SMA (200) of ATR (14).

Optimization variables are as follow:

- MA_Period: period of SMA
- PBX: size of the pullback
- PTX: size of candle
- TP: take-profit based on ATR indicator

Table 1. Optimization space of MA-Candle Cross

Variable	Start	Scanning Steps	Stop	Optimization Steps
MA_Period	10	5	50	9
PBX	50	50	200	4
PTX	50	50	200	4
TP	1	1	7	7
Optimization Space				1008

2) Breakout CTP

It is a trend following system. Trade logic is based on pivot breakout on the chart. The system detects the last peak and pit and if the price (Bid) passes the last pivot, system would trade based on the breakout direction. SL is the opposite possible pivot between the main pivot and the bid price or the high/low of the last nbars. TP is not based on price movement on the chat; it is based on time. After a certain number of bars, the system will close the position.

Optimization variables of Breakout CTP trading strategy:

- Depth: Depth of pivot
- CTP: Candle-based take-profit
- SLC: Candle-based stop-loss

Table 2. Optimization space of Breakout CTP

Variable	Start	Scanning Steps	Stop	Optimization Steps
Depth	2	2	14	7
CTP	2	unequal	12	8
SLC	0	2	2	2
Optimization Space				112

3) Bollinger Reversion

It is a counter trend following system (reversion system). The trade logic is based on crossing the upper/lower band of the Bollinger Bands [1] and then breakout with last n bars. For instance, if the last bar is closed above the upper band and bid falls below the low price of last n bars, sell signal would be generated. As it is a symmetrical trading system, the buy entry situation is just opposite the sell signal. Each entry causes stop-and-reverse situation and closes the last opposite position.

Optimization variables are as follow:

- SD: Standard deviation of Bollinger bands
- RNG: Range of breakout

Table 3. Optimization space of Bollinger Reversion

Variable	Start	Scanning Steps	Stop	Optimization Steps
SD	1	0.25	3	9
RNG	0	1	5	6
Optimization Space				54

4) RSI Reversion

It is a counter trend following system (reversion system). A Bollinger Bands indicator would be applied on RSI (14) [4]. If RSI closes above the upper band and then in the next bar it closes below the upper band sell signal is generated which results in closing last buy position. As it is a symmetrical trading system, if RSI closes below the lower band and in the next bar it closes above the lower band buy signal would be generated and also it leads to close last sell position. SL is calculated based on maximum high (for sell position) or low (for buy position) of the last 3 bars added with a coefficient of SMA (200) of ATR (14).

Optimization variables are as follow:

- Mean: RSI averaging period
- Dev: Standard deviation
- SL:ATR-based coefficient for SL determination

Table 4. RSI Reversion: Optimization space

Variable	Start	Scanning Steps	Stop	Optimization Steps
Mean	8	Unequal	62	5
Dev	1	0.2	3	11
SL	1	1	3	3
Optimization Space				165

F. *General optimization settings*

The general optimization settings during the study are as follow:

- Fitness function:Net Profit
- Robustness function: WFE
- Optimization search method: Grid search
- Optimization execution method:
    - Open prices only for MA-Candle Cross
    - Every tick for Breakout CTP, Bollinger Reversion and RSI Reversion.
- Spread: 0.3 pip (0.0001) on EURUSD and 1 pip (0.01) on Crude Oil (as fixed spread offered by the broker)
- Swap: 0 pip (based on trading account specifications offered by the broker)
- Commission: 0.5 pip (as sum of deal-in and deal-out commissions, based on trading account specifications offered by the broker)
- Trade symbols: EURUSD and Crude Oil
- Trade timeframe: H1
- Trade volume: 100,000 eurofor EURUSD and 1000 barrels for Crude Oil as one standard contract

III. RESULTS

A. *MA-Candle Cross Results*

After 82 cross validations during Walk Forward Analysis the average WFE was 9.6% with 87$ average profit of out-of-sample testing. Cumulative out-of-sample performance demonstrated 7,101$ profit after 82 walk-forward tests (each test was 2 weeks). The Ratio between in-sample to out-of-sample is 1:0.5-month.

Table 5:Walk-Forward Analysis results of MA-Candle Cross

Performance	Out-of-sample Profit $	WFE %	Drawdown $
Mean	87	9.6	705
Cumulative	7,101	***	4,101

As the WFE of the system is only 9.6 % it did not pass the Walk Forward Analysis.

B. *Breakout CTP Results*

After 9 cross validations during Walk Forward Analysis the average WFE was 66% with 20$ average loss of out-of-sample testing. Cumulative out-of-sample performance demonstrated 2,060$ loss after 9 walk-forward tests (each test was one month). The Ratio between in-sample to out-of-sample is 8:1-month.

Table 6: Walk-Forward Analysis results of Breakout CTP

Performance	Out-of-sample Profit $	WFE %	Drawdown $
Mean	-20	66	15,280
Cumulative	-2060	***	3,766

As the system showed loss during out-of-sample testing, it did not pass the Walk Forward Analysis.

C. *Bollinger Reversion Results*

After 14 cross validations during Walk Forward Analysis the average WFE was 92.05% with 2,094$ average profit of out-of-sample testing. Cumulative out-of-sample performance demonstrated 29,320$ profit after 14 walk-forward tests (each test was one month) and the ratio between in-sample to out-of-sample was 3:1-month.

Table 7: Walk-Forward Analysis results of Bollinger Reversion

Performance	Out-of-sample Profit $	WFE %	Drawdown $
Mean	2,094	92.05	3,029
Cumulative	29,320	***	12,610

As the WFE of the system is 92.05% and its out-of-sample performance shows 29,320$ profit over 14 months, the Bollinger Reversion passed Walk Forward Analysis.

Figure 7. Bollinger Reversion: Cumulative equity curve

Based on the Monte Carlo Simulation results of Table 8, initial account balance of 17,500 $ has 4.8% risk of margin failure. Among 2500 Monte Carlo tests, median annualized rate of return for 17,500$ account is 139.5% and 94% of simulations have shown profitability. As a result, by considering 17,500$ as the initial account balance, 29,320$ cumulative out-of-sample profit over 14 months means 143.60% annualized return for trading with Bollinger Reversion trading system.

Table 8. Monte Carlo Simulation of Bollinger Reversion

Initial account balance	Risk of margin failure	Median drawdown	Median profit	Median AR%	Probability of profitability
$12,500	25.1%	34.9%	22,700	181.6%	75%
$15,000	10.7%	30.5%	24,290	161.9%	88%
$17,500	4.8%	27.9%	24,420	139.5%	94%
$20,000	2.6%	25.6%	25,360	126.8%	95%
$22,500	1.2%	22.6%	24,960	110.9%	97%

D. *RSI Reversion Results*

After 14 cross validations during Walk Forward Analysis the average WFE was 66.13% with 2,353$ average profit of out-of-sample testing. Cumulative out-of-sample performance demonstrated 32,940$ profit after 14 walk-forward tests (each test was one month) and the ratio between in-sample to out-of-sample was 3:1-month.

Table 9: Walk-Forward Analysis results of RSI Reversion

Performance	Out-of-sample Profit $	WFE %	Drawdown $
Mean	2,353	66.13	3,741
Cumulative	32,940	***	15,170

As the WFE of the system is 66.13% and its out-of-sample performance shows 32,940$ profit over 14 months, the Bollinger Reversion passed Walk Forward Analysis.

Based on the Monte Carlo Simulation results in Table 10, initial account balance of 17,500$ has 9.6% risk of margin failure. Among 2500 Monte Carlo tests, median annualized rate of return for 17,500$ account is 159.9% and 89% of simulations have shown profitability. As a result, by considering 17,500$ for initial account balance, 32,940$ cumulative out-of-sample profit over 14 months means 161.33% annualized return for trading with RSI Reversion trading system.

Table 10. Monte Carlo Simulation of RSI Reversion

Initial account balance	Risk of margin failure	Median drawdown	Median profit	Median AR%	Probability of profitability
$17,500	9.6%	31.9%	27,975	159.9%	5.08
$20,000	5.4%	29.1%	27,570	137.9%	4.65
$22,500	2.9%	26.9%	28,510	126.7%	4.62

Figure 8. RSI Reversion: Cumulative equity curve

IV. CONCLUSIONS

The research has shown that the scientific approach of trading strategy development is capable to evaluate whether a trading idea is worth trading or not and it avoids curve-fitting.

Strategies with short edge window life, do not perform well under Walk Forward Analysis.

WFE is capable to be used as a robustness metric and its value is one of the major factors that should be considered during robustness evaluation.

The Walk Forward Analysis avoids algorithmic traders to apply a losing/marginal trading system on their trading account. Also, it helps profitable systems to fit their trading parameters to the recent market situation without over-fitting. It improves the adaptivity to the market changes and therefore it improves the consistency of profitability of the trading systems.

REFERENCES

[1] Bollinger, J. (2001). *Bollinger on Bollinger Bands.* McGraw-Hill.

[2] Davey, K. J. (2011). Building Winning Algorithmic Trading Systems: A Trader's Journey from Data Mining to Monte Carlo Simulation to Live Trading. John Wiley & Sons.

[3] Douglas, M. (2000). Trading in the zone: master the market with confidence, discipline and a winning attitude. New York Institute of Finance.

[4] J. Welles Wilder, J. (1978). *New Concepts in Technical Trading Systems.* Trend Research.

[5] Murphy, J. J. (1999). Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications . New York Institute of Finance.

[6] Pardo, R. (2011). *The evaluation and optimization of trading strategies* (Second ed.). John Wiley & Sons, Inc.