

Performance evaluation of trading strategies in multi-agent systems – Case of A-Trader

Jerzy Korczak*, Marcin Hernes**

*International University of Logistics and Transport, Wrocław, Poland

**Wrocław University of Economics, Poland

e-mail: {jerzy.korczak, marcin.hernes}@ue.wroc.pl

Abstract—The article presents the problem related to evaluation of Forex trading strategies in multi-agent systems. The ratios based on financial measures cannot be assumed to be only evaluation criteria because other aspects determining effectiveness of the strategies, such as, for instance, investment risk, statistics on winning, and lost transactions, transaction costs, should also be taken into consideration. The aim of this paper is to review the general financial investments performance measures in relation to the performance analysis of trading strategies. The characteristics of the commonly used performance measures are outlined. The discussion will be illustrated by solutions developed in the trading support system, called A-Trader system. The performance analysis in A-Trader is detailed on real FOREX quotations.

I. INTRODUCTION

TODAY the multi-agents systems are increasingly being used as trading support on Forex market [1, 2]. Agents operating in such systems provide strategies for open/close long/short positions with the use of various investments methods and techniques. There arises the need of constant evaluation of the agent performance to guarantee satisfactory benefits to the trader. Many of the performance evaluation methods [e.g. 2, 3, 4] and practical solutions (e.g. MetaTrader [5], Plus500 [6], MetaStock [7] or NinjaTrader [8]) are based on ratios of return (e.g. rate of return, gross profit, and the number of unprofitable transactions).

Investment decisions are made under conditions of risk and uncertainty. In the case of a trading decision, it is difficult to talk about the optimal decision as the decision bringing the highest possible rate of return for the investor, rather should be said about the decision bringing a satisfactory rate of return for the trader under a given level of other factors (e.g. risk). Many authors [e.g. 9, 10] have drawn attention to the fact that making an optimal decision is in practice very difficult in a situation of risk and uncertainty. If, on the other hand, we adopt the principle of satisfactory benefits, referred to in the literature as the principle of subjective expected utility [11], decision-making process becomes less complicated. According to this principle, for each alternative, its expected usefulness can be determined, and then the alternative that has the most

usability is to be selected. The idea is to "set the bar" at such a level that it would not be too low, because then the result would be unsatisfactory, and not too high, because it might be unattainable for a trader. Note that in this case, the trader can set the "bar" higher and higher in sequence, which will bring him closer to the optimal value. Therefore, the ratios based measures cannot be assumed as the only evaluation criterion because other aspects having influence on the effectiveness of the strategies, such as, for instance, investment risk [12], statistics on winning and lost transactions as well as transaction costs should also be taken into consideration.

The aim of this paper is to review the general financial investments performance measures in relation to the performance analysis of trading strategies. The discussion will be illustrated by solutions developed and provided by the A-Trader system [13]. A-Trader system is composed of the agents capable of generating independent trading decisions on FOREX market. It allows for High Frequency Trading (HFT). It is realized in near real time (the notion of "near real time" can be understood as a very short delay between the last quote and the time of generated trading decision; usually 10-250 ms) and characterized by high speed, short-term positions, it concentrates attention on price formation process using sophisticated algorithms based on efficient and robust indicators and modern IT [14]. High frequency traders take decisions on the basis of real-time quotes changes in order to achieve satisfactory rate of return.

In the first part of this paper, the characteristics of performance measures are outlined. Next, the methods of performance analysis in the trading support system, called A-Trader, are detailed. In the final part of the article, conclusions and future works are presented.

II. PERFORMANCE MEASURES FOR FINANCIAL INVESTMENT

There are many performance measures in related works. These measures was developed in economics, management, and finance, both by researchers and practitioners. The related works [15, 16] divide performance measures for financial investments into three main groups:

- measures based on ratios of excess returns (ratios-based),
- measures based on systematic risk measured by factor models (risk-based),
- measures based on endogenous benchmarks derived from portfolio theory (benchmarks-based).

The major differences between these groups, but also between specific measures refer to the definition of risk. The next part of the section presents characteristics of particular groups.

A. Measures based on ratios of excess returns

Ratio-based performance measures specify the return per unit of risk. Ratio-based performance measures are usually easy to compute and have only low data requirements. These measures are of high relevance in practical applications and are frequently used in publications [12, 15, 17,18].

All ratios of return-based measures follow a similar schema: a measure of the return of asset in excess of the return on the benchmark is divided by a measure of the investment risk of asset. The most popular are the following::

- arithmetic rate of return,
- logarithmic rate of return,
- the number of transaction,
- gross profit,
- gross loss,
- total profit,
- the number of profitable transactions,
- the number of profitable transactions in a row,
- the number of unprofitable transactions in a row.

All of them are available in A-Trader.

B. Measures based on systematic risk measured by factor models

Risk-based performance measures adjust for risk by computing the spread between actual returns and a hypothetical benchmark return which is determined [17].

These measures indicate whether the trader was able to beat the benchmark, strictly speaking, they do not allow for comparison of different investment products because risk-based performance measures are subject to manipulation by leverage. For identification of relevant and meaningful risk factors and computation of “fair” or expected returns, risk-based performance measures draw heavily from the asset pricing literature. This group contains measures, such as [17, 19]:

- *Sharpe Ratio*

$$S = \frac{E(r)-E(f)}{|O(r)|} \cdot 100\% \quad (1)$$

where:

$E(r)$ – arithmetic average of the rate of return,

$E(f)$ – arithmetic average of the risk-free rate of return,

$O(r)$ – standard deviation of rates of return.

- *Treynor Ratio*

$$T = \frac{E(r)-E(f)}{\beta(r)} \quad (2)$$

where:

$E(r)$ – arithmetic average of the rate of return,

$E(f)$ – arithmetic average of the risk-free rate of return,

$\beta(r)$ – beta coefficient of rates of return.

- *The Kappa ratio*

$$T = \frac{E(r)-E(f)}{\sqrt[n]{LPM(r)}} \quad (3)$$

where:

$E(r)$ – arithmetic average of the rate of return,

$E(f)$ – arithmetic average of the risk-free rate of return,

$LPM(r)$ – lower partial moments of rates of return.

- *Omega ratio*

$$\Omega = \frac{\sqrt[n]{HPM(r)}}{\sqrt[n]{LPM(r)}} \quad (4)$$

where:

$HPM(r)$ – higher partial moments of rates of return.

$LPM(r)$ – lower partial moments of rates of return.

- *Average coefficient of variation*

$$V = \frac{s}{|E(r)|} * 100\% \quad (5)$$

where:

V – average coefficient of variation,

s – average deviation of the rates of return,

$E(r)$ – arithmetic average of the rates of return.

- *Jensen Model*

$$JM = E(r) - (E(f) + \beta(r) \cdot (E(m) - E(f))). \quad (6)$$

where:

$E(r)$ – arithmetic average of the rate of return,

$E(f)$ – arithmetic average of the risk-free rate of return,

$E(m)$ – arithmetic average of the realized return of the appropriate market index,

$E(f)$ – arithmetic average of the risk-free rate of return,

$\beta(r)$ – beta coefficient of rates of return.

- *Value at Risk*

The measure known as a value exposed to the risk - that is the maximum possible loss of the market value that a financial instrument can bear in a specific timeframe and at a given confidence level.

$$VAR = P * O * k \quad (7)$$

where:

P – the initial capital,

O – volatility - standard deviation of rates of return during the period ,

k – the inverse of the standard normal cumulative distribution (assumed confidence level 95%, the value of k is 1,65).

Some of these measures are available in A-Trader (Sharpe ratio, average coefficient of variation and Value et Risk).

C. Measures based on endogenous benchmarks derived from portfolio information

This group of measures usually compares the return of each security in the portfolio to the return of a “benchmark” security in order to determine abnormal performance. The comparable securities are selected based on characteristics, or they are derived from portfolio in another time period. Consequently, data requirements are higher for these models, while the statistical concepts are relatively simple [15, 16].

- *Characteristic-Based Models*

Characteristic-Based Models are interpreted as portfolio-weighted sum of the differences in returns between the stocks and the benchmark portfolios, and can be calculated, for example by following equation:

$$CM_t = \sum_{j=1}^m w_{jt} (r_{jt} - r_{jt-1}^b) \quad (8)$$

where:

w_{jt} – weight of asset j at time t ,

r_{jt} – corresponding excess return of asset j ,

r_{jt-1}^b – the return on a benchmark portfolio that is matched to asset j measured in $t - 1$.

- *Holdings-Based Models*

These models define managerial skill as a co-variation between portfolio weights and returns of single stocks taking into consideration an omega ratio.

$$HM = \sum_{j=1}^m Cov(w_{jt}, r_{jt} | \Omega_t) \quad (9)$$

where:

w_{jt} – weight of asset j at time t ,

r_{jt} – corresponding excess return of asset j at time t ,

Ω_t – omega ratio at time t .

- *Trade-Based Models*

These models define managerial skill as a co-variation between portfolio weights and returns of single stocks.

$$TM = \sum_{j=1}^m Cov(w_{jt}, r_{jt}) \quad (10)$$

where:

w_{jt} – weight of asset j at time t ,

r_{jt} – corresponding excess return of asset j at time t ,

Performance measures presented in this section allow for a wide range of evaluation of investment strategies on Forex. There are many other measures in the related works, however we have tried to select such as are more often used in practice and which can be implemented in High-Frequency trading systems due to low computational complexity.

In A-Trader, a characteristic based models are available (based on buy and hold and random walk benchmarks).

Next part of paper presents method for performance evaluation in A-Trader multi-agent system.

III. PERFORMANCE EVALUATION METHOD IN A-TRADER

In general, A-Trader system is composed of the agents capable of generating independent trading decisions on FOREX market. It should be noted that decisions can be consistent or contradictory, e.g. two independent agents may generate buy and sell decision at the same time [20, 21]. The trading opportunities are provided by consensual advice, generated by multiple software agents that use technical and fundamental analysis as well as behavioral sentiments [22]. Trading agents in the A-Trader form the investment strategies, which advise recommended open and closed positions for online FOREX traders. There are many strategies implemented as Supervisor Agents, such as:

- Basic Strategy,
- Consensus,
- Candle genetic algorithm,
- Kohonen network,
- Growing neural gas,
- Fundamental back propagation network,
- Evolutionary algorithm.
- Deep Learning

Supervisor Agent is the most important agent in A-Trader. Its goal is to generate profitable trading advice, on the basis of three groups of The Supervisor Agent coordinates functioning of the other agents (which form a given strategy) presented, and to provide the final advice to the trader. Its other task include resolving conflicts between agents [23]. The strategies are permanently evaluated by Supervisor Agents, and those with the highest evaluation value can be taken by default or chosen by the trader.

The performance analysis in A-Trader is carried out with the consideration of the following measures (ratios):

- rate of return (ratio x_1),
- number of transactions,
- gross profit (ratio x_2),
- gross loss (ratio x_3),

- total profit (ratio x_4),
- number of profitable transactions (ratio x_5),
- number of profitable consecutive transactions (ratio x_6),
- number of unprofitable consecutive transactions (ratio x_7),
- Sharpe ratio (ratio x_8)
- average coefficient of variation (ratio x_9)
- Value at Risk (ratio x_{10})
- the average rate of return per transaction (ratio x_{11}).

There are many ways of defining the performance evaluation function. For the purpose of comparison of the agents' performance, the following simple evaluation function has been proposed:

$$y = (a_1x_1 + a_2x_2 + a_3(1-x_3) + a_4x_4 + a_5x_5 + a_6x_6 + a_7(1-x_7) + a_8x_8 + a_9(1-x_9) + a_{10}(1-x_{10}) + a_{11}x_{11}) \quad (11)$$

where x_i denote the normalized values of particular performance measures from x_1 to x_{11} . It was adopted in the test that coefficients a_1 to $a_{11}=1/11$.

It should be mentioned that these coefficients may be modified with the use of, for instance, an evolution-based method, or they could be determined by the trader in accordance with their preferences (for instance the trader may determine whether they are interested in higher rate of return with accompanying higher risk level or lower risk level but accepting a lower rate of return).

The output of the function is a value in the range [0..1],

and the agent's efficiency is directly proportional to the function value.

Figure 1 presents performance evaluation panel in A-Trader. The upper part of window presents information related to open/close positions generated by strategy in a given period. The profitable ones are marked on green and unprofitable ones are marked on red. The bottom part of window presents performance evaluation values related to selected positions (it is possible to mark all positions, or only selected positions).

Referring to the evaluation analysis related to particular measures performed in other systems (mentioned in section 1), as previously underlined, these systems only offer the functions calculating the rates of return based ratios. It should be noted that evaluation, in most cases, is performed "manually" by the trader. This work has many inconveniences. Due to its time consumption, the trader can use only selected measures of performance, and choice of these measures may be narrow. Also the trader acting under time pressure may select inadequate measures, and, in consequence, important financial losses may be generated. In addition, it is very difficult to have valid current knowledge on online trading (the trader's knowledge, in very turbulent market conditions, may be outdated or/and incomplete). These issues imply that systems operating in real time are very limited.

As has been mentioned, the evaluation function used in A-Trader enables the evaluation of performance of specific strategies. These operations are made automatically, in time close to real time, by the Supervisor Agent which may then

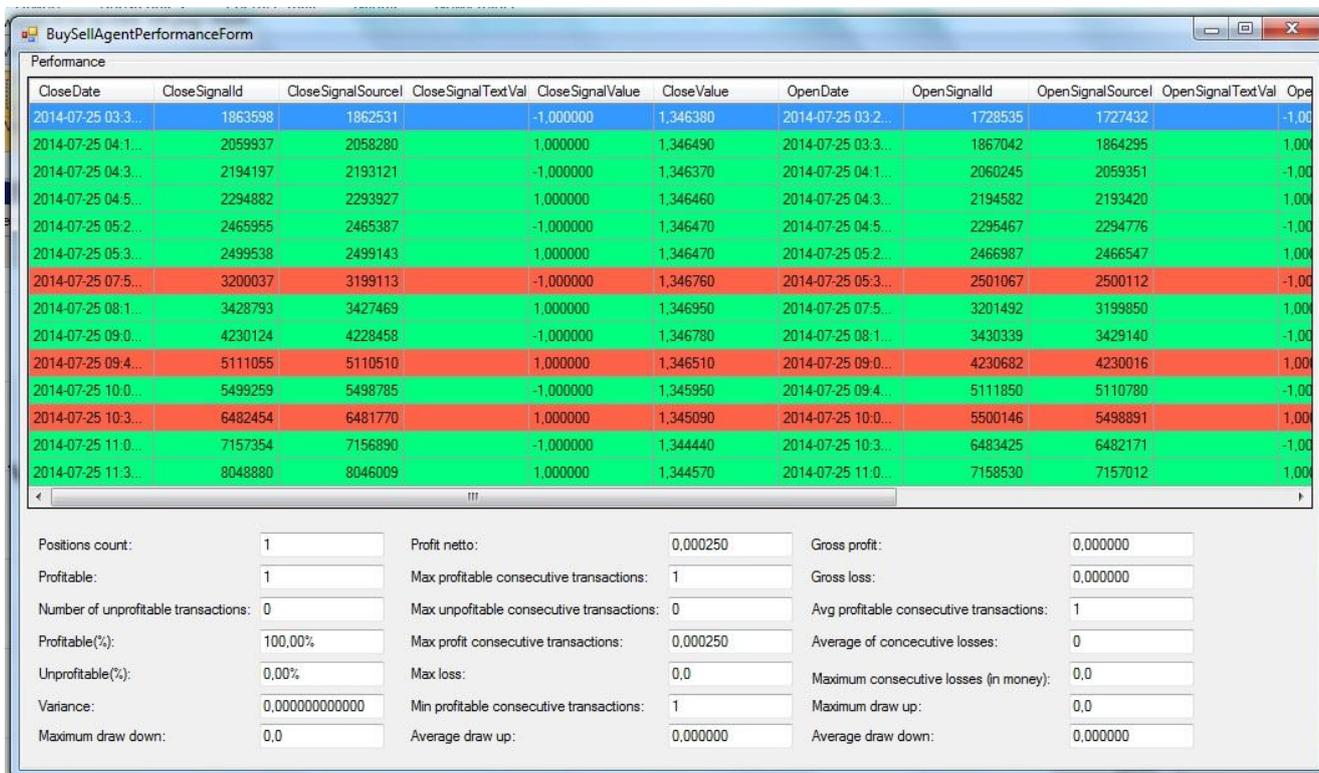


Fig. 1. Performance evaluation panel in A-Trader. Source: Own work.

suggest to the investor taking final decisions on the basis of strategy with the highest level of performance. In addition, enabling the user to change coefficients a_i and ratios x_i parameters of the evaluation function allows for considering their preference concerning the criterion of importance of particular evaluation ratios. The performance evaluation function also considers the transaction costs with the assumption that this reflects the relationship between the number of transactions and the average rate of return from the transaction. However, this simple principle cannot be adopted because a large number of transactions has impact on the reduction of the strategy's efficiency level, especially for the transactions with a high rate of return.

A-Trader uses only selected measures because using largest number of ratios require higher computing resources than are now available (in HFT this measures must be calculated near real time). In future, we plan to use the cloud computing resources, then we will be in a position to implement a larger number of measures. It should be noted that the number of performance measures in A-Trader is not limited and they can be added to evaluation function in an easy way.

IV. CONCLUSION

The strategies in the A-Trader system open/close independent long/short position use multiple criteria of trading performance, which belong to three groups of performance measures: ratios-based, risk based, and benchmark-based. As a consequence, this enables the trader to compose an evaluation function according to their preferences and to apply to the strategies of the best Supervisor Agents. The results presented in our previous research [13, 14, 21, 22, 24, 25] allow us to come to the conclusions that there is no universally accepted evaluation function nor universal measures. The choice of measures and the composition of the evaluation function is highly dependent on trader preferences and trading market. We have already demonstrated that the level of performance of particular strategies changes depending on prevailing FOREX market situation. Based on the obtained results, there is no one strategy which definitely dominates over the others.

The use of this performance evaluation function allows for automatic setting of the best strategy in time close to real time, which has, in turn, a positive influence on investment effectiveness.

Future works should concern, among others, implementation of other performance evaluation measures (presented in section 2), development of an evolution method for determining ai coefficients into the A-Trader system, and implementation of cognitive agents performing analysis experts' opinions in the scope of forecasts referring to quotations on the FOREX market..

REFERENCES

- [1] M. Aloud, E.P.K. Tsang and R. Olsen, "Modelling the FX Market Traders' Behaviour: An Agent-based Approach", [in] Alexandrova-Kabadjova B., S. Martinez-Jaramillo, A. L. Garcia-Almanza & E. Tsang (eds.), *Simulation in Computational Finance and Economics: Tools and Emerging Applications*, IGI Global, 2012, pp. 202-228. DOI: 10.4018/978-1-4666-2011-7.ch015.
- [2] R.P. Barbosa and O. Belo, "Multi-Agent Forex Trading System", [in] *Agent and Multi-agent Technology for Internet and Enterprise Systems, Studies in Computational Intelligence Volume 289*, 2010, pp. 91-118. https://doi.org/10.1007/978-3-642-13526-2_5.
- [3] L. Mendes, P. Godinho and J. Dias, "Forex trading system based on a genetic algorithm", *J. J Heuristics* 18(627), 2012, <https://doi.org/10.1007/s10732-012-9201-y>
- [4] A. Shmilovici, Y. Kahiri, I. Ben-Gal, et al., "Measuring the Efficiency of the Intraday Forex Market with a Universal Data Compression Algorithm", *Computational Economics* 33(131), 2009, <https://doi.org/10.1007/s10614-008-9153-3>
- [5] MetaTrader5, <https://www.metatrader5.com>
- [6] Plus500, <https://www.plus500.com/>
- [7] MetaStock, <https://www.metastock.com/>
- [8] NinjaTrader, https://www.ninjatradbrokerage.com/get_started/forex
- [9] F. Gul and W. Pesendorfer, "Hurwicz expected utility and subjective sources", *Journal of Economic Theory*, V 159, Part A
- [10] R. Pettigrew, "Risk, rationality and expected utility theory", *Canadian Journal of Philosophy*, 45:5-6, 798-826, DOI: 10.1080/00455091.2015.111961, 2016.
- [11] J. Shanteau and A. Pingenot, "Subjective expected utility theory", [in] M. W. Kattan (Ed.), *Encyclopedia of medical decision making* (pp. 1084-1086). Thousand Oaks, CA: SAGE Publications Ltd., 2009, doi: 10.4135/9781412971980.n312.
- [12] K. Jajuga and T. Jajuga, "Inwestycje: Instrumenty finansowe, ryzyko finansowe, inżynieria finansowa", PWN, Warszawa 2000.
- [13] J. Korczak, M. Hernes and M. Bac, "Risk avoiding strategy in multi-agent trading system", [in] *Proceedings of the 2013 Federated Conference on Computer Science and Information Systems / Ganzha Maria, Maciaszek Leszek, Paprzycki Marcin (red.), IEEE*, 2013. DOI: 978-1-4673-4471-5/\$25.00.
- [14] J. Korczak, M. Hernes and M. Bac, "Collective Intelligence Supporting Trading Decisions on FOREX Market", [in] Nguyen N., Papadopoulos G., Jędrzejowicz P., Trawiński B., Vossen G. (eds) *Computational Collective Intelligence. ICCCI 2017. Lecture Notes in Computer Science*, vol 10448. Springer, Cham, 2017. https://doi.org/10.1007/978-3-319-67074-4_12.
- [15] P. Lückoff, "Mutual Fund Performance and Performance Persistence: The Impact of Fund Flows and Manager Changes", Wiesbaden: Gabler Verlag 2011.
- [16] Y. Yao and R. Zhang, "Empirical Research on Efficiency Measure of Financial Investment in Education Based on SE-DEA". [in] Cao BY., Liu ZL., Zhong YB., Mi HH. (eds) *Fuzzy Systems & Operations Research and Management. Advances in Intelligent Systems and Computing*, vol 367. Springer, Cham, 2016. https://doi.org/10.1007/978-3-319-19105-8_35
- [17] O.K. Hussain, T.S. Dillon, F.K. and Hussain, E.J. Chang, "Risk Assessment Phase: Financial Risk Assessment in Business Activities", [in] *Risk Assessment and Management in the Networked Economy. Studies in Computational Intelligence*, vol 412. Springer, Berlin, Heidelberg, 2013. https://doi.org/10.1007/978-3-642-28690-2_6
- [18] Z. Qiu, "Discussion of Investment Analysis Method in the New Round of the China Stock Bull Market", [in] Li M., Zhang Q., Zhang J., Li Y. (eds) *Proceedings of 2015 2nd International Conference on Industrial Economics System and Industrial Security Engineering*. Springer, Singapore, 2016. https://doi.org/10.1007/978-981-287-655-3_40.
- [19] D. Hu, , G. Schwabe and X. Li, "Systemic risk management and investment analysis with financial network analytics: research opportunities and challenges", *Financial Innovation* (2015) 1: 2. <https://doi.org/10.1186/s40854-015-0001-x>.
- [20] J. Korczak, M. Bac, K. Drelczuk and A. Fafuła, "A-Trader - Consulting Agent Platform for Stock Ex-change Gamblers", [in] *Proc. FedCSIS, Wroclaw*, pp.963-968, 2012.. DOI: 978-83-60810-51-4/\$25.00.

- [21] J. Korczak and M. Hernes, "Deep Learning for Financial Time Series Forecasting in A-Trader System", [in] Proceedings of the 2017 Federated Conference on Computer Science and Information Systems / Ganzha Maria, Maciaszek Leszek, Paprzycki Marcin (eds.), Annals of Computer Science and Information Systems, vol. 11, Polskie Towarzystwo Informatyczne, Institute of Electrical and Electronics Engineers, 2017. DOI: 10.15439/2017F449.
- [22] M. Hernes and J. Sobieska-Karpińska, "Application of the consensus method in a multi-agent financial decision support system", *Information Systems and e-Business Management* 14 (1), Springer Berlin Heidelberg, 2016, DOI: 10.1007/s10257-015-0280-9.
- [23] J. Korczak, M. Hernes and M. Bac, "Fuzzy Logic as Agents' Knowledge Representation in A-Trader System", [in] Information Technology for Management. Federated Conference on Computer Science and Information Systems, ISM 2015 and AITM 2015, Lodz, Poland, Revised Selected Papers / Ziemia Ewa (red.), Lecture Notes in Business Information Processing, vol. 243, 2016, Springer, ISBN 978-3-319-30527-1, ss. 109-124, 2015.
- [24] J. Korczak, M. Hernes and M. Bac, "Risk avoiding strategy in multi-agent trading system", [in] Proceedings of the 2013 Federated Conference on Computer Science and Information Systems / Ganzha Maria, Maciaszek Leszek, Paprzycki Marcin (red.), IEEE, 2013.