2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

Sliding window-based analysis of multiple foreign exchange trading systems by using soft computing techniques

Rodrigo F. B. de Brito and Adriano L. I. Oliveira

Abstract— **Considerable effort has been made by researchers from various areas of science to forecast financial time series such as stock market and foreign exchange market. Recent studies have shown that the market can be outperformed by trading systems built with soft computing techniques. This paper aims to compare different trading systems based on support vector regression (SVR), growing hierarchical self-organizing maps (GHSOM) and genetic algorithms (G A) when tested against nine currency pairs of the foreign exchange market (Forex). The experiments were performed using the sliding window strategy. The results showed that the GA-based trading systems outperformed the SVR+GHSOM model when evaluated by four performance metrics, including an statistical test.**

I. Introduction

The Efficient Market Hypothesis (EMH) [1] states that the financial markets are unpredictable and no profit can be made, no matter what strategies are being used, since this kind of time series follows a random walk model [2]. Nevertheless, efforts are being made [3] by many fields of science (e.g., economics, physics, mathematics, computer science) aiming to predict financial time series, and the results proved the opposite from what the EMH proposes, that is, showing that soft computing techniques are able to tackle with the difficult task of time series forecasting, generating consistent results in terms of profitability [4] [5] [6] [7] [8].

Trading in foreign exchange (Forex) markets averaged US$5.3 trillion per day in April 2013 [9]. Forex is the most traded market since currencies are negotiated by multiple players, including banks, governments, business and common individuals. Besides that, it is a very liquid market. Although researches commonly focus on stock markets, some works have been proposed on Forex forecasting with the aid of soft computing [10].

Technical analysts of the financial market use technical indicators such as MACD (moving average convergence/divergence) and RSI (Relative Strength Index) to make profit, but in the long term this strategy may not give good results in terms of profit [11]. Therefore, the application of computational intelligence techniques for forecasting has been investigated by researchers, including the multilayer perceptron (MLP) neural network [11], radial basis functions (RBF) [12], self-organizing maps (SOM) [13], support vector machine (SVM) [14], extreme learning machine (ELM) [15], among others [3]. Most of the papers that appear in the literature focus on the application of such techniques for the stock market. Nevertheless, there has been an increasing interest in the application of computational intelligence techniques to Forex market rates prediction due to the importance of these markets for the international monetary scenario [16] [8].

In our previous papers [8] [17], trading systems were developed based on soft computing techniques, such as Support Vector Regression (SVR), Growing Hierarchical Self-organizing Maps (GHSOM) and Genetic Algorithms. Basically, two models were created, the SVR+GHSOM and the GA-based. In each paper, the results were evaluated against different data sets periods. Besides that, each trading system performed differently when tested in different scenarios. Although in [17] we have compared both models, the analysis were not complete due to the different scenario from [8] and the few metrics used to benchmark each model and their trading systems. This paper proposes to compare the models using larger data sets using the same periods, along with a sliding window strategy. Also, the models were analyzed with more performance indicators commonly used by investors, including an statistical test and were evaluated against nine foreign exchange currency pairs instead of only two currency pairs as in [8] and [17].

This paper is organized as follows. In Section II we briefly review the techniques used in the trading methods investigated in this paper, namely, support vector regression, growing hierarchical self-organizing maps and genetic algorithms. Section III describes the proposed hybrid models. Section IV presents the experiments and the analysis of the results. Finally, in Section V conclusions and suggestions for further research are presented.

II. Fundamentals

A. Support Vector Regression

Support vector regression (SVR) is one of the components of the model investigated in this paper for forex trading. SVR is closely related to SVM and is based on the structured risk minimization principle. It was introduced by [18], which defines the ϵ-insensitive zone in the error loss function. It has been proposed as a good alternative to MLP neural networks in time series forecasting, having a high generalization performance in time series modeling. Besides that, the solution for SVR is unique and globally optimal, in contrast to many other networks and learning algorithms which tend to be trapped to local minima. A schematic representation of the SVR using the ϵ-insensitive loss function is illustrated in Fig. 1.

Rodrigo F. B. de Brito and Adriano L. I. Oliveira are with the Department of Computing Systems, Informatics Center, Federal University of Pernambuco, Recife, Brazil (email: {rfbb, alio}@cin.ufpe.br).

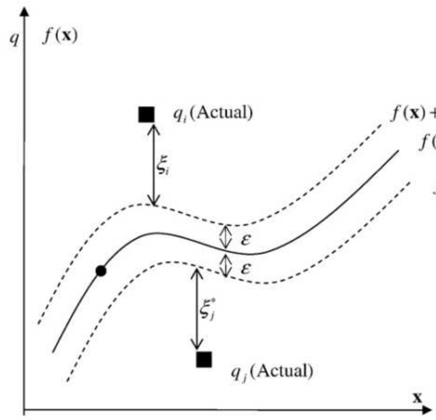

Fig. 1. An schematic representation of the SVR e-intensive loss function. (Figure adapted from [19])

A penalty is introduced when data-points are far from the predicted line, but no penalty is received when they are inside the ϵ-tube. That is, errors inside the tube are considered to be zero. When the penalty occurs, the errors are measured by the variables ξ and ξ^*.

The function that represents $f(x)$ in the case of SVR for nonlinear regression is defined by

$$f(x) = \langle w, \phi(x) \rangle + b, \tag{1}$$

where ϕ is some nonlinear function that maps the input space to a higher dimensional feature space. The weight vector w and the threshold b are chosen via optimization. The optimization problem can be stated as

$$\begin{aligned} \text{minimize} \quad & f(x) = \tfrac{1}{2}\|w\|^2 + C \sum_{i=1}^{l} (\xi_i + \xi_i^*), \\ \text{subject to} \quad & (\langle w, \phi(x_i) \rangle + b) - y_i \leq \epsilon + \xi_i, \\ & y_i - (\langle w, \phi(x_i) \rangle + b) \leq \epsilon + \xi_i^*, \\ & \xi_i, \xi_i^* \geq 0. \end{aligned} \tag{2}$$

The constant $C > 0$, which is one of the user-defined parameters of training together with ϵ, determines the trade-off between the flatness of f and the amount up to which deviations larger than ϵ are tolerated. ξ and ξ^* are the slack variables, which measure the cost of penalties on the training points. ξ and ξ^* measure the deviations from training points outside the ϵ-tube to $f(x) + \epsilon$ and $f(x) - \epsilon$ respectively, as shown in Fig. 1. The idea of SVR is to minimize an objective function which considers both the norm of the weight vector w and the losses measured by the slack variables (see (2)). Minimizing the norm of w is one of the ways to ensure the flatness of f [20].

The decision function can be computed by the inner products of $\phi(x_i)\phi(x_i)$ without explicitly mapping x into a higher dimension, which saves considerable computation efforts. Thus, the kernel function is defined as

$$K(x_i, x_j) = \langle \phi(x_i), \phi(x_i) \rangle. \tag{3}$$

In this work, we consider SVRs with the RBF kernel, which is computed as

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0, \tag{4}$$

where ,γ is a parameter of the kernel and must be defined by the user.

B. Growing Hierarchical self-organizing maps (GHSOM)

The GHSOM was proposed by [21] and is based on the self-organizing maps (SOM) [22]. GHSOM is used mainly as a clustering technique and offers as an advantage the fact that it is not necessary to specify the shape of its grid, that is, it is not required to have a previous knowledge of the problem to define the grid. The GHSOM has a hierarchical structure of multiple layers, with each layer representing an independent growing SOM, as shown in Fig. 2, which is an example of GHSOM. Layer 0 is responsible for the control of the growth process. In the example of Fig. 2, the next layer (Layer 1) consists of a 3x2 grid. All of the units of Layer 1 are expanded to six additional SOMs, creating Layer 2. In this layer, many units have not been expanded since the data representation quality was already accurate enough. It is important to note that the maps have different sizes according to the structure of the data. That is the reason why this work adopted GHSOM, since it is not necessary to define its structure beforehand.

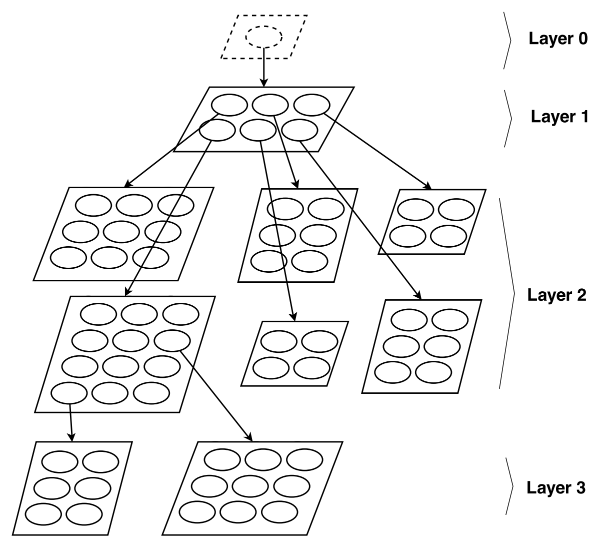

Fig. 2. Architecture of a GHSOM.

C. Genetic Algorithms

Genetic algorithms (GAs) are computer search procedures based on the principles of natural selection. These procedures were firstly brought by Holland (1975). GAs are able to find the global best solution with a high probability; they are used in many practical applications for solving complex problems, such as function optimization. In GA, the decision space is referred as the environment. The potential solutions to the optimization problem are called chromosomes, which are solutions that represent a set of decision variables. The total number of solutions is called the population size and an iteration of the optimization process is called a generation. Reproduction, crossover and mutation are the essence of a standard genetic algorithm. In the first step, the GA produces an initial population randomly, each consisting of some potential solutions. The decision variables are usually encoded as strings of binary digits or real numbers. The objective function values, called fitness function, are then calculated for each individual. For minimization problems, individuals with lower objective function values will have a higher probability of being selected for the next generation. New individuals are then created from parents of the current generation by mutation and crossover, generation the next generation. It replaces old individuals in the population and are usually similar to their parents. The individuals survive if they are fitted to the given environment. Details on Genetic Algorithms technique is provided by [23].

III. Proposed Hybrid Models

As already mentioned, this paper focus on a sliding window-based analysis of the trading systems proposed by previous works [17] [8]. Two hybrid models were proposed, the SVR+GHSOM and the GA-based one.

The GA-based method proposed in [17] is depicted in Fig. 3, which illustrates its training and test phases. The same trading systems are used in both training and test phase. In this paper we investigated fifteen trading systems. Each system consists of a combination of four technical indicators (EMA, MACD, RSI and Stochastic). The parameters of each technical indicator (TI) are optimized in the training phase using a genetic algorithm in order to maximize the fitness function, which is the profit value obtained during the training period. The best parameters of each TI are used in the test phase where the performance of the system is analyzed. This process was executed with the 15 trading systems.

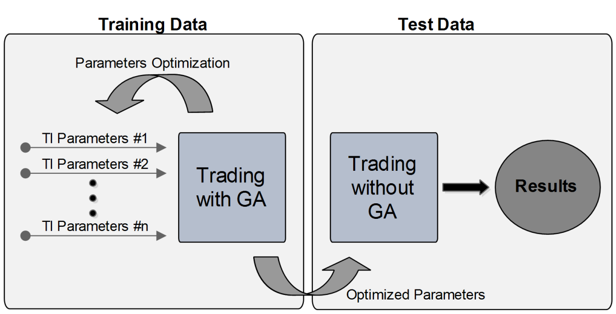

Fig. 3. Trading flow by using genetic algorithms and technical indicators.

The second model is depicted in Fig. 4 and is composed of two stages; the first one divides the entire data set into partitions with similar statistical distributions using a GHSOM network. This is done to tackle the problem of the nonstationarity of financial time series. The second stage consists of using a different SVR model for each region to make predictions. This architecture was first proposed by [24] and was inspired by the divide-and-conquer principle applied to simplify complex problems.

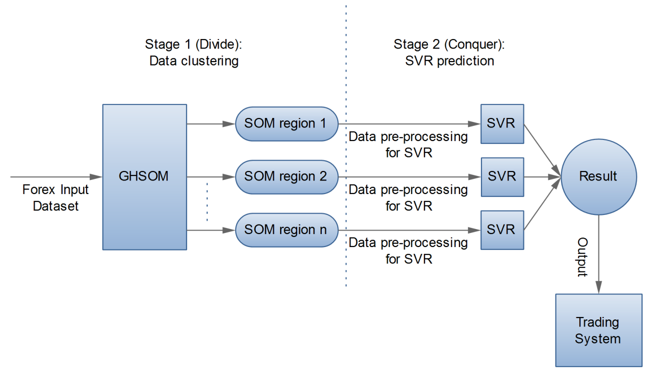

Fig. 4. The two-stage architecture.

In [17], a comparative study between both models was conducted, nevertheless a more detailed analysis should be performed by reducing the training and testing periods, since in currency market the most negotiated pairs tend to change their behavior in a short period of time. To deal with this issue, a sliding window-based approach is conducted in this work, along with the analysis of a variety of performance indicators, including popular indicators used by investors and statistical measurements. The sliding window method, presented on Fig. 5, makes it possible the discovery of new patterns throughout the time series history [25], that is, the training process is performed n times, where n corresponds to the number of sliding windows. New changes in the economic scenario can be identified by recent training phases. The use of sliding windows makes it possible to find recent frequent patterns in data streams [26] [27] [28].

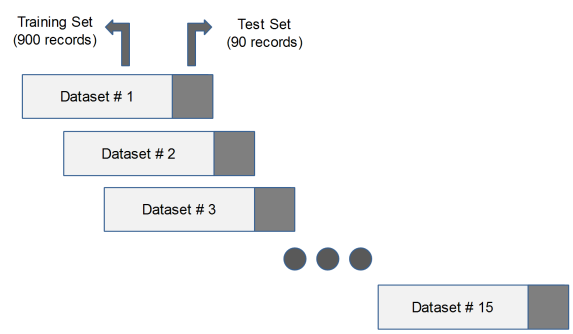

Fig. 5. Sliding Windows.

Although the experiments conducted on both models in [17] and [8] were performed on the two most negotiated currency pairs, the EUR/USD and GBP/USD, some pairs were investigated by only the GA-based model. In this paper, nine of the most popular currency pairs on foreign exchange were selected to enhance the trading systems validation with the aid of sliding windows. Details regarding both models (e.g., parameters, preprocessing) can be verified in [17] and [8].

IV. Experiments and comparison

In [17] and [8], multiple trading systems were developed combining some soft computing technique, such as GHSOM, SVR and Genetic Algorithms. The SVR+GHSOM obtained good results when tested against the data set described in [17], but didn't succeed when tested against the data set described in [8], in contrast with the GA-based model which presented good results. Since both papers were developed and tested against different data sets periods, due to the broker data availability, a more complete investigation should be conducted to select the best hybrid model and trading systems based on some performance indicators.

A. Data sets

As previously mentioned, the experiments were conducted on the most negotiated Forex currency pairs and are described on Table I comprehending an average period of 4.3 years. The time series are the EUR/USD, GBP/USD, USD/CHF, USD/JPY, AUD/USD, USD/CAD, EUR/CHF, EUR/JPY and the GBP/JPY. Besides that, two new data sets from stock market were included, the Dow Jones Industrial Average (DJIA) and the Bovespa (BVSP). The currency data set was obtained from OANDA (a financial institution dedicated to Forex trading and currency information services) [29] from a living account and the stock market indexes from Yahoo Finance [30]. All the data sets were split into 15 windows, with each window containing 900 price quotes for training and 90 quotes for testing, that is, 90% for training and the remaining for testing (see Fig. 5). After the first window execution, the next window is initiated removing the first 90 quotes from the previous window and aggregating the next 90 quotes for testing. Note that the testing period used in the first window is now used in the training process of the second window. That way, it is possible to comprise a greater testing period, since it was not possible to obtain a greater data set from OANDA.

TABLE I

Data set period comprising 15 sliding windows.

Data set	Period
EUR/USD (E/U)	2007-12-25 to 2012-03-26
GBP/USD (G/U)	2008-01-02 to 2012-04-12
USD/CHF (U/CHF)	2008-01-01 to 2012-04-13
USD/JPY (U/J)	2008-01-01 to 2012-04-15
AUD/USD (A/U)	2007-12-29 to 2012-03-31
USD/CAD (U/CAD)	2008-01-03 to 2012-04-18
EUR/CHF (E/C)	2008-01-08 to 2012-04-23
EUR/JPY (E/J)	2008-01-06 to 2012-04-17
GBP/JPY (G/J)	2008-01-05 to 2012-04-25
Dow Jones (DJIA)	2007-12-26 to 2012-04-25
Bovespa (BVSP)	2007-12-26 to 2012-04-25

The daily time frame (1D) was used to generate the data sets. It is important to notice that the 990 quotes within a window does not mean 990 contiguous days, since the financial market does not open some days, as in holidays and weekends. All data sets were scaled to the range [-0.9,0.9].

B. Trading systems

Trading systems (TS) were developed and details of implementation are described in [17] and [8]. In a nutshell, the SVR+GHSOM (SG) trading system open trades based on the hybrid model prediction for the close price of five days ahead. The order can be a SELL or a BUY operation, with the possibility to make profit in each direction. Orders are open all day, regardless if there are already open trades. All orders have an expiration of five days, that is, all orders are closed in the fifth day, regardless on the current trend.

The TSs for the GA-based model are composed by a series of rules involving technical indicators such as: (i) EMA (Exponential Moving Average), (ii) MACD (Moving Average Convergence / Divergence), (iii) RSI (Relative Strength Index) and (iv) the Stochastic Oscillator. The rules are detailed in [17]. A total of 15 TSs were developed by combining these technical indicators, as described in Table II. The parameters of each technical indicator were selected by genetic algorithms (GA). The order can be a BUY, a SELL or a NEUTRAL operation in case of no signals are emitted (matching rules).

TABLE II

GA-based trading systems.

TS	TS Rules
GA1	EMA3
GA2	MACD
GA3	RSI
GA4	STOCH
GA5	EMA3 + MACD
GA6	EMA3 + RSI
GA7	EMA3 + STOCH
GA8	MACD + RSI
GA9	MACD + STOCH
GA10	RSI + STOCH
GA11	EMA3 + MACD + RSI
GA12	EMA3 + MACD + STOCH
GA13	EMA3 + RSI + STOCH
GA14	MACD + RSI + STOCH
GA15	EMA3 + MACD + RSI + STOCH

In addition to these 16 TSs, another TS was developed using the same SVR+GHSOM model forecasting, though using the stop-and-reverse strategy, i.e., while the system is emitting a BUY signal, the order is maintained intact until a SELL or NEUTRAL signal is emitted, closing the current order and reverting the operation in case of a SELL signal or just staying NEUTRAL. The same process occurs in the case of a recurrent SELL signal output, reverting the order in the case of a BUY signal or just closing the trade and staying NEUTRAL. That specific TS was named SVR+GHSOM Stop and Reverse, or just "SGSR".

C. Performance indicators

In this paper we aim to compare the methods investigated from a financial perspective, therefore, we selected the performance indicators shown in Table III. The DS (direction success) presented in our previous paper is not used in this paper due to its irrelevance, demonstrated in [8]. It is important to remember that in the Forex market profits can be obtained in both movements, that is, in falling or rising. The ROI (return on investment) can be used as a measure of profitability, where the financial return is measured in percentage. The ROI can be compared to other types of investments and can be easily understood by market participants. The MD (maximum drawdown) shows, in percentage, the greatest loss during the period tested, that is, the maximum percentage of an account which could be lost after a series of losing trades in a period. Thus, it is an important measure to determine the risk involved in the investment.

In this paper, we extended the performance indicators by including the wilcoxon signed-rank test [31] and the profit factor (PF). The wilcoxon test is a distribution-free, non-parametric technique that does not require any underlying distributions in the data, and deals with the signs and ranks of the values and not with their magnitude (thus not influenced by outlier data points). It is one of the most commonly adopted tests in evaluating the predictive capabilities of two different models to see whether they are statistically significant difference between them [19]. The objective is to determine if the ROI obtained from the trading systems are part of a distribution with the population mean different from zero, that is, if the TSs are statistically profitable. The profit factor is a metric used by investors to measure the profit capability of a trading system. Values of PF above 2.0 are desirable [32], notwithstanding some investors suggests that TSs with PF above 1.5 are acceptable [33].

TABLE III

PERFORMANCE INDICATORS.

Metric	Calculation
Return on Investment	$ROI = 100 * \frac{(Gain - Investment)}{Investment}$
Maximum Drawdown	$MD = \Vert \frac{balance\ valley - balance\ peak}{balancepeak} \Vert$
Profit Factor	$PF = \frac{Gross\ Profit}{Gross\ Loss}$

D. Trading configuration

Different levels of risk can be assumed in the operations in the Forex market. The choice of the type of trader one wants to operate, that is, more conservative or more aggressive, depends on the amount of lots invested in the transaction. In the experiments, we used one mini lot. Each mini lot corresponds to a certain value in US$ per pipette. This value can be obtained from [34]. The term "pip" is an acronym for price interest point and is used in the Forex market to determine the minimum price change in floating foreign exchange rates. For example, a change in a quoted price from 1.6998 to 1.6999 is equal to one pip. A pipette is a fractional pip, which is equal to 1/10th of a pip. It is represented by the 5th decimal digit. In the experiments, we consider a fictitious account with an initial deposit of US$ 10,000.00.

E. Results

Initially, the data sets were tested against a simple buy-and-hold strategy to justify the use of sophisticated models such as the proposed in this paper (soft computing). The buy-and-hold strategy is commonly used by conservative investors and it is based on the assumption that the economy is in an ascending trend. Basically, the trader buy some assets and holds the position for a long period of time (usually for years). In the experiments, the buy-and-hold strategy buys each asset in the first quote and closes the operation in the last quote of each data set. The results are summarized in Table IV, and it can be observed that for only two data sets, the ROI were positive (the highlighted results). Also, the MD of DJIA and BVSP presented values above 50%, which means that half of the initial deposit were lost during the full data set period, yet recovered in the end. Trading systems with high MDs are inadvisable due to their high risks.

TABLE IV

BUY-AND-HOLD STRATEGY RESULTS.

Data set	ROI	MD
EUR/USD	-7.36	25.614
GBP/USD	-19.62	32.714
USD/CHF	-18.68	40.977
USD/JPY	-27.62	32.179
AUD/USD	17.32	37.777
USD/CAD	0.25	27.285
EUR/CHF	-26.59	36.912
EUR/JPY	-33.41	42.671
GBP/JPY	-38.85	45.509
DJIA	-3.4	51.09
BVSP	-3.95	51.39

Table V lists the ROI performance of each trading system applied against each currency pair, with the positive values highlighted. It can be observed that multiple TSs have outperformed the buy-and-hold strategy of Table IV and that the SVR+GHSOM with Stop and Reversal (SGSR) strategy showed to be the best TS when considering the ROI only.

It can be noticed in Table V that the values returned when tested against the USD/CAD are far superior when used the GA-based strategy, with most of the TSs showing positive values. Also, for the SG model, the ROI obtained presented the best performance in the ROI benchmarking. Thus, it is important to present more detailed results for such TSs and currency pairs. In Table VI, all trading systems were used combined to open orders for each data set. Notice that it is not a advisable to use all TSs, since the ROI and MD for some of them eliminated the whole initial deposit. In the other hand, for the USD/CAD and EUR/JPY, the results showed results of 258.65% and 162.19% respectively for ROI, but with a high risk due to the high MD. In Table VII, the same is performed, but using the SG and SGSR TSs. Although for some data sets the results showed to be positive, the MD for the SG TSs achieved high values, differently from the SGSR TSs which achieved acceptable values of MD.

It is also possible to apply the TSs to the nine currency pairs, without any portfolio management. That way, the TSs

TABLE V

ROI RESULTS PER TRADING SYSTEM AND CURRENCY PAIR.

TS	E/U	G/U	U/CHF	U/J	A/U	U/CAD	E/C	E/J	G/J
GA1	-36.85	-26.42	-60.42	-1.73	-16.53	5.81	-26.54	-62.96	-74.64
GA2	17.62	5.92	-24.12	10.16	-15.56	16.38	-17.56	38.14	6.9
GA3	-23.78	-42.53	-21.69	40.4	-14.8	30.97	-14.71	27.68	-21.71
GA4	22.89	-6.68	14.39	-11.54	-27.72	22.99	-5.64	-0.17	4.93
GA5	-15.21	-37.7	-37.3	-14.95	-10.78	16.18	-35.16	19.11	54.42
GA6	-55.09	-19.59	-27	7.49	-18.21	-0.31	-18.49	-7.74	72.75
GA7	14.66	1.2	-6.74	-2.04	-20.35	13.25	-12.17	52.64	-48.31
GA8	-12.45	-7.98	-26.95	-13.4	-32.29	19.32	-15.64	52.85	-21.89
GA9	6.42	-6.17	7.28	-8.02	-16.26	17.63	-4.19	7.04	51.06
GA10	-8.55	-24.28	12.04	4.74	-32.04	33.37	-5.22	27.6	1.6
GA11	-28.73	-39.1	-33.44	-3.37	-15.79	21.41	-14.49	-37.04	46.84
GA12	7.57	6.61	-21.54	-21.38	-4.04	22.57	-3.98	15.16	-14.65
GA13	-6.7	2.91	-5.14	-7.11	-2.17	9.88	-3.91	6.34	-13.56
GA14	4.06	-12.95	5.4	-13.68	-1.03	21.96	2.42	24.9	-14.71
GA15	8.58	-6.32	6.2	11.95	-20.11	7.25	-3.49	-1.34	-18.92
SG	-100	-50.35	-34.4	-71.49	84.17	-100	13.24	66.1	248.24
SGSR	-5.4	-26.15	-5.61	-6.65	49.52	2.61	19.29	51.27	40.23

TABLE VI

ROI AND MD OBTAINED WHEN USED ALL GA-BASED TSS AGAINST EACH DATA SET.

Data set	ROI	MD
EUR/USD	-100	100
GBP/USD	-100	100
USD/CHF	-100	100
USD/JPY	-100	100
AUD/USD	-100	100
USD/CAD	258.65	53.07
EUR/CHF	-100	100
EUR/JPY	162.19	88.77
GBP/JPY	-100	100

TABLE VII

ROI AND MD OBTAINED WHEN USED ALL SVR+GHSOM TSS AGAINST EACH DATA SET.

Data set	ROI SG	MD SG	ROI SGSR	MD SGSR
EUR/USD	-100	100	-5.4	25.87
GBP/USD	-50.35	94.08	-26.15	30.83
USD/CHF	-34.4	84.43	-5.61	30.29
USD/JPY	-71.49	84.37	-6.65	19.61
AUD/USD	84.17	75.93	49.52	11.59
USD/CAD	-100	100	2.61	17.76
EUR/CHF	13.24	44.8	19.29	24.13
EUR/JPY	66.1	72.36	51.27	25.99
GBP/JPY	248.24	51.51	40.23	34.14

can be evaluated individually, regardless of the currency pair. As we can see from the results in Table VIII, the SGSR outperformed the SG model and the GA-based trading systems, maintaining the low rate of MD, specially when compared to the MD of a conservative strategy as the buy-and-hold (Table IV). Again, the highlighted results shows the positive results.

Yet, performance indicators such as ROI and MD are not the only metrics used in financial markets. The profit factor (PF) performance metric relates the amount of profit per unit of risk, with values greater than one indicating a profitable system. It is a widely used metric by real traders [33]. The results of PF obtained in the experiments are detailed in Table IX with the recommended TSs presented by highlighted cells due to their values superior to 1.5 [33]. Notice that the number of acceptable TSs have reduced when compared to the ROI results of Table V. Although some results in Table V presented high values of ROI, e.g., the SG TS when tested against the GBP/JPY currency pair, the PF achieved poor values. Table X summarizes the best trading systems based on the PF metric for each currency pair. Notice that the best TSs were GA-based ones, and that no SVR+GHSOM model have succeeded. We give emphasis to the results obtained by the GA10 when tested against the USD/CAD currency pair, achieving PF of 3.62. Although its ROI value reached a rate of 33.37% (see Table V), it could be increased by raising the number of lots invested.

TABLE VIII

ROI AND MD PERFORMANCE WHEN APPLIED EACH TRADING SYSTEM TO ALL DATA SETS COMBINED.

TS	ROI	MD
GA1	100	-100
GA2	37.89	71.45
GA3	-100	100
GA4	13.44	58.4
GA5	-100	100
GA6	-66.2	99.88
GA7	-7.84	90.44
GA8	-58.43	85.32
GA9	54.78	26.18
GA10	9.26	53.29
GA11	-100	100
GA12	-13.68	61.89
GA13	-19.46	75.04
GA14	16.38	47.03
GA15	-16.2	70.07
SG	-100	100
SGSR	119.12	32.673

In the task of searching for the best model, some trading systems succeeded when using the SVG+GHSOM model

TABLE IX

PF RESULTS PER TRADING SYSTEM AND CURRENCY PAIR.

TS	E/U	G/U	U/CHF	U/J	A/U	U/CAD	E/C	E/J	G/J
GA1	0.72	0.7	0.51	0.97	0.8	1.09	0.64	0.6	0.66
GA2	1.18	1.03	0.81	1.07	0.91	1.14	0.84	1.12	1.02
GA3	0.66	0.63	0.65	1.59	0.85	1.7	0.64	1.16	0.92
GA4	1.42	0.9	1.63	0.76	0.48	1.63	0.87	1	1.04
GA5	0.83	0.66	0.52	0.83	0.88	1.26	0.47	1.16	1.34
GA6	0.3	0.71	0.54	1.18	0.65	0.99	0.53	0.91	1.85
GA7	1.75	1.02	0.82	0.95	0.6	1.28	0.73	1.66	0.68
GA8	0.79	0.93	0.67	0.81	0.67	1.29	0.75	1.58	0.89
GA9	1.18	0.91	1.26	0.76	0.55	1.4	0.89	1.15	1.75
GA10	0.83	0.7	1.57	1.13	0.44	3.62	0.86	1.41	1.02
GA11	0.46	0.35	0.48	0.91	0.63	1.52	0.71	0.59	1.43
GA12	1.23	1.18	0.62	0.54	0.85	1.72	0.87	1.26	0.81
GA13	0.71	1.05	0.83	0.75	0.91	1.39	0.83	1.11	0.88
GA14	1.22	0.66	1.15	0.59	0.96	1.79	1.17	1.56	0.81
GA15	1.51	0.84	1.26	1.82	0.38	1.3	0.84	0.96	0.76
SG	0.82	0.96	0.98	0.92	1.09	0.8	1.02	1.04	1.13
SGSR	0.98	0.92	0.98	0.97	1.24	1.01	1.12	1.13	1.09

TABLE X

BEST TRADING SYSTEMS WHEN EVALUATED AGAINST THE PROFIT FACTOR PERFORMANCE INDICATOR.

Data set	TS
EUR/USD	GA7
GBP/USD	-
USD/CHF	GA4
USD/JPY	GA15
AUD/USD	-
USD/CAD	GA10
EUR/CHF	-
EUR/JPY	GA7
GBP/JPY	GA6

TABLE XI

BEST TRADING SYSTEMS WHEN EVALUATED BY THE WILCOXON TEST.

Data set	TS (p-value)
EUR/USD	-
GBP/USD	-
USD/CHF	-
USD/JPY	-
AUD/USD	-
USD/CAD	GA3 (0,00021049) GA10 (0,0015431)
EUR/CHF	-
EUR/JPY	GA7 (0,007712) GA10 (0,026779)
GBP/JPY	GA6 (0,042066)

and others when trading with technical indicators optimized by genetic algorithms (GA-based), depending on the performance metric. Since the ROI metric differs from the PF metric, we also analyze statistical results by applying the wilcoxon test to the trading systems. The objective is to verify whether the trading systems are statistically profitable. It can be observed from Table XI that only some of the GA-based trading systems have passed the test for specific currency pairs, reducing even more the number of acceptable trading systems. Also, the GA6 TS on GBP/JPY obtained the highest p-value.

V. CONCLUSIONS

The work presented a comparison between multiple trading systems based on soft computing techniques, such as SVR, GHSOM and Genetic Algorithms. They were analyzed by commonly used performance indicators in financial market, such as the return on investment, maximum drawdown, profit factor and by a statistical test, the wilcoxon signed-rank test. The results showed that the technical indicators strategy, with their parameters optimized by genetic algorithms, in this work named GA-based trading systems, outperformed the SVR+GHSOM model when considered all performance metrics. Also, both models showed to be superior to a simple buy-and-hold strategy and proved the opposite from what the efficient market hypothesis proposes. Therefore, one of the most popular tools known by financial market players, the technical indicators, showed to be successful in the task of making consistent profit results when the right parameters are chosen. To tackle with the problem of choosing the best parameters, genetic algorithms can be a powerful soft computing technique.

Future work may be performed in order to improve the results. To this end, the trading system could be modified to manage the risk involved in the operation, adjusting the number of lots for each transaction and using automated closing operations, such as the use of trailing stops, stop loss and take profit, thus, using the concepts of money and risk management.

References

[1] E. F. Fama, *Random walks in stock-market prices (University of Chicago. Graduate School of Business Selected papers)*. Graduate School of Business, University of Chicago, 1965.

[2] L. Kilian and M. P. Taylor, "Why is it so difficult to beat the random walk forecast of exchange rates?" *Journal of International Economics*, vol. 60, no. 1, pp. 85–107, 2003.

[3] G. S. Atsalakis and K. P. Valavanis, "Surveying stock market forecasting techniques Part II: Soft computing methods," *Expert Systems with Applications*, vol. 36, no. 3, pp. 5932–5941, Apr. 2009.

[4] P. Paranjape-Voditel and U. Deshpande, "A stock market portfolio recommender system based on association rule mining," *Applied Soft Computing*, vol. 13, no. 2, pp. 1055–1063, 2013.

[5] V. Kodogiannis and A. Lolis, "Forecasting Financial Time Series using Neural Network and Fuzzy System-based Techniques," *Neural Computing & Applications*, vol. 11, no. 2, pp. 90–102, Oct. 2002.

[6] A.-S. Chen and M. T. Leung, "Regression neural network for error correction in foreign exchange forecasting and trading," *Computers & Operations Research*, vol. 31, no. 7, pp. 1049–1068, 2004.

[7] L. Cao, "Support vector machines experts for time series forecasting," *Neurocomputing*, vol. 51, pp. 321–339, 2003.

[8] R. F. B. de Brito and A. L. I. Oliveira, "A foreign exchange market trading system by combining GHSOM and SVR," *The 2012 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7, Jun. 2012.

[9] BIS, "Triennial Central Bank Survey of foreign exchange and derivatives market activity in 2013," 2013. [Online]. Available: http://www.bis.org/publ/rpfx13.htm

[10] S. Theodoridis and K. Koutroumbas, "Pattern Recognition, Fourth Edition," Sep. 2008.

[11] M. Bramer, N. OConnor, and M. G. Madden, "A neural network approach to predicting stock exchange movements using external factors," *Knowledge-Based Systems*, vol. 19, no. 5, pp. 371–378, 2006.

[12] L. Yu, K. Lai, and S. Wang, "Multistage RBF neural network ensemble learning for exchange rates forecasting," *Neurocomputing*, vol. 71, no. 16-18, pp. 3295–3302, Oct. 2008.

[13] S.-H. Hsu, J. P.-A. Hsieh, T.-C. Chih, and K.-C. Hsu, "A two-stage architecture for stock price forecasting by integrating self-organizing map and support vector regression," *Expert Systems with Applications*, vol. 36, no. 4, pp. 7947–7951, May 2009.

[14] H. Ni and H. Yin, "Exchange rate prediction using hybrid neural networks and trading indicators," *Neurocomputing*, vol. 72, no. 13-15, pp. 2815–2823, Aug. 2009.

[15] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489–501, Dec. 2006.

[16] P. Vitale, "Foreign exchange intervention: how to signal policy objectives and stabilise the economy," *Journal of Monetary Economics*, vol. 50, no. 4, pp. 841–870, 2003.

[17] R. F. de Brito and A. L. Oliveira, "Comparative Study of FOREX Trading Systems Built with SVR+GHSOM and Genetic Algorithms Optimization of Technical Indicators," in *2012 IEEE 24th International Conference on Tools with Artificial Intelligence*, vol. 1. IEEE, Nov. 2012, pp. 351–358.

[18] A. S. Vladimir Vapnik, Steven E. Golowich, "Support Vector Method for Function Approximation, Regression Estimation, and Signal Processing."

[19] C.-J. Lu, T.-S. Lee, and C.-C. Chiu, "Financial time series forecasting using independent component analysis and support vector regression," *Decision Support Systems*, vol. 47, no. 2, pp. 115–125, 2009.

[20] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, Aug. 2004.

[21] A. Rauber, D. Merkl, M. Dittenbach, O. . Y. E. Øø, and O. O. E. O. O. O. Óò, "The growing hierarchical self-organizing map: exploratory analysis of high-dimensional data." *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, vol. 13, no. 6, pp. 1331–41, Jan. 2002.

[22] T. Kohonen, "The self-organizing map," *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464–1480, 1990.

[23] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Professional, 1989.

[24] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, and G. E. Hinton, "Adaptive Mixtures of Local Experts," *Neural Computation*, vol. 3, no. 1, pp. 79–87, Feb. 1991.

[25] S. K. Tanbeer, C. F. Ahmed, B.-S. Jeong, and Y.-K. Lee, "Sliding window-based frequent pattern mining over data streams," *Information Sciences*, vol. 179, no. 22, pp. 3843–3865, 2009.

[26] J. H. Chang and W. S. Lee, "Finding recent frequent itemsets adaptively over online data streams," in *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '03*. New York, New York, USA: ACM Press, Aug. 2003, p. 487.

[27] C. Leung and Q. Khan, "DSTree: A Tree Structure for the Mining of Frequent Sets from Data Streams," in *Sixth International Conference on Data Mining (ICDM'06)*. IEEE, Dec. 2006, pp. 928–932.

[28] J. Li, D. Maier, K. Tufte, V. Papadimos, and P. A. Tucker, "No pane, no gain," *ACM SIGMOD Record*, vol. 34, no. 1, p. 39, Mar. 2005.

[29] OANDA, "The Currency Site, Foreign Currency Data." [Online]. Available: http://www.oanda.com

[30] Yahoo, "Yahoo! Finance," 2011. [Online]. Available: http://finance.yahoo.com

[31] R. F. Woolson, "Wilcoxon signed-rank test," *Wiley Encyclopedia of Clinical Trials*, pp. 1–3, 2008.

[32] B. Vanstone and G. Finnie, "An empirical methodology for developing stockmarket trading systems using artificial neural networks," *Expert Systems with Applications*, vol. 36, no. 3, pp. 6668–6680, Apr. 2009.

[33] J. Katz and D. McCormick, *The Encyclopedia of Trading Strategies*, 2000.

[34] Forex Trading Profit Calculator, "Forex Trading Profit Calculator," 2014. [Online]. Available: http://fxtrade.oanda.com/lang/pt/analysis/profit-calculator/