

Modeling and Trading FTSE100 Index Using a Novel Sliding Window Approach Which Combines Adaptive Differential Evolution and Support Vector Regression

Konstantinos Theofilatos, Andreas Karathanasopoulos, Peter Middleton,
Efstratios Georgopoulos, Spiros Likothanassis

► **To cite this version:**

Konstantinos Theofilatos, Andreas Karathanasopoulos, Peter Middleton, Efstratios Georgopoulos, Spiros Likothanassis. Modeling and Trading FTSE100 Index Using a Novel Sliding Window Approach Which Combines Adaptive Differential Evolution and Support Vector Regression. Harris Papadopoulos; Andreas S. Andreou; Lazaros Iliadis; Ilias Maglogiannis. 9th Artificial Intelligence Applications and Innovations (AIAI), Sep 2013, Paphos, Greece. Springer, IFIP Advances in Information and Communication Technology, AICT-412, pp.486-496, 2013, Artificial Intelligence Applications and Innovations. <10.1007/978-3-642-41142-7_49>. <hal-01459685>

HAL Id: hal-01459685

<https://hal.inria.fr/hal-01459685>

Submitted on 7 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Modeling and Trading FTSE100 Index Using a Novel Sliding Window Approach Which Combines Adaptive Differential Evolution and Support Vector Regression

Konstantinos Theofilatos¹, Andreas Karathanasopoulos², Peter Middleton³, Efstratios Georgopoulos⁴, Spiros Likothanassis¹

¹Department of Computer Engineering and Informatics, University of Patras, Greece
{theofilk@ceid.upatras.gr, likothan@ceid.upatras.gr}

²Business School, University of East London, Unighted Kingdom
{a.karathanasopoulos@uel.ac.uk}

³Management School, University of Liverpool, Unighted Kindgom
{peter.william.middleton@gmail.com}

⁴Technological Educational Institute of Kalamata, Greece
{sfg@teikal.gr}

Abstract. The motivation for this paper is to introduce a novel short term trading strategy using a machine learning based methodology to model the FTSE100 index. The proposed trading strategy deploys a sliding window approach to modeling using a combination of Differential Evolution and Support Vector Regressions. These models are tasked with forecasting and trading daily movements of the FTSE100 index. To test the efficiency of our proposed method, it is benchmarked against two simple trading strategies (Buy and Hold and Naïve Strategy) and two modern machine learning methods. The experimental results indicate that the proposed method outperforms all other examined models in terms of statistical accuracy and profitability. As a result, this hybrid approach is established as a credible and worth trading strategy when applied to time series analysis.

Keywords: Differential Evolution, Support Vector Regression, Confirmation Filters, FTSE100, Daily Trading

1 Introduction

Modeling and trading financial indices remains a very interesting but challenging topic in econometrics. Forecasting financial time series is a difficult task because of their complexity and their dynamic and noisy nature. All traditional linear methods and even more sophisticated non-linear machine learning models have failed to capture the complexity and the nonlinearities that exist in financial time series. This is particularly the case during times of uncertainty such as during the credit crisis in 2008. More robust and intuitive models have since been researched and applied to finan-

cial times series in order to overcome these inefficiencies associated with previous models [1].

Non-linear machine learning models have three main limitations. The first disadvantage is that most are calibrated to search for a global optimal estimator which in most cases does not exist due to the dynamic nature of financial timeseries. Another drawback is that the algorithms which are used for modeling financial time-series have a lot of parameters which need to be optimized and if this procedure is not performed correctly then the extracted prediction models will produce unsatisfactory results due to the data-snooping effect. Another mistake which is commonly made by practitioners is that, most of the time the training of a prediction model is performed separately from the construction of a trading strategy and thus the overall performance is reduced.

The purpose of this paper is to present a novel methodology which is capable of overcoming the aforementioned limitations. This methodology is based on a sliding window approach for the one day ahead prediction of the FTSE100 returns. To forecast every single return, the proposed trains a machine learning model using a sliding window of contemporary limited historical data. Thus the proposed method searches for the optimal predictor for each day. The machine learning model which was applied is an adaptive hybrid combination of Differential Evolution and the nu-Support Vector Regression (SVR) algorithm [2]. The adaptive Differential Evolution [3] (DE) was used for selecting the optimal feature subset, optimizing the parameters of nu-SVRs and at the same time optimizing the confirmation threshold for the confirmation filters which are used as parameters for our trading strategy. Moreover, a specialized fitness function was designed and used for the evaluation step of the adaptive Differential Evolution Algorithm which takes into account metrics for both statistical accuracy and trading performance.

The performance of the proposed methodology is compared with two traditional linear models and two non-linear machine learning approaches. The empirical analysis reveals that our proposed methodology clearly outperforms the other existing models ranking the highest across all of the examined metrics.

The novelty of the proposed approach is twofold. The first contribution lies in the application of a sliding window training method and the second offers an original adaptive hybrid machine learning methodology. On review of existing literature, the latter is unique and original as this is the first time that an adaptive Differential Evolution algorithm and nu-SVRs are combined into one model used for forecasting financial time series. Moreover, our proposed machine learning method is not only a simple combination of existing methods. The Differential Evolution algorithm optimizes not only the feature subset and the parameters of nu-SVRs but also a parameter of our trading strategy. When this was combined with the usage of a new fitness function specialized in trading tasks, it enabled our proposed method to fill the gap between financial forecasting and trading.

The rest of the paper is organized as follows: Section 2 describes the dataset used for our experiments. In Section 3 the proposed methodology was described in detail and in Section 4 the benchmark models are described and the experimental results are

presented. Finally, Section 5 presents concluding remarks and some interesting future directions for research.

2 Related Financial Data

The FTSE 100 index is a weighted according to market capitalization which currently comprises of 101 large cap constituents listed on the London Stock Exchange. For the purpose of our trading simulation the iShares FTSE100 exchange traded funds traded to capture daily movements of the FTSE100 index. Positions are initiated on the open of a trading day and closed at or around 16:30 GMT. The cash settlement of this index is simply determined by calculating the difference between the traded price on the open and the closing price of the index on the . When the model forecasts a negative return then a short position (sale) is assumed on the open and when it forecasts a positive return a long position (purchase) is taken. Profit / loss is realised on a daily basis and positions are not held overnight.

The FTSE 100 daily time series is non-normal (Jarque-Bera statistics confirms this at the 99% confidence interval), containing slight negative skewness and relatively high kurtosis. Arithmetic returns were used to calculate daily returns and they are estimated using equation 1. Given the price level P_1, P_2, \dots, P_t , the arithmetic return at time t is formed by:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \quad (1)$$

In table 1 we present the full dataset used.

Table 1.Total dataset

Name of Period	Trading Days	Beginning	End
Total Dataset	1260	1 January 2008	31 December 2012
In sample Dataset	756	1 January 2008	31 December 2010
Out of Sample Set	504	1 January 2011	31 December 2012

As inputs to our algorithms, we selected a combination of autoregressive inputs, moving average time series of the FTSE100 returns, a FTSE100 volume time series, daily highs and lows of FTSE100 index, the VIX realized volatility index, and two metric time series which capture the aggregate advancing and declining volumes as a daily percentage as provided by FACTSET(2013) The inputs used in our modeling process are described in table 2.

Table 2.Explanatory Variables

Number	Variable	Lag
1	Daily High of FTSE100 all share return	1
2	Daily Low of FTSE100 all share return	1
3	FTSE100 all share return	1

4	FTSE100 all share return	2
5	FTSE100 all share return	3
6	FTSE100 all share return	4
7	FTSE100 all share return	5
8	FTSE100 all share return	6
9	FTSE100 all share return	7
10	5 day Moving Average of FTSE100 all share returns	1
11	15 day Moving Average of FTSE100 all share returns	1
12	30 day Moving Average of FTSE100 all share returns	1
13	FTSE100 daily volume	1
14	5 day Moving Average of FTSE100 daily volume	1
15	15 day Moving Average of FTSE100 daily volume	1
16	30 day Moving Average of FTSE100 daily volume	1
17	VIX Realized Volatility Index	1
18	5-day Moving Average of the VIX Index	1
19	15-day Moving Average of the VIX Index	1
20	30-day Moving Average of the VIX Index	1
21	Aggregate Advancing Volume Percentage metric	1
22	5-day Moving Average of Aggregate Advancing Volume Percentage metric	1
23	15-day Moving Average of Aggregate Advancing Volume Percentage metric	1
24	30-day Moving Average of Aggregate Advancing Volume Percentage metric	1
25	Aggregate Declining Volume Percentage metric	1
26	5-day Moving Average of Aggregate Declining Volume Percentage metric	1
27	15-day Moving Average of Aggregate Declining Volume Percentage metric	1
28	30-day Moving Average of Aggregate Declining Volume Percentage metric	1

In contrast to other non-linear approaches we have incorporated the VIX index to capture volatility. It is believe that this will be of particular benefit during times of higher volatility such as period of crisis. Moreover, some advanced volume metrics where studied as they are considered significant for capturing the markets liquidity. For instance, the advancing volume metric provides a sum of daily volume for those companies with advancing prices during a particular day. The declining volume metric provides a sum of daily volume for those companies with a declining price during a specific day.

Advancing volume percentage is defined as:

((Sum of Volume (Day t)) for all companies

where $\text{Price (Day } t) > \text{Price (Day } t-1) / \text{Sum of Volume for all constituents}$ (2)

Declining volume percentage is defined as:

((Sum of Volume (Day t)) for all companies

where $\text{Price (Day } t) < \text{Price (Day } t-1) / \text{Sum of Volume for all constituents}$ (3)

These calculations use daily prices and volumes for each constituent as provided by the FactSet Pricing Database (2013).

Finally, the set of explanatory variables were normalized in the interval of [-1,1] to avoid overrating inputs which takes higher absolute values.

3 Proposed Method

The proposed methodology is a sliding window approach which is designed to perform daily forecasting and trading. To forecast future values of the FTSE100 returns it trains a machine learning model using a window of a specific amount (name sliding window size) of historical prices for the examined inputs. By this way it outperforms other classical methodologies as it uses most recent data to update prediction models when forecasting next day returns.

The machine learning methodology which was used is based on the nu-Support Vector Regression predictors [4].

Support vector machines (SVM) are a group of supervised learning methods that can be applied to classification or regression. SVMs represent an extension to nonlinear models of the generalized algorithm developed by Vapnik [2]. They have been developed into a very active research area and have already been applied to many scientific problems. For instance, SVM have already been applied in many prediction and classification problems in finance and economics [5, 6] although they are still far from mainstream and the few financial applications so far have only been published in statistical learning and artificial intelligence journals.

SVM models were originally defined for the classification of linearly separable classes of objects. For any particular linear separable set of two-class objects SVM are able to find the optimal hyperplanes that separates them providing the bigger margin area between the two hyperplanes. However, SVM can also be used to separate classes that cannot be separated with a linear classifier. In such cases, the coordinates of the objects are mapped into a feature space using nonlinear functions. Every object is projected in a high-dimensional space feature space in which the two classes can be separated with a linear classifier.

In the task of forecasting financial indexes SVMs can be used for forecasting the directional movement of the examined index. However, these forecasts are not easily transformed to an effective trading strategies as the application of confirmation filters is not straightforward. The introduction of the ϵ -sensitive loss function by Vapnik(1995) [2] improves the Support Vector Regression approach and provides a robust technique for solving difficult regression problems. An improvement of the classical e-SVM is the more flexible nu-SVR [4] which is adopted for this particular study. When training nu-SVRs, the features which should be used as inputs should be selected carefully to avoid the curse of dimensionality. Moreover, the parameters C (regularization parameter), gamma (Radial Basis Function parameter) and nu should be optimized.

In more recent years a variety of meta-heuristic optimization problems have been proposed such as Genetic Algorithms [7] and Differential Evolution [3]. The most important problem they encounter is the fact that their own parameters should be calibrated in such a manner to enable them to effectively explore the search space while at the same time performing effective local searches.

The DE algorithm is currently one of the most powerful and promising stochastic real parameter optimization algorithms [3]. It belongs to the wider family of evolutionary algorithms as its operation is based on an iterative application of selection, muta-

tion and crossover operators. DE is mainly based on a specific mutation operator. This operator randomly selects individuals from the population based on scaled differences between other randomly selected and distinct population members.

The representation schema which DE uses is the continuous gene representation. Thus, candidate solutions are represented as strings of continuous variables comprising of feature and parameter variables. For every candidate input of our model a feature variable is added to the representation of a member of DE's population. These genes take values from 0 to 2 with values higher than 1 indicating that this feature should be used as input to the SVR model. Four parameter variables are included in the representation of a candidate solution of DEs in the proposed methodology. These are C (values in the interval [0,1024]), gamma (values in the interval [0, 1024]), nu (values in the interval [0,1]) and the optimal confirmation threshold (values in the interval [0, 0.01]).

The mutation operator which was selected for our proposed methodology selects for every population member X_i three random distinct members of the population ($X_{1,i}$, $X_{2,i}$, $X_{3,i}$) and produces a donor vector using the equation (3):

$$V_i = X_{1,i} + F * (X_{2,i} - X_{3,i}) \quad (3)$$

where F is called mutation scale factor.

The crossover operator applied was the binomial one. This operator combines every member of the population x_i with its corresponding donor vector V_i to produce the trial vector U_i using the equation (4).

$$U_i(j) = \begin{cases} V_i(j) & , \text{ if } (rand_{i,j}[0,1] \leq Cr) \\ X_i(j) & , \text{ otherwise} \end{cases} \quad (4)$$

Where: $rand_{i,j}[0,1]$ is a uniformly distributed random number and Cr is the crossover rate.

Next the selection operator is applied. Every trial vector U_i is evaluated and if it suppresses the corresponding member of the population X_i it takes its position in the population. To evaluate the candidate solutions of the proposed methodology we used the following specialized fitness function:

$$\text{Fitness Function} = \text{Correct_rate} - 1000 * \text{MSE} + \text{Annualized_Return} \quad (5)$$

Where: Correct_Rate is the percentage of correct predictions, Annualized Return is the annualized return of the extracted trading strategy when taking into account the transaction costs and applying the corresponding conformation filter. MSE is the Mean Square Error (we multiplied it with 1000 to normalize its values in the magnitude of the ones of Correct Rate and Annualized Return). The aforementioned fitness function enables us to achieve high statistical metrics in our forecasting task while at the same time extracting optimized profitable trading strategies.

The termination criterion which was used was a combination of the maximum number of generations to be reached with a convergence criterion. The convergence criterion terminates the algorithm when the performance of the best member of the population is less than 5% away than the mean performance of the population.

The most important control parameters of a DE algorithm are the mutation scale factor F and the crossover rate Cr. Parameter F controls the size of the differentiation quantity which is going to be applied to a candidate solution from the mutation opera-

tor. Parameter Cr determines the number of genes which are expected to change in a population member. Many approaches have been developed to control these parameters during the evolutionary process of the DE algorithm [3]. In our adaptive DE version, we deployed one of the most recent promising approaches [8]. This approach randomly selects values during every iteration. For the F parameter selected from a uniform distribution with mean value 0.5 and standard deviation 0.3 and a random value for the parameter Cr from a uniform distribution with mean value Crm and standard deviation 0.1. Crm is initially set to 0.5. The Crm is replaced during the evolutionary process with values that have generated successful trial vectors. As a result, this approach replaces the sensitive user defined parameters F and Cr with less sensitive parameters like their mean values and their standard deviation.

The parameters of our method which needed to be optimized where the population size, the maximum number of generations to reach and the sliding window size. These parameters were optimized with thorough experimentation using only the in-sample dataset and examining the values of the fitness function described in equation 5. The optimal values which were found and used where the following: population size = 30, maximum number of generations to reached = 200, sliding window size = 252.

4 Comparative Experimental Results

4.1 Benchmark Models

The simple benchmark trading strategies which were used for comparison reasons were Naive Strategy and Buy and Hold Strategy. Buy and Hold is a simple strategy, where traders buy the index (asset) at the beginning of the review period and sell it at the end of a predetermined period or once a price target has been reached. The naive strategy simply takes the most recent period change as the best prediction of the future change. Then it goes long if the forecasting is positive and short if it is negative.

From the machine learning benchmark models the ones which were used for comparative reasons, were the hybrid combination of Genetic Algorithms and Artificial Neural Networks proposed in [9] and the hybrid methodology combining Genetic Algorithms and Support Vector Machines [10]

Neural networks exist in several forms in the literature. The most popular architecture is the Multilayer Perceptrons (MLP). Their most important problem is that they require a feature selection step and their parameters are hard to be optimized. For these reasons outline by [9] Genetic Algorithms [7] were used to select suitable inputs. The Levenberg–Marquardt back propagation algorithm [11] is employed during the training procedure which adapts the learning rate parameter during this procedure.

In the second machine learning model which was used for comparative reasons [reference] the authors proposed a hybrid GA and SVM model which is designed to overcome some of the limitations of Artificial Neural Networks and simple SVMs. More specifically in this methodology, a genetic algorithm is used to optimize the SVM parameters and on parallel to find the optimal feature subset. Moreover, this

approach used a problem specific fitness function which is believed to produce more profitable prediction models.

4.2 Trading Performance

In this section we present the results of each model from trading the FTSE100 index. The trading performance of all the models considered in the out-of-sample subset is presented in table 3. The trading strategy for the GA-MLP and the GA-SVM is simple and identical for both of them: go long on the open when the forecast return is above zero and go short when the forecast return is below zero. Each position is held for only one trading day. The trading strategy for our model is identical with the previous result except it is more selective when trading as we apply a confirmation filter. The confirmation filter restricts the model for trading when the forecasted value is less than the optimal confirmation threshold for its sliding window period. Because non-linear methodologies are stochastic by nature a single forecast is not sufficient enough to represent a credible result. For this reason, an average of ten estimations where executed to represent each model as presented in table 3.

Table 3.Trading Results

	Naive Strategy	Buy and Hold Strategy	GA-MLP	GA-SVM	Proposed Method
<i>Annualized Return (excluding costs)</i>	6,68%	-1,64%	11,61%	16,27%	23,56%
<i>Positions Taken (annualized)</i>	64	0,5	56	60	47
<i>Transaction Costs</i>	6,41%	0,05%	5,66%	6,01%	4,7%
<i>Annualized Return (including costs)</i>	0,27%	-1,69%	5,94%	10,26%	18,86%
<i>Annualized Volatility</i>	17,98%	18,04%	18,04%	18,03%	13,69%
<i>Information Ratio (including costs)</i>	0,01	-0,09	0,33	0,57	1,38
<i>Maximum Drawdown</i>	-31,29%	-22,42%	-30,15%	-30,15%	-15,32%
<i>Correct Directional Change</i>	49,30%	50,10%	53,28	54,08%	56,75%

As it was expected the proposed methodology clearly outperformed the existing models with leading results across all the examined metrics.

To further examine the findings of the proposed methodology, table 4 presents the percentage of which each input was selected during the sliding windows training period.

Table 5.Percentage of Selection for each Variable

Number	Variable	Percentage of selection
1	Daily High of FTSE100 all share return	45,02%
2	Daily Low of FTSE100 all share return	49,40%
3	FTSE100 all share return	43,43%

4	FTSE100 all share return	35,06%
5	FTSE100 all share return	58,57%
6	FTSE100 all share return	46,61%
7	FTSE100 all share return	39,44%
8	FTSE100 all share return	45,02%
9	FTSE100 all share return	37,45%
10	5 day Moving Average of FTSE100 all share returns	58,96%
11	15 day Moving Average of FTSE100 all share returns	50,20%
12	30 day Moving Average of FTSE100 all share returns	43,43%
13	FTSE100 daily volume	49,80%
14	5 day Moving Average of FTSE100 daily volume	46,61%
15	15 day Moving Average of FTSE100 daily volume	33,07%
16	30 day Moving Average of FTSE100 daily volume	31,47%
17	VIX Realized Volatility Index	49,40%
18	5-day Moving Average of the VIX Index	51,00%
19	15-day Moving Average of the VIX Index	43,82%
20	30-day Moving Average of the VIX Index	39,04%
21	Aggregate Advancing Volume Percentage metric	52,59%
22	5-day Moving Average of Aggregate Advancing Volume Percentage metric	47,81%
23	15-day Moving Average of Aggregate Advancing Volume Percentage metric	47,81%
24	30-day Moving Average of Aggregate Advancing Volume Percentage metric	45,82%
25	Aggregate Declining Volume Percentage metric	49,40%
26	5-day Moving Average of Aggregate Declining Volume Percentage metric	46,22%
27	15-day Moving Average of Aggregate Declining Volume Percentage metric	34,26%
28	30-day Moving Average of Aggregate Declining Volume Percentage metric	40,24%

From table 5 we see which inputs were more influential in explaining the directional changes of the FTSE100 during the period of January 2008 – December 2012. Although a relatively short time period was examined for the moving average series it can be seen that the shorter term moving average series offer more explanation for daily variations of the FTSE100 index. One possible explanation for this finding is that long term moving averages converge to a constraint value which is slightly varying.

5 Conclusions and Future Work

In the present paper we introduced a novel methodology for acquiring profitable and accurate trading strategies when speculatively trading the FTSE100 index. This methodology is a sliding window combination of an adaptive Differential Evolution with nuSVRs. It not only addresses the limitations of existing non-linear models but it also displays the benefits of using an adaptive hybrid approach to utilizing two algorithms.

Furthermore, this investigation also fills a gap in current financial forecasting and trading literature. This was accomplished by using a specialized fitness function and deploying differential evolution to optimize the confirmation threshold of the applied trading strategy on parallel of optimizing the nu-SVR model.

Experimental results proved that the proposed technique clearly outperformed the examined linear and machine learning techniques in terms of an information ratio and net annualized return. This technique is now a proven and profitable technique when applied to forecasting a major equity index. Further applications will be made to test the robustness of our model by trading other equity indices and a wider range of asset classes. In addition, the universe of inputs will be expanded in future research to include returns from specific stocks, fixed income time series, commodities and various other explanatory variables.

6 References

1. Li, Y., and Ma, W. (2010) Applications of Artificial Neural Networks in Financial Economics: A Survey. In Proceedings of International Symposium on Computational Intelligence and Design (ISCID), vol. 1, pp. 211-214
2. Vapnik, V. N. (2000) *The Nature of Statistical Learning Theory*, Springer, United States of America
3. Das, S., and Suganthan, P. (2011) Differential Evolution: A survey of the State-of-the-Art. *IEEE Transactions on evolutionary computation*. 15(1), pp. 4-30.
4. Chang, C., and Lin, C., (2002) Training nu-support vector regression: theory and algorithms. *Neural Computation* 14(8), pp. 1959-77
5. Ince, H. and Trafalis, T. (2008) Short Term Forecasting with Support Vector Machines and Application to Stock Price Prediction. *International Journal of General Systems*. 37(6), pp. 677-687.
6. Huang, W., Nakamori, Y. and Wang, S. (2005) Forecasting Stock Market Movement Direction With Support Vector Machine. *Computers & Operations Research*. 32, pp. 2513-2522.
7. Holland, J. (1995) *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence*. Cambridge, Mass: MIT Press.
8. Qin, A., Huang, V., and Suganthan, P. (2009) Differential evolution algorithm with strategy adaptation for global numerical optimization. *IEEE Transactions of Evolutionary Computation*. 13(2), pp. 398-417
9. Karathanasopoulos, A., Theofilatos, K., Leloudas, P., and Likothanassis, S. (2010) Modeling the ASE 20 Greek index using artificial neural networks combined with genetic algorithms. *Lecture notes in computer science LNCS (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics) (part 1, vol. 6352, pp. 428-435)*.
10. Dunis, C., Likothanassis, S., Karathanasopoulos, A., Sermpinis, G., and Theofilatos, K. (2013) A hybrid genetic algorithm-support vector machine approach in the task of forecasting and trading. *Journal of Asset Management* 14, pp. 52-71.
11. Burney, S., Jilani, T., and Ardil, C. (2005) Levenberg-Marquardt algorithm for karachi stock exchange share rates forecasting. *World Academy of Science, Engineering and Technology* 3, pp. 171-176.5