ON THE PREDICTION OF EXCHANGE RATE DOLLAR/EURO WITH AN SVM MODEL

CIOBANU Dumitru[1], BAR Mary Violeta[2]

University of Craiova

Abstract

Developing new methods for predictive modeling of time series and application of the existing techniques in other areas will be a permanent concern for both researchers and companies that are interested in gaining competitive advantages. In this paper I present the construction of an artificial intelligence model, based on Support Vector Machines that predict the exchange rate DOLLAR/EURO. For simulations I've used Matlab software suite.

Keywords: *Prediction, Exchange Rate, Support Vector Machines, Matlab.*

JEL classification: *C45, C53, C63.*

1. Introduction

Monitoring business processes results in obtaining time series. Analysis of available data in these time series can provide valuable information about the time evolution of the process monitored. A good forecast of future evolutions for economic activities can bring significant improvement to the technological process.

Accurate prediction for the future evolution of financial time series can produce direct gains.

[1] *Ph.D. student, Faculty of Economics and Business Administration, University of Craiova, Craiova, Romania, ciobanubebedumitru@yahoo.com*

[2] *Ph.D. student, Faculty of Economics and Business Administration, University of Craiova, Craiova, Romania, petrescumary_odobleja@yahoo.com*

Under these conditions, the analysis and understanding of economic phenomena are of interest for the researchers and firms wanting highly accurate predictions.

Most companies and organizations from nowdays collect data on a large scale. This huge amount of data can be used to extract knowledge that can represent a real advantage to business managers. Analyzing accurately and timely this large amount of data is a difficult task, in general, not possible with traditional methods. The ability to analyze and use the massive amounts of data remained far behind the possibilities of storing them. This raises new challenges for businessmen and researchers for the purposes of extracting useful information (Wang, Liao & Rees, 2002).

Increasing accuracy for predictions can save significant amount of money for a company and is a major motivation for using the methods of forecasting and systematic investigation for new models and techniques to improve the current results (Zhang, 2003).

If a number of explanatory variables must be identified and predicted, using time series approach has the advantage of easy preparation of data collection and modeling.

In time series prediction, historical data is collected and analyzed to produce models that capture the relationships between observed variables. The model is then used to predict future values of time series. There have been numerous efforts to develop and improve methods for time series prediction.

Linear approach assumes a linear process that characterizes the generation of data. There are a number of linear prediction models such as moving average, exponential smoothing, time series regression and time series decomposition.

One of the most important models is Autoregressive Integrated Moving Average - ARIMA, which was developed by Box and Jenkins (Box & Jenkins, 1976) in the 70s. Often, ARIMA is called Box-Jenkins model.

Although ARIMA models are quite flexible and can be used for a large number of time series the main limitation is given by the assumption of a linear form for the model. This means that an autocorrelated linear structure is supposed to be granted before according the model with data. Therefore, an ARIMA model cannot capture nonlinear behavior that is quite common in economic time series. Approximation with linear models of complex real-world problems is not always satisfactory as was stressed at a publicized M-competition in the early 80 (Makridakis, et al., 1982).

The approach of nonlinear time series modeling is suitable for most real-world problems. The real world is nonlinear and complex rather than linear because there are so many possible nonlinear structures and relationships. Most nonlinear models developed during the last two decades are parametric. In order to use the parametric model, it must be specified before. Therefore, they cannot be used if the data characteristics do not match the model assumptions. Parametric approach is very suitable for nonlinear problems with a complex structure but there is a lack of theories that suggest a specific form for the structure.

Artificial neural networks are algorithms and techniques that can be used for statistical modeling and is an alternative to linear regression models, the most common approach for developing predictive models.

Neural networks have several advantages including less need for formal statistical training, ability to detect, implicitly, complex nonlinear relationships between dependent and independent variables, ability to detect any possible interactions between predictor variables and the existence of a wide variety of training algorithms (Ciobanu, 2012b).

Disadvantages of neural network include the nature of "black box" computing, inclination for memorizing the data (network loses the ability to generalize), and the empirical nature of the model developed.

Another alternative that came from Artificial Intelligence is SVR (Support Vector Regression) that is an adaptation of the SVM (Support Vector Machines) for regression.

Support vector machines are a fairly simple, but very powerful concept, very well behaved in comparative tests with other popular classifiers (Meyer 2002, 2003) and have been successfully applied for problems in many fields.

Some examples of applications in which support vector machines have proven their superiority are: identifying images, medical image classification, face recognition and visually speech recognition.

Besides solving some problems that many learning methods are facing, such as small samples, overtraining, large dimensions and local minimum, support vector machines have shown a power of generalization (in the case of support vector classification - SVC) or prediction (if support vector regression - SVR) better than artificial neural networks (Ciobanu, 2012c).

Unlike neural networks, support vector machines have far fewer parameters to be set which makes it easier to determine a suitable structure for a studied problem.

In this paper, I used Matlab programming environment to create a SVM model capable to predict the following values of the exchange rate dollar/euro series. I used the built model to determine the following predictions of exchange rate and analyzed the evolution of predictions compared with observed data.

2. Support Vector Machines

Machine learning involves designing and developing algorithms that allow computers to simulate the behaviour based on empirical data.

Machine learning uses the learning process and examples to capture the interest features of the unknown probability distribution of data and perform tasks that are difficult or impossible to achieve using classical algorithms.

The problem of machine learning is the need to develop techniques that enable the machine to learn from past experience and to predict the future.

Supervised learning objective is to automatically generate rules from a database of examples already treated to make predictions on new cases. Learning database is a set of input-output pairs (x_n , y_n) with $x_n \in X$ and $y_n \in Y$, which we consider to be prepared in accordance to an unknown law on $X \times Y$.

We have a regression problem when the output values are in a continuous subset of real numbers, $Y \in \mathbb{R}$ and a classification problem when the set has finite cardinal output values $Y=\{y_1, y_2, ..., y_r\}$.

A Support Vector Machine (SVM) is a learning machine that can be used in classification problems (Cortes & Vapnik, 1995) and regression problems (Smola, 1996).

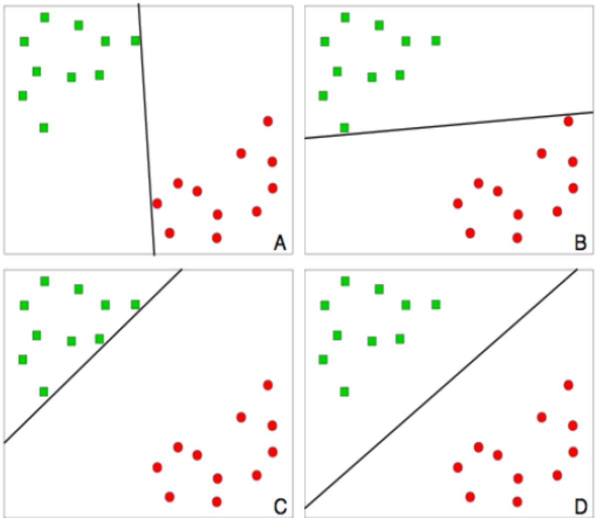
In order to perform classification, SVMs seek an optimal hyperplane that separates data into two classes.

In Figure 1. some possibilities of linear separation of two sets of elements are presented.

Support vector machines are also called classifiers with maximum edge. This means that the resulted hyperplane maximizes the distance between the closest vectors from different classes taking into account the fact that a greater margin provides increased SVM generalization capability.

The elements closest to the optimal separating hyperplane are called support vectors and only they are considered by the SVMs for the classification task. All other vectors are ignored.

Figure 1. Different variants of linear separation of two sets.

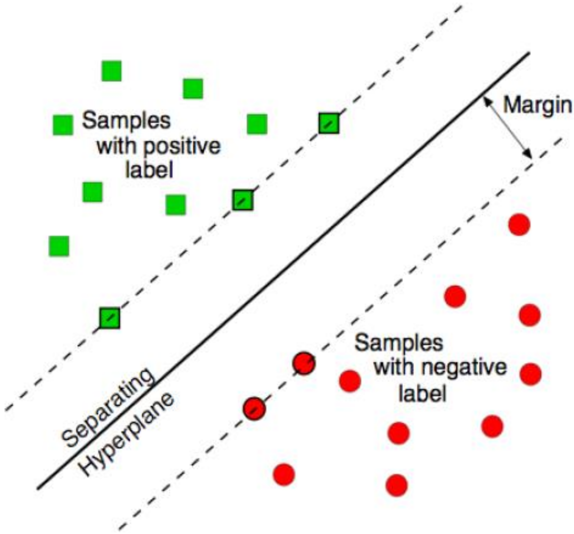


(Guggenberger, 2008)

SVM is one of the most promising algorithms in the machine learning field and there are many examples in which SVMs are successfully used, for example, text classification, face recognition, Optical Character Recognition (OCR), Bioinformatics. SVMs apply very well on these datasets and often exceed the performance of other traditional techniques. Of course, this is not a magic solution as set forth in (Bennett & Campbell, 2000), there are still some open issues, such as incorporation of domain knowledge, a new model selection and interpretation of results produced by SVMs.

Figure 2. illustrates a classifier with maximum margin and support vectors, those located on dotted lines from both sides of the optimal separating hyperplane.

Figure 2 Optimal separating hyperplane. The vectors on dotted lines are support vectors.



(Guggenberger, 2008)

SVMs have been used in several real-world problems:

- classification of text (and hypertext);
- image classification;
- in bioinformatics (protein classification, classification of types of cancer);
- classification of music;
- handwritten character recognition.

In (Chen, Jeong & Härdle, 2008), authors propose a method GARCH (Generalised AutoRegressive Conditional Heteroscedasticity) based on recurrent SVR whose performance exceeds other approaches such as moving average (MA), recurrent Neural Networks (NN) and parameterized GARCH in terms of their ability to Forecast financial market volatility.

3. Support Vector Regression

Support Vector Machines have been developed to solve the problem of classification. A problem of regression differs from classification in the sense that observations are associated with numeric values and not a label from a discrete set.

Given this difference we can easily adapt the application of support vector machines to regression problems where we are dealing with numerical observations.

Thus, the regression problem can be formulated as follows:

with data

- A universe of data *X*,
- A sample set *S*, $S \subset X$,
- A target function $f : X \to \mathbb{R}$,
- A training set *D*, where $D = \{(x, y) | x \in S \text{ si } y = f(x)\}$.

We need to determine a model $\hat{f}$: $X \to \mathbb{R}$ using *D* so that $\hat{f}(x) \cong f(x)$ for any $x \in X$.

The basic idea of machine learning is retained, namely the determination of a model (function) that fits best with the target function for all data elements in the universe.

As in the case of classification *X* is a multidimensional real data set that is $X \subset \mathbb{R}^n$ with $n \geq 1$.

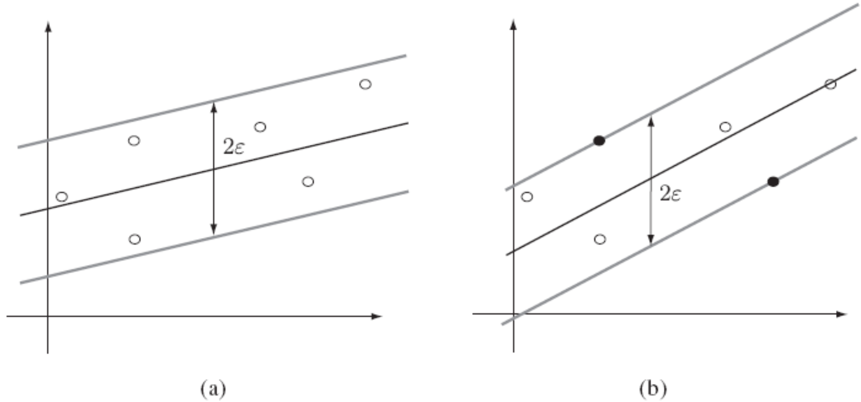
To develop support vector machine in regression context the maximization of the margin is used. The idea used is the same as for classification maximizing the margin.

It gives a hyperplane and maximizes distances from observations to it. For a regression problem where the observations of training set

$$D = \{(\bar{x}_1, y_1), (\bar{x}_2, y_2), \dots, (\bar{x}_l, y_l)\} \subset \mathbb{R}^n \times \mathbb{R}, \tag{1}$$

are contained in a hypertube with width 2ε and $\varepsilon > 0$ (Figure 3.), this can be interpreted as a regression model considering that there is a hyperplane positioned in the center of the hypertube that approximate observations. Usually there are several ways of positioning the hypertube with width 2ε to contain all the training observations. There is one optimal positioning of hypertube so as more observations are pushed closer to the outside of hypertube. Optimal Hypertube alignment is obtained when the distances from the observations to the central hyperplane are maximized. This is illustrated in Figure 3., Where filled circles represent observations that act as constraints in optimization problem.

Figure 3. Linear regression modeling using hypertubes with width 2ε (a) a hypertube containing all observations, (b) optimal regression model with maximum margin.



(Hamel, 2009)

This is very similar to the problem of maximizing the margin of decision area, and we can use the same optimization problem to determine the optimal alignment by adjusting the hyperplane with corresponding constraints.

The primal problem is considered first,

$$\min \Phi(\overline{w}, b) = \min_{\overline{w}, b} \frac{1}{2} \overline{w} \cdot \overline{w} \ , \tag{2}$$

so that constraints

$$\begin{aligned} y_i - \hat{f}(\overline{x}_i) &\le \varepsilon, \\ \hat{f}(\overline{x}_i) - y_i &\le \varepsilon, \end{aligned} \tag{3}$$

Are satisfied for $i = 1,\dots,l$ and $\hat{f}(\overline{x}) = \overline{w} \cdot \overline{x} + b$.

The two inequalities (3) can be written as

$$\left| y_i - \hat{f}(\overline{x}_i) \right| \le \varepsilon \ . \tag{4}$$

Geometric interpretation of optimization is that of making rotation (updating weights) and shift or displacement (bias update) until the distance from the observations to the central hyperplane is maximized with constraint to maintain the observations inside of hypertube.

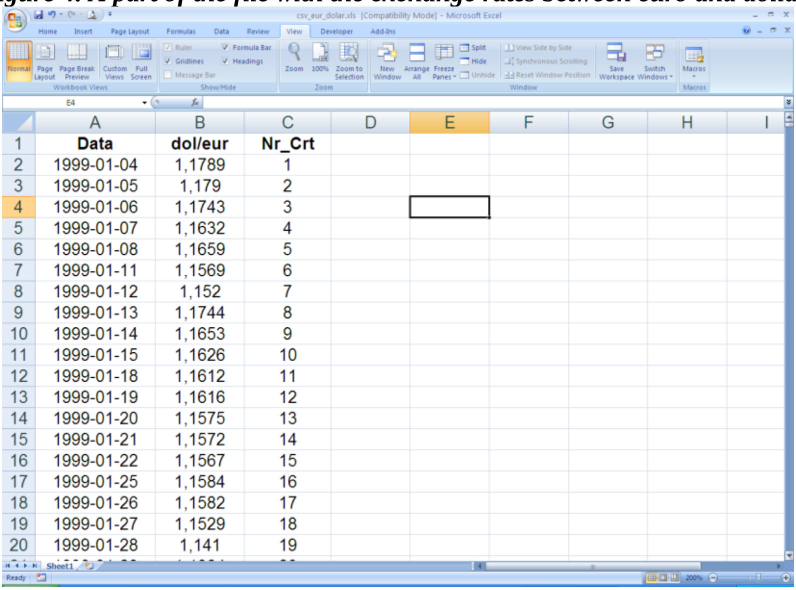
4. Creating the SVR model for predicting the dollar/euro exchange rate

The time series used was the euro-dollar exchange rate and was downloaded from the website of European Central Bank at http://sdw.ecb.europa.eu/browseSelection.do?DATASET=0&sfl1=4&FREQ=D&sfl3=4&CURRENCY=USD&node=2018794&SERIES_KEY=120.EXR.D.USD.EUR.SP00.A .

After the removal of lines, unhelpful to our analysis, the Microsoft Excel file with values of exchange rate between euro and dollar looks like Figure 4.

The program used to perform simulations was Matlab version 7.12.0 (R2011a).

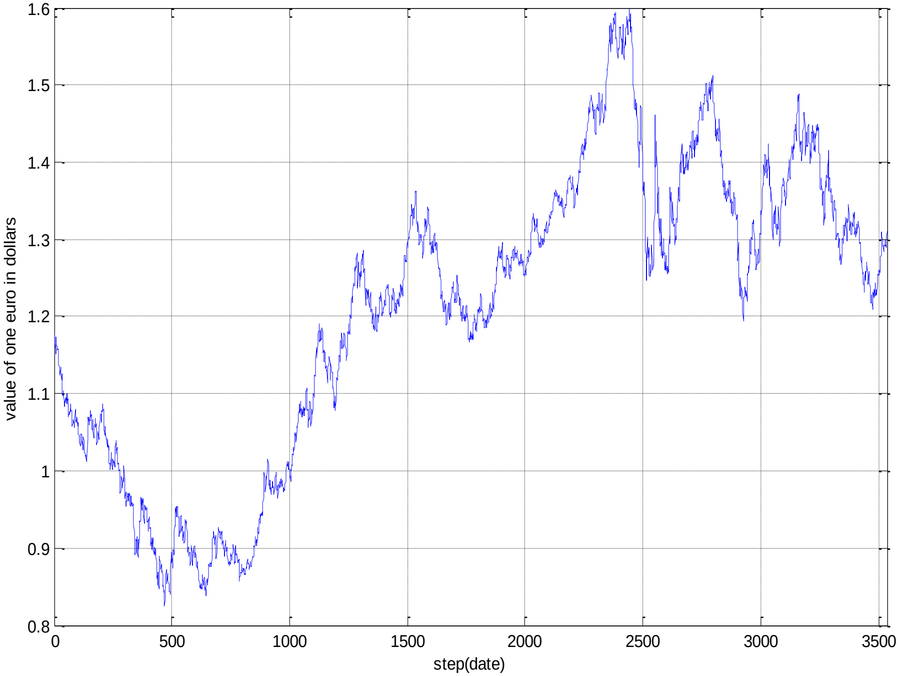
Figure 4. A part of the file with the exchange rates between euro and dollar.



	A	B	C
1	Data	dol/eur	Nr_Crt
2	1999-01-04	1,1789	1
3	1999-01-05	1,179	2
4	1999-01-06	1,1743	3
5	1999-01-07	1,1632	4
6	1999-01-08	1,1659	5
7	1999-01-11	1,1569	6
8	1999-01-12	1,152	7
9	1999-01-13	1,1744	8
10	1999-01-14	1,1653	9
11	1999-01-15	1,1626	10
12	1999-01-18	1,1612	11
13	1999-01-19	1,1616	12
14	1999-01-20	1,1575	13
15	1999-01-21	1,1572	14
16	1999-01-22	1,1567	15
17	1999-01-25	1,1584	16
18	1999-01-26	1,1582	17
19	1999-01-27	1,1529	18
20	1999-01-28	1,141	19

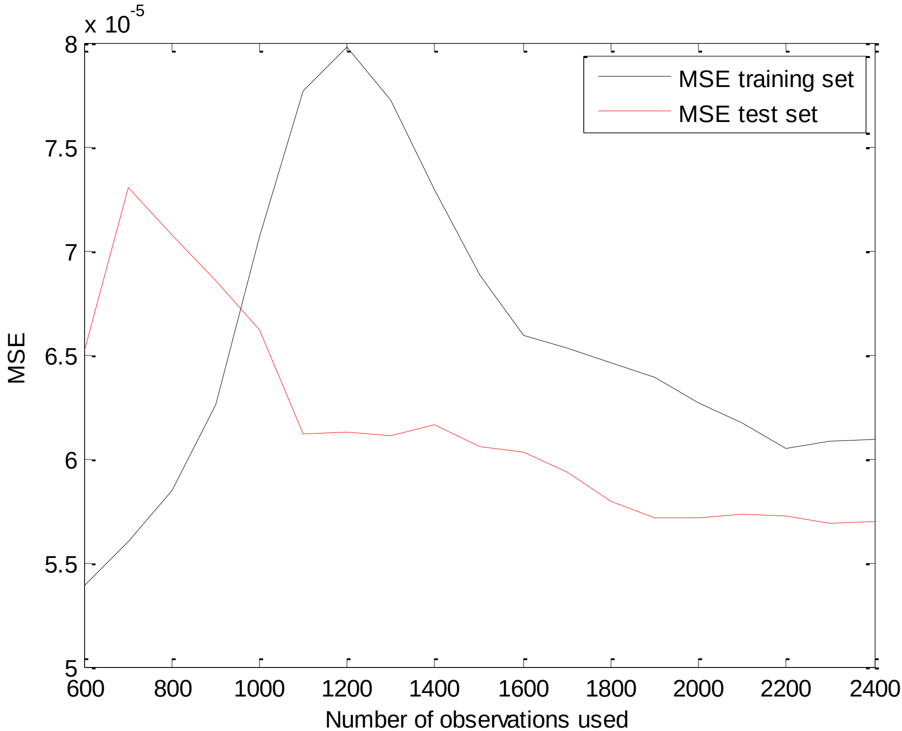
I imported into Matlab data from column 2 of the excel file. The data set consists of 3536 observations between 1999-01-04 and 2012-10-17 with a mean of 1,2093545 and a standard deviation of 0,190516.

Figure 5. The evolution of exchange rate dollar/euro, 3536 observations.



In Figure 5. is shown the evolution of exchange rate between euro and dollar for 3536 observations. From this, I kept 100 observations for verification, 100 for test and use the other for training.

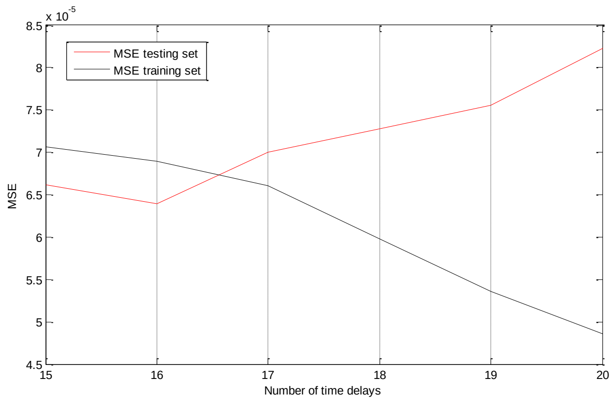
Figure 6. The evolution of MSE for training set and testing set when varying the number of observation used.



I measured the model performance using MSE (Mean Squared Error).

I want that the two MSEs to have approximately the same value so, I will use 1000 observations because the intersection of this graphs is close to this value (Figure 6).

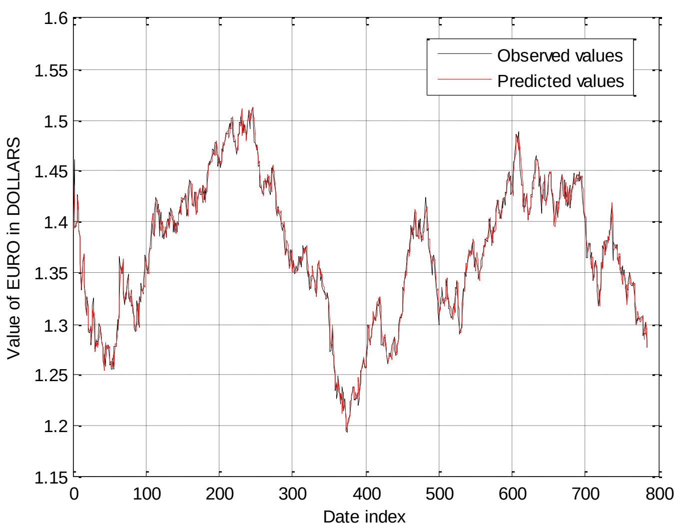
Figure 7. The evolution of MSE for training set and testing set when varying the number of time delays.



Using the same idea I set the number of time delays to 16 (Figure 7).

With these two values for number of observations and time delays I trained a SVM

Figure 8. The observed values and predictions for training set.



For the training data considered in my example I obtained MSE = $6.8933 \cdot 10^{-5}$. MSE for the testing set was $6.3922 \cdot 10^{-5}$.

Figure 9. The observed and the predicted values for testing set. Using as inputs 16 measured values the model returns us the next prediction. The case of one step ahead prediction.

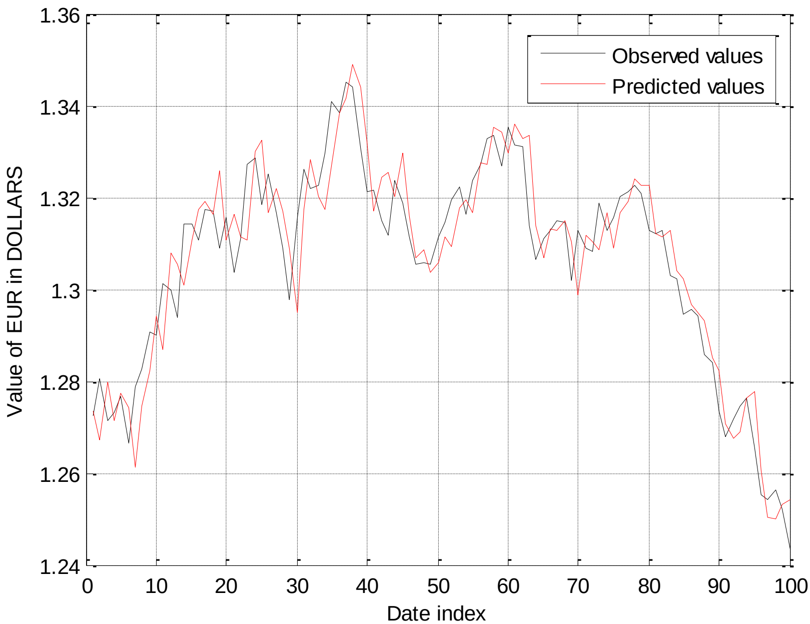
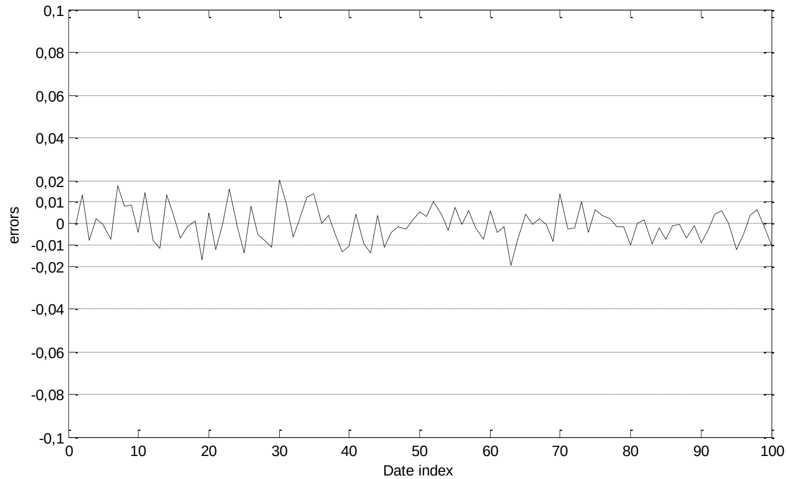


Figure 10. The errors for the testing set. The case of one step ahead prediction.



I used the model to predict the next 100 values (set of verification). First I used the observed values as inputs, introducing 16 measured values and the model provided our next prediction for the exchange rate.

Figure 11. The observed and the predicted values for the set of verification. Using as inputs 16 measured values the model returns us the next prediction. The case of one step ahead prediction.

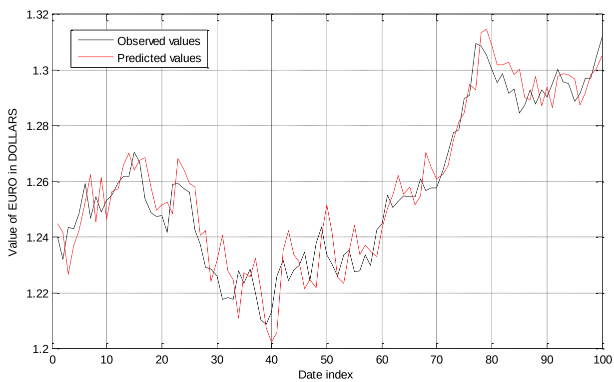
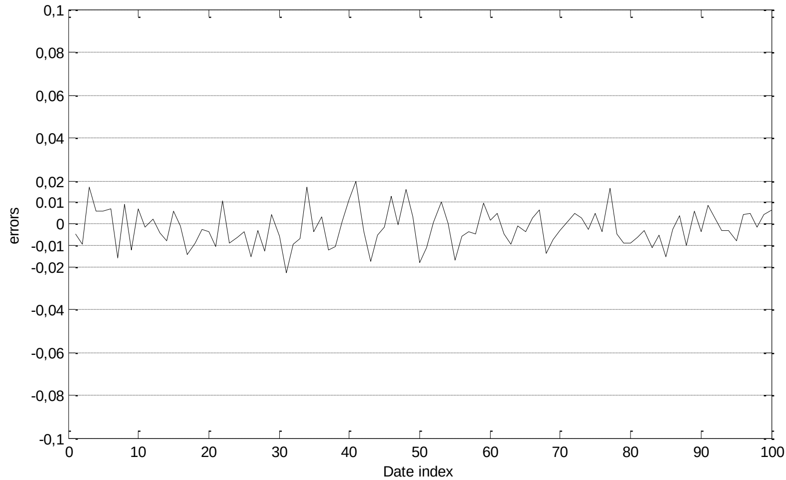


Figure 12. The errors for the set of verification. The case of one step ahead prediction.



Then, I used the model to predict the next 100 values but using as input the predictions obtained, the result can be seen in Figure 13 and Figure 14.

Figure 13. The observed and the predicted values for the set of verification. The case of 100 step ahead prediction

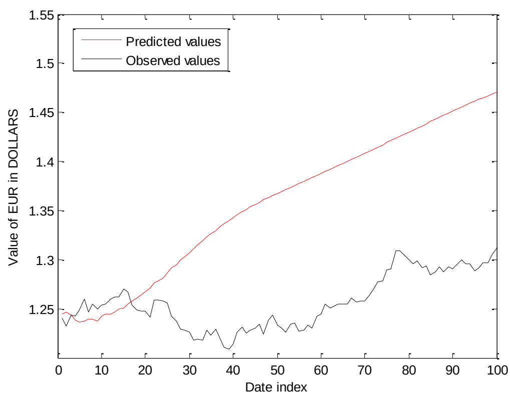
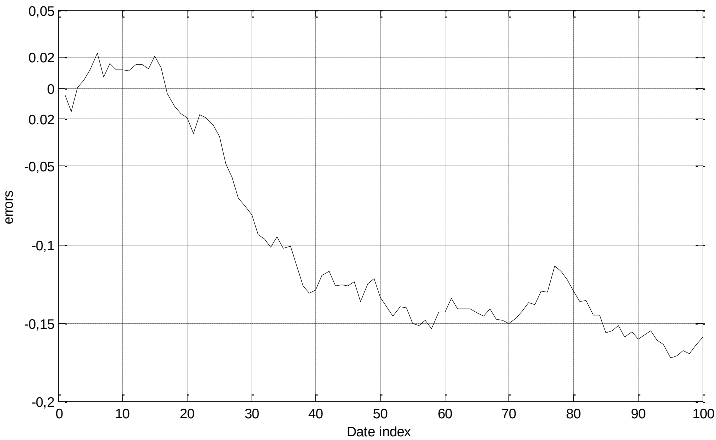


Figure 14. The errors for the set of verification. The case of 100 step ahead prediction.



5. Conclusions

Using SVM for prediction is a good alternative for both traditional methods as well as those methods arising from computational intelligence like Neural Networks (NN). To adjust the model with data we must determine just a few parameters and this represents a major advantage of SVM over NN.

In the case of prediction for many steps ahead I have performed many tests and each time the series of predictions showed large deviations from the observed values after a relatively small number of steps. I concluded that all models obtained using SVM to predict exchange rate shows a chaotic behaviour. The fact that in this case the predictions for a longer period did not work is not a minus for SVM models but underlines the chaotic nature of time series euro-dollar.

Thus, considering the chaotic nature of exchange rate time series, prediction with an acceptable error can be done only for a few steps forward.

In the case of exchange rate euro-leu one of the authors has identified many signs of chaotic behaviour like positive Lyapunov exponents and noninteger correlation dimension (Ciobanu, 2012a).

It remains as a future direction to underlie the chaotic behavior of exchange rate euro-dollar.

Another direction for future works is to identify exogenous variables that influence the exchange rate euro-dollar and use them to improve predictions.

6. References

- Bennett, K. P.; Campbell, C., Support vector machines: hype or hallelujah?, SIGKDD Explorations Newsl., Vol 2, No. 2, pp. 1–13, 2000.
- Box, G. E. P.; Jenkins, G. (1976) *Time Series Analysis: Forecasting and Control*, Holden-Day, 1976.
- Ciobanu, D. (2011) *Using Neural Networks for Prediction in Business Planning*, Proceedings of the international conference ECOTREND 2011 – Exit from the crisis and revival of sustainable growth, November 25-26, Târgu Jiu, Romania, 2011.
- Ciobanu, D. (2012a) *EVIDENCE OF CHAOS FOR THE TIME SERIES OF EXCHANGE RATE EUR-LEU*, International Scientific Conference ECO-TREND 2012 IXth Edition, November 23-24, Târgu Jiu, Romania, 2012.
- Ciobanu, D. (2012b) On a model for predicting the exchange rate euro-leu with a NAR neural network, Proceedings of the 19th International Economic Conference – IECS 2012.
- Ciobanu, D. (2012c) Using SVM for Classification, *Acta Universitatis Danubius. Œconomica*, Print ISSN: 2065-0175, Online ISSN: 2067 – 340X, Vol. 8, No. 5, pp. 209-224, 2012.
- Chen, S.; Jeong, K.; Härdle, W., Support Vector Regression Based GARCH Model with Application to Forecasting Volatility of Financial Returns, SFB 649 "Economic Risk", Humboldt-Universität zu Berlin, 2008.
- Cortes, C.; Vapnik, V., Support vector networks, *Machine Learning*, 20(3), pp. 273-297, 1995.

- Guggenberger, A., (2008) *Another Introduction to Support Vector Machines*, Available online at http://mindthegap.googlecode.com/files/AnotherIntroductionSVM.pdf accesed May 2012.
- Hamel, L., (2009) Knowledge Discovery with Support Vector Machines, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2009.
- Kadilar, C.; Simsek, M.; Aladag, C. H. (2009) *Forecasting The Exchange Rate Series With Ann: The Case Of Turkey*, Istanbul University Econometrics and Statistics Journal, vol. 9, issue 1, pp. 17-29, 2009.
- Makridakis, S.; Anderson, A.; Carbone, R.; Fildes, R.; Hibdon, M.; Lewandowski, R.; Newton, J.; Parzen, E.; Winkler, R. (1982) *The accuracy of extrapolation (time series) methods*, Results of a forecasting competition. *Journal of Forecasting*, 1, pp. 111-153, 1982.
- Meyer, D.; Leisch, F.; Hornik, K., The support vector machine under test, *Neurocomputing*, 55: 169–86, 2003.
- Smola, J., Regression estimation with support vector learning machine, Master's thesis, Technische Universitat Munchen, 1996.
- Sumathi, S.; Surekha, P. (2010) *Computational Intelligence Paradigms: Theory and Applications using Matlab*, CRC Press, ISBN 978-4398-0902-0, New York, 2010.
- Wang, M.; Rees, S.J.; Liao, S.Y. (2002) *Building an online purchasing behavior analytical system with neural network*, Edited by Zanasi, Brebbia and Melli, Data Mining III., WIT Press, 2002.
- Widrow, B.; Lehr, M.A. (1990) *30 Years of Neural Networks: Perceptron, Madaline and Backpropagation*, Proceedings of the IEEE, 78:1415–1442, 1990.
- Wilshaw, D.J.; Von Der Malsburg, C. (1976) *How Patterned Neural Connections can be Set up by Self-Organization*, in Proceedings of the Royal Society of London B, vol. 194, pp. 431-445, 1976.
- Yang, N. (2011) *Research on prediction of chaotic exchange rate time series applying dynamic component predicting model*, 2nd

International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), Zhengzhou, China, 2011.

- Zhang, G. P. (2003) *Time series forecasting using a hybrid ARIMA and neural network model*, Neurocomputing, vol. 50, pp. 159-175, 2003.
- Zhang, G.; Patuwo, B. E.; Hu, M. Y. (1998) *Forecasting with artificial neural networks: The state of the art*, International Journal of Forecasting, 14, pp. 35-62, 1998.