Pis'ma v ZhETF, vol. 102, iss. 12, pp. 958 – 961

FOREX trades: can the Takens algorithm help to obtain steady profit at investment reallocations?

V. Yu. Petrov[+*], M. I. Tribelsky[+×1)]

[+]Lomonosov MSU, 119991 Moscow, Russia

[*]Central Economics and Mathematics Institute of the RAS, 117418 Moscow, Russia

[×]MIREA, 119454 Moscow, Russia

Submitted 26 October 2015

We report our preliminary results of application of the Takens algorithm to build a FOREX trade strategy, resulting in a steady long-time gain for a trader. The actual historical rates for pair EUR vs. USD are used. The values of various parameters of the problem including the "stop loss" and "take profit" thresholds are optimized to provide the maximal gain during the training period. Then, these values are employed for trades. We have succeeded to get the steady gain, if the spread is neglected. It proves that the FOREX market is predictable.

DOI: 10.7868/S0370274X1524011X

Introduction. During the recent decades *econophysics* has been quite appreciated in the physical community as a new branch of science, where the approaches of non-equilibrium statistical physics and theory of dynamical systems are applied to explain economical, financial, and social phenomena. The corresponding papers are published by the top-rank general and physical journals, see, e.g., [1–5]. Some physical journals, for example, Physica A, include this topic into the list of covered issues. Prestigious scientific publishers printed books, devoted to this subject matter [6], etc.

However, as long as the actual market dynamics is a concern, most publications are devoted just to descriptions of various statistical properties of the financial time series. The number of publications trying to predict certain events in this dynamics (e.g., market crashes) are much fewer. Regarding the publications, where the methods of econophysics are applied to build up a successful strategy for financial speculations, we do not know any. Among other things, it is explained by obvious commercial importance of such results. It is difficult to imagine a person, who has succeeded to build up such a strategy and then published all details in a scientific paper. Moreover, if such a person appeared, he/she immediately would have received a lot of followers, whose action in the market would change its dynamics, so that, eventually, the developed algorithms would stop working.

On the other hand, apart their practical importance, the questions: (i) whether the dynamics of an actual market is predictable by the econophysical methods; if so, in which sense and (ii) whether such predictions could be employed to get a steady long-time profit, have deep scientific meaning. In the present letter we try to answer these questions, focussing just on the FOREX (Foreign currency Exchange) market. There are some arguments *pro* and *contra* the predictability of this market. On one hand the FOREX market is almost an ideal economical example of an extremely large ensemble of interacting "particles" (individual trades). According to data of *the Bank of International Settlements* the average daily turnover of the global FOREX market in April 2013 is $ 5.3 trillion [7]. For comparison the US GDP in 2013 was $ 16.163 trillion [8], i.e., the FOREX 3-day-turnover is equivalent to the GDP produced in the entire US in a year! Such a huge turnover guarantees self-averaging of the FOREX market and its statistical independence of actions of any individual trader. It may be regarded as grounds to apply a statistical approach to description of this market.

The known phycological facts that the behavior of a mob is much simpler than that of an individual person, and that in a similar situation people, as a rule, act in a similar manner, say in favor of the hypothesis of predictability of the market. Finally, the fact of existence of a ceratin fraction of the FOREX traders, who can manage to get a steady profit by trading for many years is the best empirical evidence of the predictability.

1) e-mail: tribelsky@mirea.ru

On the other hand, the statistical nature of the FOREX market and even its predictability, if any, do not guarantee that the prediction could be made with help of the approaches of the non-equilibrium statistical physics, especially, bearing in mind that the system in question by no means may be regarded as ergodic, or close to ergodic. However, since the proof of the pudding is in the eating, the only way to find the answers to the questions raised above is to try to develop a constructive description of the market based upon the approaches of the non-equilibrium statistical physics and to check if it works out. The results of such a study are presented below.

Motivation. The study is motivated by the old results of one of the authors [9, 10]. These publications reported about the predictions of the dynamics of the FOREX market based on the Takens algorithm [11]. Nonetheless, there is a dramatic difference in approaches of Ref. [9, 10] and the one discussed in the present Letter. First, while in Ref. [9, 10] the predicted price is the daily average, now we are focus on a minute scale, characterized by a completely different dynamics. Second, the realization of the prediction algorithm in Ref. [9, 10] is achieved with help of an original, very complicated approach employed adaptive neural networks. It is very time consuming. To make a prediction for the next week almost the entire present week of non-stop work of a standard commercial PC is required. In contrast to that, now a standard free package TISEAN [12] is employed and the time, required for a prediction is shorted to a few minutes.

However, the most important difference is that in the approach developed in Ref. [9, 10] a prediction is the final result. The decision how the prediction should be employed for a trade is made by a human being. Though this approach does work with high efficiency (the net gain of a test trading session reported in Ref. [10] was 63 % relative to the initial deposit), owing to the crucial role of a human factor plaid in this experiment, it cannot be regarded quite satisfactory from the scientific point of view. Counterpoint to that, now the trade process is fully automated. Moreover, the maximization of the gain during the training period is the criterion to select the best values of the tuning parameters of the algorithm; for more detail see below.

Surrogate data. The initial database employed was tick data for pair EUR vs. USD for the period form 2010 to 2014 (about 10^7 records) [13]. Based on it, we calculate the *minute average*, which is the average of all *bid* and *ask* prices for every tick in a given minute. The database $\{x_n\}$ of the minute average was the surrogate data used in our calculations.

The next step is to smoothen the surrogate data by the moving widow method, i.e., to obtain

$$X_n = \frac{\sum_{i=0}^{N-1} w_i x_{n-i}}{\sum_{i=0}^{N-1} w_i}. \tag{1}$$

We have used several monotonically decreasing wight functions w_i, satisfying conditions $w_0 = 1$ and $w_N = 0$. Since results practically do not depend on the specific choice of these functions, in what follows we will focus on the case of the simple linear one.

Then, the standard deviation of x_n from X_n in the form

$$\delta x_n = \frac{\sum_{i=0}^{N-1} w_i \sqrt{(x_{n-i} - X_{n-i})^2}}{\sum_{i=0}^{N-1} w_i}, \tag{2}$$

regarded as quantitative measure of the market volatility is introduced.

Method. The idea of application of the Takens algorithm to the problem in question is based on the hypothesis that the observed market dynamics is a discrete 1D projection of a trajectory of a complicated dynamical system in the corresponding multidimensional phase space. If so, to predict the behavior of this system in future one has to reconstruct the topological structure of the trajectories embedded in this multidimensional phase space and to find in past a segment of the phase trajectory the closest to the most recent segment of it. Then, for a while the continuation of the recent trajectory in future will follow this selected trajectory in the past. For a chaotic attractor the period of the closeness of the two trajectories is defined by the largest Lyapunov exponent, which, eventually, determines the prediction horizon.

To recover the topological structure of the phase trajectories 1D series $\{X_n\}$ is employed to generate the series of M-dimensional Takens vectors $\{\mathbf{Y}_m\}$, where

$$\mathbf{Y}_m = \{X_m, X_{m-d}, X_{m-2d}, \ldots, X_{m-(M-1)d}\}. \tag{3}$$

Here M is the *embedding dimension* and integer d is called the *time delay*, or *lag*. Then, the phase trajectory is approximated by a smooth curve, sequentially passing through the ends of every Takens vector.

Immediately a number of questions arises. Fist of all, the concept of a phase space is applicable for the autonomous systems only. In other words, the coefficients of the unknown to us dynamical system governing the FOREX market must be time-independent. It hardly could be the case, because of the global progress of the mankind. The best, what we could hope for, is that the changes related to the global progress are rather slow and do not affect much the market dynamics at the minute time-scale. However, even if it is so, the problem of the optimal length of the historical dataset remains. On one hand, if the dataset is too small, the accuracy of calculations is poor. On the other hand, if the dateset is too large the non-steadiness of the dynamics becomes essential. It violates the applicability conditions for the Takens algorithm and also results in bad predictability. Calculations show that in our case the best results correspond to the length of $\{X_n\}$ in several thousands of terms.

Regarding the time delay, formally, it could have any non-zero value. However, actually, to provide the best accuracy of the approximation for the phase trajectories the sizes of the attractor along any direction in the embedding space should be approximately the same.

As for the embedding dimension, if it is selected too small (smaller than the actual dimension of the attractor), the attractor will be distorted. It results in poor predictions. If it is selected too large, the set of the available Takens vectors (remember, that owing to the aforementioned constraint the length of the series $\{\mathbf{Y}_m\}$ is limited) will be scattered in a larger embedding space. It gives rise to a large approximation errors in the fitting of the phase trajectories and, eventually, to bad predictions too. To select the best value of M we employ a certain modification of the *False Nearest Neighbors* (FNN) method [14].

To optimize the values of these and other parameters, required for implementation of the TISEAN package we need a quantitative measure of the quality of the predictions made. In the case of a prediction of binary events (say, a sequence of "1" and "0") such a measure could be the ratio of the correct predictions to the total number of predictions. However, in the case of a prediction of a behavior of a continuous curve to build such a measure is not a straightforward matter. Since we are going to employ the predictions for successful trades, a natural measure of the prediction quality could be the net balance of the trades for a certain rather an extended period.

Then, a strategy to employ the predictions for the trades is required. We have adopted the following strategy. First, a continuous segment S with a length about 4–5 weeks is randomly selected from the series $\{x_n\}$. Next, a subsegment s with the length of a few thousands records, adjacent to the left boundary of S is selected. This subsegment is employed as a historical database to train TISEAN and to obtain a prediction of $\{X_n\}$, beginning from the right boundary of the subsegment s. The values of the parameters of TISEAN used for this prediction were chosen in agreement with the principles discussed above. Then, the subsegment s is shifted to one record forward ($n \to n+1$) and a new prediction with the same, unchanged values of the TISEAN parameters is made, and so on, until the right boundary of s meets the one of S. Thus, a sequence of the overlapping predictions $\{P_n\}$ is obtained.

Among all prediction $\{P_n\}$ only those with pronounced trends (either growing or decreasing) are selected. For the selected predictions the instantaneous value of x_n at the beginning of the corresponding prediction ($n = n_0$) is compared with $X_{n_0} - \delta x_{n_0}$ for the growing trend and $X_{n_0} + \delta x_{n_0}$ for decreasing. If $x_{n_0} > X_{n_0} - \delta x_{n_0}$ in the former case, or $x_{n_0} < X_{n_0} + \delta x_{n_0}$ in the latter, no action is taken. In contrast, at the opposite signs of the inequalities "buy" or "sell" positions are opened, respectively. The size of the positions always remains the same. The "take profit" threshold for the positions of both types equals X_{n_0} (i.e., it corresponds to the moment when the local rate fluctuations relax to the smoothed "equilibrium" value). The "stop loss" thresholds are $X_{n_0} \pm 2\delta x_{n_0}$, respectively, see Fig. 1 as an illustration.

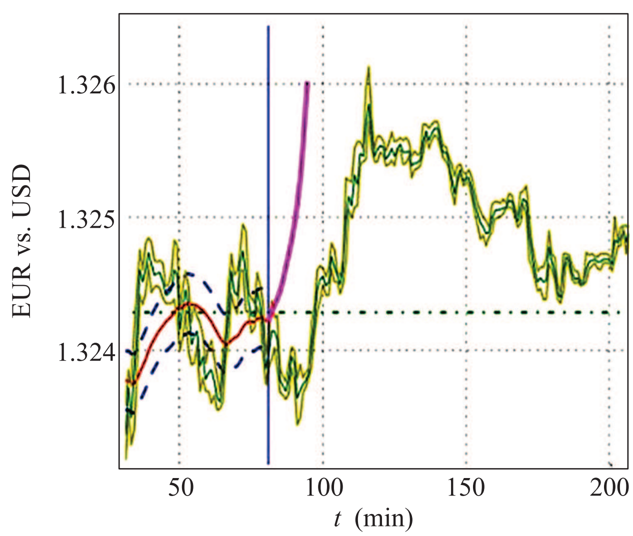

Fig. 1. (Color online). Illustration to the trade strategy. The dark green line indicates the average minute rates x_n. Two light green lines, framing it, correspond to the actual bid and ask prices. The red line is the smoothed rates X_n. The prediction of the smoothed rates is shown as a pink line. Two blue dashed lines correspond to $X_n \pm \delta x_n$. A solid vertical blue line indicates the beginning of the prediction ($n = n_0$). The value of X_{n_0} is shown with a horizontal dashed-dotted line. Since $x_{n_0} < X_{n_0} - \delta x_{n_0}$ a "buy" position is opened. The "take profit" and "stop loss" thresholds are set at X_{n_0} and $X_{n_0} - 2\delta x_{n_0}$, respectively

As a result, in the end of segment S we get a certain balance, which is a function of the values of the problem parameters. Then, the procedure is repeated with different values of these parameters in order to maximize the balance according to the standard procedure of a numerical maximization of a function of several variables. Finally, the optimized values of the parameters for S are obtained. These values are used for the next (relative to the end of S) week to mimic actual trades. After the end of this week the entire procedure is repeated.

Results. The approach described above has been applied to a several hundred segments S randomly selected for the pair EUR vs. USD in the historical interval 2010–2014. In all the cases the approach has given rise to a steady gain, provided the spread is neglected, i.e., the *bid* and *ask* prices are supposed to be equal each other. A typical example of such a kind is presented in Fig. 2. However, as soon as the actual non-zero spread

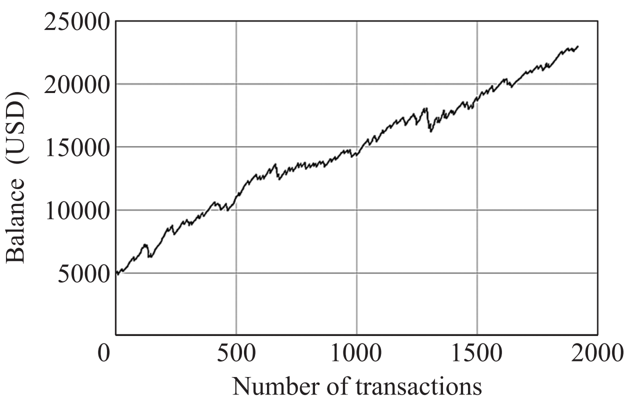

Fig. 2. Growth of the balance in time. Spread is neglected

is considered, the steady gain is replaced with a steady loss.

Conclusions. The performed analysis has provided a convincing evidence of predictability of the FOREX market by the methods discussed, at least for the pair EUR vs. USD, the minute time-scale and the period 2010–2014.

In our view the failure to obtain a steady gain with help of the developed approach, when the actual spread is considered, is explained by the fact that while the FOREX market itself is a global self-averaging phenomenon, the spread is connected with a specific features of a specific trading platform and hence, it does not obey statistical laws.

Nonetheless, there is a hope that the fatal role of the spread may be overcome if the dynamics with another time-scale is considered and the optimization process is improved. The corresponding study is in progress and will be reported elsewhere.

The financial support of this study through Russian Science Foundation grant # 14-18-01999 *Measurements and Forecasting in Digital Economics* is acknowledged.

1. V. Plerou, P. Gopikrishnan, and H. E. Stanley, Nature **421**, 130 (2003).
2. X. Gabaix, P. Gopikrishnan, V. Plerou, and H. E. Stanley, Nature **423**, 267 (2003).
3. V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, and H. E. Stanley, Phys. Rev. Lett. **83**, 1471 (1999).
4. D. Sornette, F. Deschatres, T. Gilbert, and Y. Ageon, Phys. Rev. Lett. **93**, 228701 (2004).
5. Y. Yura, H. Takayasu, D. Sornette, and M. Takayasu, Phys. Rev. Lett. **112**, 098703 (2014).
6. R. N. Mantegna and H. E. Stanley, *Introduction to Econophysics: Correlations & Complexity in Finance*, Cambridge University Press, Cambridge (2000).
7. https://www.bis.org/press/p130905.htm.
8. http://www.tradingeconomics.com/united-states/gdp.
9. M. Tribelsky, Y. Harada, N. Makarenko, and Y. Kuandykov, *Predictability of market prices*, in *Empirical Science of Financial Fluctuations. The Advent of Econophysics*, ed. by H. Takayasu, Springer, Tokyo, Berlin, etc. (2001), p. 241.
10. M. I. Tribelsky, *New complex approach to prediction of market prices*, in *Toward Control of Economic Change – Application of Econophysics*, ed. by H. Takayasu, Springer, Tokyo, Berlin, etc. (2003), p. 131.
11. F. Takens, *Detecting strange attractors in turbulence*, in *Springer Lecture Notes in Mathematics*, Springer, Tokyo, Berlin, etc. (1981), v. 898, p.366.
12. http://www.mpipks-dresden.mpg.de/tisean/archive_3.0.0.html.
13. http://ratedata.gaincapital.com/.
14. M. B. Kennel, R. Brown, and H. D. I. Abarbanel, Phys. Rev. A **45**, 3403 (1992).