Physica A 392 (2013) 6414–6426

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

The string prediction models as invariants of time series in the forex market

R. Pincak *

Institute of Experimental Physics, Slovak Academy of Sciences, Watsonova 47, 043 53 Kosice, Slovak Republic
Bogoliubov Laboratory of Theoretical Physics, Joint Institute for Nuclear Research, 141980 Dubna, Moscow Region, Russia

HIGHLIGHTS

- We apply a new approach of string theory to the real financial market.
- The models are constructed with an idea of prediction based on string invariants.
- The models are compared to support vector machines and artificial neural networks.
- Comparisons were done on artificial and a financial time series.
- The presented string models could be useful for portfolio creation and risk management.

ARTICLE INFO

Article history:
Received 13 May 2013
Available online 9 August 2013

Keywords:
Finance forex market
Nonlinear statistics
String theory
Trading strategy
Financial forecasting

ABSTRACT

In this paper we apply a new approach of string theory to the real financial market. The models are constructed with an idea of prediction models based on the string invariants (PMBSI). The performance of PMBSI is compared to support vector machines (SVM) and artificial neural networks (ANN) on an artificial and a financial time series. A brief overview of the results and analysis is given. The first model is based on the correlation function as invariant and the second one is an application based on the deviations from the closed string/pattern form (PMBCS). We found the difference between these two approaches. The first model cannot predict the behavior of the forex market with good efficiency in comparison with the second one which is, in addition, able to make relevant profit per year. The presented string models could be useful for portfolio creation and financial risk management in the banking sector as well as for a nonlinear statistical approach to data optimization.

1. Introduction

Time series forecasting is a scientific field under continuous active development covering an extensive range of methods. Traditionally, linear methods and models are used. Despite their simplicity, linear methods often work well and may well provide an adequate approximation for the task at hand and are mathematically and practically convenient. However, the real life generating processes are often non-linear. This is particularly true for financial time series forecasting. Therefore the use of non-linear models is promising. Many observed financial time series exhibit features which cannot be explained by a linear model.

There are plenty of non-linear forecast models based on different approaches (e.g. GARCH [1], ARCH [2], ARMA [3], ARIMA [4] etc.) used in financial time series forecasting. Currently, perhaps the most frequently used methods are based on Artificial Neural Networks (ANN, which covers a wide range of methods) and Support Vector Machines (SVM). A number of

* Correspondence to: Institute of Experimental Physics, Slovak Academy of Sciences, Watsonova 47, 043 53 Kosice, Slovak Republic. Tel.: +421 907936147; fax: +421 55 633 62 92.
E-mail addresses: pincak@saske.sk, pincak@sors.com.

0378-4371/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.physa.2013.07.048

research articles compare ANN and SVM to each other and to other more traditional non-linear statistical methods. Tay and Cao [5] examined the feasibility of SVM in financial time series forecasting and compared it to a multilayer Back Propagation Neural Network (BPNN). They showed that SVM outperforms the BP neural network. Kamruzzaman and Sarker [6] modeled and predicted currency exchange rates using three ANN based models and a comparison was made with the ARIMA model. The results showed that all the ANN based models outperform the ARIMA model. Chen et al. [7] compared SVM and BPNN taking the auto-regressive model as a benchmark in forecasting the six major Asian stock markets. Again, both the SVM and BPNN outperformed the traditional models.

While the traditional ANN implements the empirical risk minimization principle, SVM implements the structural risk minimization [8]. Structural risk minimization is an inductive principle for model selection used for learning from finite training data sets. It describes a general model of capacity control and provides a trade-off between hypothesis space complexity and the quality of fitting the training data (empirical error). For this reason SVM is often chosen as a benchmark to compare other non-linear models to. Also, there is a growing number of novel and hybrid approaches, combining the advantages of various methods using for example evolutionary optimization, methods of computational geometry and other techniques (e.g. Refs. [9,10]).

In this paper we apply the string model and approaches described in Ref. [11] to the real finance forex market. This is an extension of the previous work [11] into the real finance market. We derive two models for predictions of EUR/USD prices on the forex market. This is the first attempt for real application of string theory in the field of finance, and not only in high energy physics, where it is established very well. Firstly we describe briefly some connections between these different fields of research.

We would like to transfer modern physics ideas into the neighboring field called econophysics. The physical statistical viewpoint has proved to be fruitful, namely in the description of systems where many-body effects dominate. However, the standard, accepted by physicists, bottom-up approaches are cumbersome or outright impossible to follow the behavior of complex economic systems, where autonomous models encounter intrinsic variability.

The modern digital economy is founded on data. Our primary motivation comes from the actual physical concepts [12,13]; however, our realization differs from the original attempts in various significant details. Similarly as with most scientific problems, the representation of data is the key to efficient and effective solutions. The string theory development over the past 25 years has achieved a high degree of popularity among physicists [14].

The underlying link between our approach and string theory may be seen in switching from a local to a non-local form of data description. This line passes from the single price to the multivalued collection, especially the string of prices from the temporal neighborhood, which we term here as the string map. It is the relationship between more intuitive geometric methods and financial data. Here we work on the concept that is based on projection data into higher dimensional vectors in the sense of the works [15,16].

The present work exploits time series which can build the family of string-motivated models of boundary-respecting maps. The purpose of the present data-driven study is to develop statistical techniques for the analysis of these objects and moreover for the utilization of such string models onto the forex market. Both of the string prediction models in this paper are built on the physical principle of the invariance in time series of the forex market. Founding of a stationary state in the time series of the market was studied in Ref. [17].

2. Definition of the strings

By applying standard methodologies of detrending we suggest to convert the original series of quotations of the mean currency exchange rate $p(\tau)$ onto a series of returns defined by

$$\frac{p(\tau+h)-p(\tau)}{p(\tau+h)}, \tag{1}$$

where h denotes a tick lag between currency quotes $p(\tau)$ and $p(\tau+h)$, τ is the index of the quote. The mean $p(\tau) = (p_{\mathrm{ask}}(\tau)+p_{\mathrm{bid}}(\tau))/2$ is calculated from $p_{\mathrm{ask}}(\tau)$ and $p_{\mathrm{bid}}(\tau)$.

In the spirit of string theory it would be better to start with the 1-end-point open string map

$$P^{(1)}(\tau,h) = \frac{p(\tau+h)-p(\tau)}{p(\tau+h)}, \quad h \in \langle 0, l_s \rangle \tag{2}$$

where the superscript (1) refers to the number of endpoints.

The variable h may be interpreted as a variable which extends along the extra dimension limited by the string size l_s. For the natural definitions of the string to be fulfilled the boundary condition

$$P^{(1)}(\tau,0) = 0, \tag{3}$$

holds for any tick coordinate τ. We want to highlight the effects of rare events. For this purpose, we introduce a power-law Q-deformed model

$$P_q^{(1)}(\tau,h) = \left(1 - \left[\frac{p(\tau)}{p(\tau+h)}\right]^Q\right), \quad Q > 0. \tag{4}$$