Available online at www.sciencedirect.com
SCIENCE DIRECT®
Physica A 336 (2004) 538–548
PHYSICA A
www.elsevier.com/locate/physa

Robust methods for stock market data analysis

I. Antoniou[a,b,*], P. Akritas[a,b,c], D.A. Burak[d], V.V. Ivanov[a,b,e], A.V. Kryanev[a,d], G.V. Lukin[d]

[a] *International Solvay Institutes for Physics and Chemistry, CP-231, ULB, Bd. du Triomphe, 1050 Brussels, Belgium*
[b] *Department of Mathematics, Chaos and Innovation Research Unit, Aristoteles University of Thessaloniki, 54124 Thessaloniki, Greece*
[c] *Institut Superieur de Technologie, 6 rue Richard Coudenhove-Kalergi, L-1359 Luxembourg*
[d] *Moscow Engineering and Physical Institute, 115409 Moscow, Russia*
[e] *Laboratory of Information Technologies, Joint Institute for Nuclear Research, 141980 Dubna, Russia*

Received 31 October 2003

Abstract

We consider the problem of extraction of trend and chaotic components from irregular stock market time series. The proposed methods also permit to extract a part of chaotic component, the so-called anomalous term, caused by the transient short-time surges with high amplitudes. This provides more accurate determination of the trend component. The methods are based on the M-evaluation with decision functions of Huber and Tukey type. The iterative numerical schemes for determination of trend and chaotic components are briefly presented, resulting in an acceptable solution within a finite number of iterations. The optimal level for extraction of the chaotic component is determined by a new numerical scheme based on the fractal dimension of the chaotic component of the analyzed series. Forecasting from the realized part of the analyzed series and a priori expert information is also discussed.

PACS: 89.65.Gh; 05.45.Tp; 02.05.−r

Keywords: Stock market; Robust methods; Analysis and forecasting

* Corresponding author. Department of Mathematics, Chaos and Innovation Research Unit, Aristoteles University of Thessaloniki, Thessaloniki 54124, Greece. Fax: +30-2310-997929.
E-mail address: iantonio@math.auth.gr (I. Antoniou).

0378-4371/$ - see front matter
doi:10.1016/j.physa.2003.12.052

1. Introduction

The problem of stock market series analysis aiming to identify the structural changes in dynamics of the underlying process and eventually to predict the time series behavior in the future is a very difficult task. The processes underlying stock market data, noise level and time series volatility are "regime shifting", i.e., non-stationary. The commonly accepted viewpoint is that this problem is compatible to inventing a *perpetuum mobile* or solving problems like the *quadrature of the circle* [1].

A first step in stochastic time series analysis is the decomposition of the time series $x(t)$ as a sum of a predictable, deterministic and a chaotic component:

$$x(t) = \tilde{y}_{det}(t) + \tilde{y}_{ch}(t) , \tag{1}$$

where $\tilde{y}_{det}(t)$ is the deterministic (or trend) component and $\tilde{y}_{ch}$ is the chaotic component [2–4].

The deterministic component reflects the time-series changes due to the influence of some defined causes which may not be clear enough. However, as a rule their cumulative influence can be predictable during relatively long periods of time. In that case we have the possibility of forecasting.

The chaotic or stochastic part usually concerns a high frequency "*noise*" where the successive elements are practically uncorrelated. This means that this component is not predictable.

Besides, the analyzed series may also involve the anomalous component $y_{an}(t)$ reflecting the structural changes in the process dynamics. In this case (1) becomes

$$x(t) = y_{det}(t) + y_{an}(t) + y_{ch}(t) . \tag{2}$$

The terms $y_{det}(t)$, $y_{ch}(t)$ in (2) may significantly differ from the corresponding components in (1).

Unfortunately, traditional methods do not allow the effective extraction of the anomalous term. Moreover, various traditional methods for extraction of chaotic components may lead to the results which are qualitatively different (see, for example Ref. [5]).

Moreover, traditional methods for the determination of the trend component are not stable with respect to short-time surges of a high amplitude. This may result in significant distortion of the trend component itself and not permit to extract correctly the anomalous component $y_{an}(t)$ that very often is of interest.

The aim of this work is the creation of effective numerical schemes for estimation of trend, chaotic and anomalous components for satisfactory forecasting taking also into account a priori expert information. In Section 1 we describe the numerical schemes for extracting of deterministic and chaotic terms using expansions in orthogonal polynomials. The numerical schemes for determination of trend and chaotic terms on the basis of robust linear splines are presented in Section 2. Section 3 is devoted to the optimal determination of the chaotic component based on its fractal dimension. Section 4 presents first results concerning the application of the forecasting scheme which uses a priori expert estimation.