Mathematica Aeterna, Vol. 6, 2016, no. 2, 261 - 280

Time Series Analysis for a $1/t^\beta$ Memory Function and Comparison with the Lyapunov Exponent using Volatility Scaling

Paddy Walsh
ESB International Ltd., 18-21 St. Stephens's Green, Dublin 2, Ireland.
Dublin Institute of Technology, Kevin Street, Dublin 8, Ireland.

Jonathan Blackledge
DVC (Research), University of KwaZulu-Natal, South Africa.
SFI Stokes Professor, Dublin Institute of Technology, Ireland.

Abstract

Being able to provide accurate forecasts on the trending behaviour of time series is important in a range of applications involving the real-time evolution of signals, most notably in financial time series analysis, but control engineering in general. This paper reports on the use of an indicator that is based on a Memory Function of the form $\sim 1/t^\beta,\ \ \beta > 0$, and, in terms of a comparative analysis, the Lyapunov Exponent λ coupled with an approach whereby both parameters (i.e. λ and $\beta - 1$) are scaled according to the corresponding Volatility σ of the time series. A 'back-testing' procedure is used to evaluate and compare the performance of the indices $(\beta - 1)/\sigma$ and λ/σ for forecasting and quantifying trends over a range of time scales. However, in either case, a critical solution for providing high accuracy forecasts is the filtering operation used to identify the position in time at which a trend occurs subject to a time delay factor that is inherent in the filtering strategy used. The paper explores this strategy and presents some example results that provide a quantitative measure of the accuracy obtained.

Mathematics Subject Classification:
91B84, 91B02, 91B26, 91B64, 91B70, 91B82

Keywords:
Financial Time Series Analysis, Generalised Kolmogorov-Feller Equation, Memory Functions, Lyapunov Exponent, Volatility Scaling, Filtering Strategies.

1 Introduction

Continuous Time Random Walk models are important in developing algorithms for both simulating and analysing time series data. Most models of this type are based on Einstein's evolution equation, and, on the basis of this equation, we consider a continuous time solution based on a Memory Function of the form $1/t^\beta$ and show that the first order fundamental solution is equivalent to considering a Lévy distributed system with a $\delta(t)$ memory function. This is the subject of Sections 2 and 3, respectively. The purpose of this is to understand the origins of parameters such as $\beta > 0$, the Lévy index $\gamma \in (0, 2]$ and, for comparison, the Lyapunov Exponent λ (which is briefly considered in Section 3), for example, in terms of their use as indicators in regard to the analysis of an evolving time series. In this context, we explore the use of these parameters for forecasting both the type and stability of time series trends by scaling them with the corresponding Volatility. For this purpose, we provide a short derivation of the short time Volatility in Section 4 using a 'Phase Only Condition'. Coupled with a novel filtering strategy, it is shown that the key to forecasting the onset of time series trends are the zero crossings associated with $\beta - 1$, $(1/\gamma) - 1$ and λ, in the latter case, for example, the demarcation between an upward and a downward trend being predicated on whether $\lambda > 0$ or $\lambda < 0$, respectively. This is discussed in Section 5 which includes example results on trend forecasting for an energy commodities time series and quantifies the accuracy of the forecasts obtained using a back-testing procedure presented in Section 5.1.

The work reported in this paper can be contextualised in terms of the application of memory functions to complex systems analysis. In complex systems, the elements adapt to the aggregate pattern they co-create. As the components react, the aggregate changes, as the aggregate changes the components react anew. Barring the reaching of some asymptotic state or equilibrium, complex systems keep evolving to producing stochastic fields. Such systems arise naturally in an economy. Economic agents, be they banks, firms, or investors, continually adjust their market strategies to the macroscopic economy which their collective market strategies create. It is important to appreciate that there is an added layer of complexity within the economic community. Unlike many physical systems, economic elements (human agents) react with strategy and foresight by considering the implications of their actions, i.e. the decisions taken are subject to risk management and although we can not be certain whether this fact changes the resulting behaviour, we can be sure that it introduces feedback which is the basis for many complex systems models that generate chaotic fields with self-affine structures. This foresight is based to a certain extent on the 'memory' of past events and hence, the concept of defining a 'memory function' to analyse a complex system has a natural synergy with the underlying issues associated with modelling complex systems.

Complex systems can be split into two categories: equilibrium and non-equilibrium. Equilibrium complex systems, undergoing a phase transition, can lead to 'critical states' that often exhibit random self-affine structures in which the statistics of the stochastic field are scale invariant. Non-equilibrium complex systems give rise to 'self organised critical states' and financial markets can be considered to be non-equilibrium systems because they are constantly driven by transactions that occur as the result of new economic information over a range of time scales. They are complex systems because the market also responds to itself, often in a highly non-linear fashion, and would carry on doing so (at least for some time) in the absence of new information. The 'price change field' is highly non-linear and very sensitive to exogenous shocks and it is probable that all shocks have a long term effect for which a long term memory function may be sought. Market transactions generally occur globally at the rate of hundreds of thousands per second. It is the frequency and nature of these transactions that dictate the behaviour of stock market indices nearly all of which are statistically self-affine.

2 The Generalised Kolmogorov-Feller Equation

For a Probability Density Function (PDF) $p(x)$, Einstein's evolution equation is, [1]

$$u(x,t+\tau)=u(x,t)\otimes_x p(x) \tag{1}$$

where $u(x,t)$ is a 'density function' representing the concentration of a canonical ensemble of particles undergoing elastic collisions and $\otimes_x$ denotes the non-causal convolution integral. For arbitrary values of τ, Taylor expansion allows us to write

$$\tau\frac{\partial}{\partial t}u(x,t)+\frac{\tau^2}{2!}\frac{\partial^2}{\partial t^2}u(x,t)+...\equiv\tau m(t)\otimes_t\frac{\partial}{\partial t}u(x,t)=-u(x,t)+u(x,t)\otimes_x p(x) \tag{2}$$

where $m(t)$ is a Memory Function and $\otimes_t$ is taken to denote the causal convolution integral over t, convolution with a memory function replacing the infinite series representation through Taylor expansion of the function $u(x,t+\tau)$. Equation (2) is the Generalised Kolomogorov-Feller Equation (KFE) which reduces to the Classical KFE when $m(t)=\delta(t)$ and is equivalent to the case of considering $\tau<<1$ in the Taylor series expansion of equation (1), [2], [3].

2.1 Othonormality

For any memory function for which there exists a function or class of functions of the type $n(t)$, say, such that $n(t)\otimes_t m(t)=\delta(t)$ we can write equation (2) in the form

$$\tau\frac{\partial}{\partial t}u(x,t)=-n(t)\otimes_t u(x,t)+n(t)\otimes_t u(x,t)\otimes_x p(x)$$

where the Classical KFE is recovered when $n(t) = \delta(t)$. Any solution obtained to the Generalised KFE is dependent upon the choice of memory function $m(t)$ used. There are a number of choices that can be considered, each of which is taken to be a 'best characteristic' of a stochastic time series in terms of the influence of its time history. However, it may be expected that the time history of physically significant random systems is relatively localised in time.

We consider a memory function of the type [4]

$$m(t) = \frac{1}{\Gamma(1-\beta)t^{\beta}}, \quad \beta > 0 \ \text{ where } \ n(t) = \frac{1}{\Gamma(\beta-1)t^{2-\beta}}$$

given that

$$\int_0^{\infty} \frac{\exp(-st)}{\Gamma(\beta)t^{1-\beta}} dt = \frac{1}{s^{\beta}} \ \text{ and } \ \int_0^{\infty} \delta(t)\exp(-st)dt = 1$$

2.2 Fundamental (Green's Function) Solution

By writing equation (2) in the form

$$\tau\frac{\partial}{\partial t}u(x,t) + u(x,t) = u(x,t) - n(t)\otimes_t u(x,t) + n(t)\otimes_t u(x,t)\otimes_x p(x)$$

the Green's function solution is given by

$$u(x,t) = g(t)\otimes_t u(x,t) - g(t)\otimes_t n(t)\otimes_t u(x,t) + g(t)\otimes_t n(t)\otimes_t u(x,t)\otimes_x p(x), \quad (3)$$

the Green's function being given by

$$g(t) = \frac{1}{\tau}\exp(-t/\tau), \quad t > 0$$

which is the solution to

$$\tau\frac{\partial}{\partial t}g(t) + g(t) = \delta(t)$$

assuming initial conditions $u(x,0) = 0$ and $g(0) = 0$.

Theorem 2.1

If the Laplace transform of the function $n(t)$ exists, then a solution of equation (3) is

$$u(x,t) = h(t)\otimes_t u(x,t)\otimes_x p(x), \quad h(t) \leftrightarrow \frac{\bar{n}(s)}{\tau s + \bar{n}(s)}$$

where $\leftrightarrow$ denotes the Laplace transformation, i.e. the mutual transformation from t-space to s-space.

Proof of Theorem 2.1

Using the convolution theorems for the Fourier and Laplace transforms, respectively, equation (3) can be written as

$$\widetilde{\bar{u}}(k,s) = \bar{g}(s)\widetilde{\bar{u}}(k,s) - \bar{g}(s)\bar{n}(s)\widetilde{\bar{u}}(k,s) + \bar{g}(s)\bar{n}(s)\widetilde{\bar{u}}(k,s)\widetilde{p}(k)$$

where

$$\widetilde{\bar{u}}(k,s) = \int_0^\infty \int_{-\infty}^\infty u(x,t)\exp(-ikx)dx\exp(-st)dt, \quad \bar{g}(s) = \int_0^\infty g(t)\exp(-st)dt,$$

$$\bar{n}(s) = \int_0^\infty n(t)\exp(-st)dt \quad \text{and} \quad \widetilde{p}(k) = \int_{-\infty}^\infty p(x)\exp(-ikx)dx$$

Thus, noting that $\bar{g}(s) = (1+\tau s)^{-1}$, we can write

$$\widetilde{\bar{u}}(k,s) = -\frac{\bar{g}(s)}{1-\bar{g}(s)}\bar{n}(s)\widetilde{\bar{u}}(k,s) + \frac{\bar{g}(s)}{1-\bar{g}(s)}\bar{n}(s)\widetilde{\bar{u}}(k,s)\widetilde{p}(k)$$

$$= -\frac{\bar{n}(s)}{\tau s}\widetilde{\bar{u}}(x,t) + \frac{\bar{n}(s)}{\tau s}\widetilde{\bar{u}}(k,s)\widetilde{p}(k) = \bar{h}(s)\widetilde{\bar{u}}(k,s)\widetilde{p}(k)$$

which, upon inverse transformation yields

$$u(x,t) = h(t)\otimes_t u(x,t)\otimes_x p(x) \tag{4}$$

Theorem 2.2

For an initial solution $u_0(x,t)$, the convergent N^{th}-order iterative solution of equation (4) is

$$u_N(x,t) = \prod_{j=1}^{N}{}_{\otimes}\, p(x) \prod_{k=1}^{N}{}_{\otimes}\, h(t)\otimes_x \otimes_t u_0(x,t), \quad \|h(t)\|_2 \times \|p(x)\|_2 < \frac{1}{\sqrt{2\pi}} \tag{5}$$

where

$$\prod_{j=1}^{N}{}_{\otimes}\, f(t) \equiv f(t)\otimes_t f(t)\otimes_t f(t)\otimes_t ...$$

denotes the N^{th} convolution of $f(t)$.

Proof of Theorem 2.2

Consider an iteration of equation (4) as defined by

$$u_{n+1}(x,t) = h(t)\otimes_t u_n(x,t)\otimes_x p(x), \quad n = 1,2,...,N$$

Using the convolution theorem, the equivalent iteration in Fourier-Laplace space is

$$\bar{\widetilde{u}}_{n+1}(k,s) = \bar{h}(s)\bar{\widetilde{u}}_n(k,s)\widetilde{p}(k)$$

for initial solution $\bar{\widetilde{u}}_0(k,s)$. Thus, after N iterations, we can write

$$\bar{\widetilde{u}}_N(k,s) = [\bar{h}(s)]^N[\widetilde{p}(k)]^N\bar{\widetilde{u}}_0(k,s)$$

so that upon inverse Fourier-Laplace transformation, the result is obtained.

The criterion for convergence as stated in Theorem 2.2 is obtained by considering an n^{th} order error function $\epsilon_n(x,t)$ so that $u_n(x,t) = u(x,t) + \epsilon_n(x,t)$. We can then write

$$\bar{\widetilde{\epsilon}}_{n+1}(k,s) = \bar{h}(s)\widetilde{p}(k)\bar{\widetilde{\epsilon}}_n(k,s)$$

so that

$$\bar{\widetilde{\epsilon}}_n(k,s) = [\bar{h}(s)\widetilde{p}(k)]^n\bar{\widetilde{\epsilon}}_0(k,s)$$

and it is clear that, since we require $\bar{\widetilde{\epsilon}}_n \to 0$ and $n \to \infty$, $[\bar{h}(s)\widetilde{p}(k)] < 1 \quad \forall(k,s)$. The condition for convergence therefore becomes

$$\|\bar{h}(s)\widetilde{p}(k)\| \leq \|\bar{h}(s)\| \times \|\widetilde{p}(k)\| < 1$$

or, for Euclidian norms, and, using Rayleigh's theorem,

$$\|\bar{h}(s)\|_2 \times \|p(x)\|_2 < \frac{1}{\sqrt{2\pi}}$$

In (k,t)-space

$$\widetilde{\epsilon}_n(k,t) = \prod_{k=1}^{n}{}^{\otimes} \; h(t)[\widetilde{p}(k)]^n \otimes_t \widetilde{\epsilon}_0(k,t)$$

so that, using Hölder's inequality,

$$\|\widetilde{\epsilon}_n(k,t)\| \leq \| \prod_{k=1}^{n}{}^{\otimes} \; h(t)[\widetilde{p}(k)]^n\| \times \|\widetilde{\epsilon}_0(k,t)\| \leq \|h(t)\|^n \times \|\widetilde{p}(k)\|^n \times \|\widetilde{\epsilon}_0(k,t)\|$$

from which the condition for convergence is thus derived.

2.3 First Order Impulse Response Function

Form equation (5), if the initial solution is an impulse, i.e. $u_0(x,t) = \delta(x)\delta(t)$, then the Impulse Response Function (IRF), denoted by $R_N(x,t)$, say, is given by

$$R_N(x,t) = \prod_{j=1}^{N}{}^{\otimes} \; p(x) \prod_{k=1}^{N}{}^{\otimes} \; h(t)$$

with 'transfer function'

$$\tilde{\bar{R}}_N(k,s) = [\bar{h}(s)\tilde{p}(k)]^N$$

The first order IRF is

$$R_1(x,t) = p(x)h(t)$$

and it is then clear that the 'space integrated IRF' is given by $h(t)$, i.e.

$$h(t) = \int_{-\infty}^{\infty} R_1(x,t)dx, \quad \int_{-\infty}^{\infty} p(x)dx = 1$$

For memory function

$$m(t) \leftrightarrow \frac{1}{s^{1-\beta}}, \quad \bar{h}(s) = \frac{1}{1+\tau s^{\beta}} \sim \frac{1}{\tau s^{\beta}}, \quad \tau >> 1 \Rightarrow h(t) \sim \frac{1}{\tau \Gamma(\beta) t^{1-\beta}}$$

Thus, if we consider an initial solution $u_0(x,t) = \delta(x)r(t)$ where $r(t)$ is some random varying function, then we arrive at a model for $R(t)$ given by (ignoring scaling by $1/\tau$)

$$R(t) = \frac{1}{\Gamma(\beta)t^{1-\beta}} \otimes_t r(t)$$

This model is based on the Riemann-Liouville integral transform which has self-affine properties, properties that exhibit 'Stochastic Trending Characteristics'. In other words $R(t)$ defines a random fractal function whose IRF is $1/t^{1-\beta}$, a result that is, in light of the above analysis, been shown to be a PDF independent first order solution to the Generalised KFE for a memory function of the type (ignoring scaling) $1/t^{\beta}$.

2.4 Application to Stochastic Time Series Analysis

By considering a short time model for a time series which is predicated on the scaling law $t^{\alpha}, \ \alpha = \beta - 1$, regression methods can be used to estimate the parameter α on a moving window basis, a least squares estimate, for example, being given by, for a uniformly sampled time series $R(t_i) > 0 \forall t_i, \ \ i = 1, 2, ..., N$

$$\alpha = \frac{\sum\limits_{i=1}^{N} \ln R(t_i) \sum\limits_{i=1}^{N} \ln t_i - N \sum\limits_{i=1}^{N} \ln R(t_i) \ln t_i}{\left(\sum\limits_{i=1}^{N} \ln t_i\right)^2 - N \sum\limits_{i=1}^{N} \ln t_i^2}$$

In this context, and, from the point of view detecting trends in time series data $R(t_i)$, $\alpha > 0$ indicates a positive upward trend and $\alpha < 0$ indicates a negative downward trend. The case when $\alpha \sim 0$ is an indication a non-trending behaviour. The transition points between the start and end of any given trend can therefore be identified by those positions in time at which $\alpha \sim 0$ through the application of a moving window process where an estimate is computed at each position of the windowed data.

2.5 Equivalence with the Lévy Index

Consider a (symmetric) Lévy distribution $p(x)$ defined in terms of the Characteristics Function

$$\widetilde{p}(k) = \exp(-a \mid k \mid^{\gamma}) \sim 1 - a \mid k \mid^{k}, \quad \gamma \in (0, 2]$$

where $a << 1$ is a constant and γ is the 'Lévy index'. Application of a first order Taylor series to equation (1) under the condition that $\tau << 1$ and application of the Convolution Theorem then yields the Fractional Diffusion Equation

$$\left(\frac{\partial^{\gamma}}{\partial \mid x \mid^{\gamma}} - \frac{\partial}{\partial t}\right) u(x,t) = 0; \quad \frac{\partial^{\gamma}}{\partial \mid x \mid^{\gamma}} u(x,t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widetilde{u}(k,t) \mid k \mid^{\gamma} \exp(ikx) dk$$

where $\tau := \tau / a = 1$ (for normalised units).

Theorem 2.3

The Green's function for this equation has the operational form

$$g(x,t) = \frac{i}{2\Gamma(1/\gamma) t^{1-1/\gamma}} \otimes_t \left[\delta(t) + \sum_{n=1}^{\infty} \frac{(ix)^n}{n!} \frac{\partial^{n/\gamma}}{\partial t^{n/\gamma}}\right] = \frac{i}{2\Gamma(1/\gamma) t^{1-1/\gamma}}, \quad x \to 0$$

Proof of Theorem 2.2

For the source function $-\delta(x,t)$, computation of the Green's function by Fourier and Laplace transformation in x-space and t-space, respectively, yields the result (for the pole at $\mid k \mid = s^{1/\gamma}$)

$$g(x,t) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \frac{i \exp(is^{1/\gamma} x)}{2 s^{1/\gamma}} \exp(st) ds$$

Thus, by writing

$$\exp(is^{1/\gamma} x) = 1 + \sum_{n=1}^{\infty} \frac{(ix)^n}{n!} s^{n/\gamma}$$

the result of obtained, given that

$$s^{n/\gamma} \leftrightarrow \frac{\partial^{n/\gamma}}{\partial t^{n/\gamma}}$$

This result shows that the IRF for a Lévy distributed system is determined by the function $1/t^{1-1/\gamma}$ $t > 0$ for $x \to 0$, and, in this sense, it is clear that $\gamma = 1/\beta$ for the range $1/2 \leq \beta < \infty$.

3 The Lyapunov Exponent and Kolmogorov-Sinai Entropy

For a set of discrete time steps $t_n,\ n = 1, 2, ...$, equation (1) can be considered to be an Iteration Function System defined as

$$u(x_m, t_{n+1}) = p(x_m) \otimes_x u(x_m, t_n)$$

where the density function u is also cast in terms of a set of discrete steps in space $x_m,\ m = 1, 2, ...$ and $\otimes_x$ is taken to denote the 'Convolution Sum'. Suppose that after many time steps, this iteration converges to the function $\phi(x_m, t_\infty)$, say. We can then represent the iteration in the form

$$u(x_m, t_{n+1}) = \phi(x_m, t_\infty) + \epsilon(x_m, t_n)$$

where $\epsilon(x_m, t_n)$ denotes the error at any time step n. Convergence to the function $\phi(x_m, t_\infty)$ then occurs if $\epsilon(x_m, t_n) \to 0\ \forall\ m$ as $n \to \infty$.

Consider a 'model' for the error at each time step given by (for some real constant ε)

$$\epsilon(x_m, t_{n+1}) = \varepsilon \exp(t_n \lambda) \Rightarrow \epsilon(x_m, t_{n+1}) = \epsilon(x_m, t_n) \exp(\lambda)$$

We can then construct an equation for λ given by

$$\lambda = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \ln \left| \frac{\bar{\epsilon}(t_{n+1})}{\bar{\epsilon}(t_n)} \right| \quad \text{where} \quad \bar{\epsilon}(t_n) = \lim_{M\to\infty} \frac{1}{M} \sum_{n=1}^{M} \epsilon(x_m, t_n)$$

If λ is negative, then the iterative process is stable since we can expect that as $N \to \infty$, $\bar{\epsilon}(t_{n+1})/\bar{\epsilon}(t_n) < 1$ and thus $\log[\bar{\epsilon}(t_{n+1})/\bar{\epsilon}(t_n)] < 0$. However, if λ is positive then the iterative process will diverge. This criterion for convergence/divergence is of course dependent on the exponential 'model' used to represent the error function at each iteration, and, within this context, λ is known as the Lyapunov Exponent. Note that we have purposely derived an expression for this exponent in regard to the evolution equation - equation (1) - in order to demonstrate the connectivity between the parameter λ and the parameters $\beta - 1$ and $1 - 1/\gamma$ (as discussed in Sections 2.4 and 2.5, respectively). In this sense, the 'unifying framework' for all such parameters is equation (1).

3.1 The Lyapunov Exponent

In a general context, the Lyapunov Exponent, denoted by $\lambda(t_0)$ for some initial condition in time t_0, is a quantitative measure of the exponential divergence of trajectories starting from the neighbourhood of t_0 [5]. For a one-dimensional system, i.e. a time series function $f(t)$, say,

$$|f_n(t_0+\varepsilon) - f_n(t_0)| = \varepsilon \exp[n\lambda(t_0)],$$

where ε is a small perturbation from the initial condition t_0 and n is the number of iterations. Generally, λ depends on the initial condition, and thus we can only estimate its average value. In a measure-preserving system λ is constant for all trajectories and is given by the limit

$$\lambda(t_0) = \lim_{n\to\infty} \lim_{\varepsilon\to 0} \frac{1}{n} \log \left| \frac{f_n(t_0+\varepsilon) - f_n(t_0)}{\varepsilon} \right|$$

or

$$\lambda(t_0) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \log |f'(t_k)| = \lim_{n\to\infty} \frac{1}{n} \log \prod_{k=1}^{n} |f'(t_k)|$$

For each k, $f'(t_k)$ tells us how much the function f is changing with respect to its argument at the point t_k. This derivative expresses the magnitude of change in the transition from t_k to t_{k+1}. The limit of the average of the derivative logarithms over n iterations is taken to provide a measure of how fast the orbit changes as (discrete) time propagates. In the context of a discrete stochastic time series $R(t_1), R(t_2), ..., R(t_N)$, computation of the exponent is typically undertaken using the result

$$\lambda = \frac{1}{N} \sum_{n=1}^{N} \log \left| \frac{R(t_{n+1})}{R(t_n)} \right| \qquad (6)$$

and provides a method of indicating the trending behaviour of a time series where positions in time at which the exponent crosses zero are an indication of the transition points between the start and end of a trend.

3.2 Kolmogorov-Sinai Entropy and Lyapunov Exponents

For a d-dimensional system we have a set $\lambda = \{\lambda_1, \ldots, \lambda_d\}$ and more complex behaviour, but still qualitatively the same as the one dimensional case [6]. However, a more accurate measure is the Kolmogorov-Sinai (KS) entropy because it considers the resolution (the precision) under which the system is observed. The Lyapunov Exponents measure how fast we loose the capability to predict the behaviour of a stochastic system. The disadvantage is that this measure does not consider the resolution under which the system is observed, unlike the Kolmogorov-Sinai entropy [5], [7] and [8].

Let the partition $\beta = \{T_1, T_2, ..., T_m\}$ be the observer's resolution. Looking at the system state t, the observer can only determine the fact that $t \in T_i$ and reconstruct the symbolic trajectory $\alpha_n = \{s_{m_1}, s_{m_2}, \ldots, s_{m_n}\}$ corresponding to the regions visited. The entropy of a trajectory α_n with respect to partition β is given by

$$H_n^\beta = -\sum_{\alpha_n} \Pr(\alpha_n) \log_{|\mathcal{A}|} \Pr(\alpha_n)$$

where $\Pr(\alpha_n)$ is the probability of occurrence of the substring α_n. The conditional entropy of the $(n+1)^{\text{th}}$ symbol provided the previous n symbols are known is defined as

$$h_n^\beta = h_{n+1|n}^\beta = \begin{cases} H_{n+1}^\beta - H_n^\beta, & n \geq 1; \\ H_1^\beta, & n = 1. \end{cases}$$

The entropy for a partition β is given by

$$h^\beta = \lim_{n\to\infty} h_n^\beta = \lim_{n\to\infty} \frac{1}{n} H_n^\beta$$

and the Kolmogorov-Sinai entropy is the supremum over all possible partitions

$$h_{KS} = \sup_\beta h^\beta$$

The KS entropy is zero for regular systems, is finite and positive for a deterministic chaos and infinite for a random process and is related to the Lyapunov exponents by

$$h_{KS} = \sum_{1\leq d\leq D} \lambda_d$$

being proportional to the time horizon T on which the system is predictable.

The behaviour of the Lyapunov Exponent when applied to a time series is similar to that of the 'α-index' defined in Section 2.4. Both parameters are related to other metrics associated with the fields of non-Gaussian stochastic systems in regard to the equivalence of the numerical ranges to be expected that differentiate between persistent and anti-persistent behaviour. In this context, Table 1 quantifies the comparative ranges of some selected metrics (subject to their upper and lower bounds).

α-Index	Levy Index	Lyapunov Exponent	Hurst Exponent	Fractal Dimension
$\alpha = 0$	$\gamma = 2$	$\lambda = 0$	$H = \frac{1}{2}$	$D = 1.5$
$\alpha > 0$	$\gamma < 1$	$\lambda > 0$	$H > \frac{1}{2}$	$D < 1.5$
$\alpha < 0$	$\gamma > 1$	$\lambda < 0$	$H < \frac{1}{2}$	$D > 1.5$

Table 1: Equivalence of the numerical ranges associated with the α-parameter, the Lévy index ($0 < \gamma \leq 2$) and the Lyapunov Exponent with the Hurst exponent ($0 < H < 1$) and the Fractal Dimension ($1 < D < 2$). .

4 The Volatility and Volatility Scaling

Consider the short time rate equation

$$f(t) = \frac{d}{dt} \ln R(t) = \sigma u(t)$$

where σ is the 'Volatility' and $u(t)$ is a stochastic signal. The Volatility is a measure of the randomness associated with $R(t)$ and we therefore require an estimate for σ in terms of the function $R(t)$ through the elimination $u(t)$. This requires a conditional estimate of σ to be formulated.

4.1 Phase Only Condition

Theorem 4.1

If $u(t)$ is a phase only function with unit amplitude, bandwidth Ω and compact support T then

$$\sigma = \sqrt{\frac{2\pi}{\Omega}} \|f(t)\|_2, \quad \|f(t)\|_2 := \left(\int_{-T/2}^{T/2} \mid f(t) \mid^2 dt \right)^{\frac{1}{2}} \tag{7}$$

Proof of Theorem 4.1

If $u(t)$ is a phase only function (assuming unit amplitude), then

$$\widetilde{u}(\omega) = \exp[i\theta(\omega)]$$

where $\theta(\omega)$ is the 'Phase Spectrum'. Noting that $\sigma \|u(t)\|_2 = \|f(t)\|_2$, using Parseval's Theorem, we have

$$\int_{-T/2}^{T/2} \mid u(t) \mid^2 dt = \frac{1}{2\pi} \int_{-\Omega/2}^{\Omega/2} \mid \widetilde{u}(\omega) \mid^2 d\omega = \frac{\Omega}{2\pi}$$

from which the result is thus derived.

Theorem 4.2

For a uniformly sampled discrete function $f_(t_n), \quad n = 1, 2, 3, ..., N,$ equation (7) becomes

$$\sigma = \frac{2\pi}{\Omega}(\|f(t_n)\|_2, \quad \|f(t_n)\|_2 := \left(\sum_{n=1}^{N} \mid f(t_n) \mid^2\right)^{\frac{1}{2}}$$

Proof of Theorem 4.2

For a uniform sampling interval of Δt, say, the discrete version of equation (7) is

$$\sigma = \sqrt{\frac{2\pi\Delta t}{\Omega}}\|f(t_n)\|_2$$

The sampling interval Δt of $f(t_n)$ is related to the sampling interval $\Delta\omega$ of the Discrete Fourier Transform of $f(_n)$ by the equation

$$\Delta t \Delta\omega = \frac{2\pi}{N}$$

and since the bandwidth of the discrete spectrum of f_n is $N\Delta\omega$ is is clear that $\Delta t = 2\pi/\Omega$ from which the result is thus derived.

Corollary

The scaling constant $2\pi/\Omega$ can be used to define a re-scaled the Volatility given by $\sigma := \sigma\Omega/2\pi$ thereby yielding the expression

$$\sigma = \left[\sum_{n=1}^{N-1} \left|\ln\left(\frac{R(t_{n+1})}{R(t_n)}\right)\right|^2\right]^{\frac{1}{2}} \tag{8}$$

Comparing equation (8) with equation (6), we observe a similarity in both forms with regard to the commonality of the quotient $R(t_{n+1})/R(t_n)$ and the logarithmic operation but where $\lambda < 0$ or $\lambda \geq 0$ but where $\sigma \geq 0 \forall n$.

4.2 Volatility Scaling

The zero crossings associated with computing of the α-index (as discussed in Section 2) and/or the Lyapunov exponent (Section 3) on a moving window basis provides the positions in time where there is a transition in the trend type. The value of the Volatility indicates the 'stability' of the time series, the temporal characteristics of all indicators being dependent of the size of the window or 'period' used. This suggests scaling the indices with the Volatility, i.e. computing the quotient $\alpha_\sigma = \alpha/\sigma$ and $\lambda_\sigma = \lambda/\sigma$, in order to assess not only changes in the direction of a trend but the corresponding stability of that trend. This idea has obvious applications to a range of time series but especially in regard to financial time series analysis where forecasting both the type and characteristics of a trend is of fundamental importance, a positive trend with low volatility indicating a good investment horizon, for example.

5 Filtered Zero-Crossings Analysis

On the basis of the ideas considered in the previous section, the critical points at which a trend forecasting decision is made are the zero crossing points associated with λ_σ (or α_σ). We consider the case associated with parameter λ_σ but note that the arguments and analysis presented in this section applies in equal measure to the parameter α_σ, a comparison of the two parameters being given later on in Section 5.2.

By computing $\lambda_\sigma(t)$ on a moving window basis where t is the position in time of the window, identification of the zero crossings denoted by the function $z_c(t)$ involves the follow basic procedure:

$$z_c(t) = \begin{cases} +1, & \lambda_\sigma(t) < 0 \;\; \& \;\; \lambda_\sigma(t+\varepsilon) \geq 0; \\ -1, & \lambda_\sigma(t) > 0 \;\; \& \;\; \lambda_\sigma(t+\varepsilon) \leq 0; \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

where ε is a small perturbation in time. This procedure generates a series of Kronecker delta functions whose polarity determines the position(s) in time at which a trend is expected to be positive or negative. Thus the function $z_c(t)$ identifies the zero crossings associate with the end of an upward trend and the start of a downward trend as is the case when $z_c(t) = -1$ and the end of downward trend and the start of an upward trend as the case $z_c(t) = +1$. This is a 'critical indicator' in regard to forecasting the trending behaviour of a time series, and, with regard to the back-testing algorithm to be discussed in the following section, is the primary performance evaluator.

In practice, the points at which the zero crossings are evaluated according equation (9) depend on the accuracy of the algorithm used to compute λ_σ which in turn, depends on the intrinsic noise associated with the time series used. This can yield errors in the positions at which the zero-crossings are computed especially with regard to changes associated with very short time micro-trends. In the context of longer term macro-trends, such micro-trends may legitimately be interpreted as noise although, in the context of financial times series analysis, for example, the term 'noise' must be understood to include legitimate price values. To overcome this effect, and, since the same argument applies to the computation of the Volatility, λ_σ is filtered thus:

$$\Lambda_\sigma(\tau) = \frac{1}{T} \int_{-T/2}^{T/2} w(t+\tau)\lambda_\sigma(t)dt \tag{10}$$

where

$$w(t) = \begin{cases} 1, & \mid t \mid \leq T/2 \\ 0, & \mid t \mid > T/2 \end{cases}$$

which defined a basic moving average filter. However, in the context of an evolving time series such as financial time series, a filtering strategy must be used given that

$$\lambda_\sigma(t) \exists \forall t \in (-\infty, \xi];\ \lambda_\sigma(t) = 0 \forall t > \xi \tag{11}$$

where ξ defines the finite extent of the time series at any point in time beyond which the series is undefined and can thereby be set to zero as given in condition (11) . For this reason, we consider a modification to equation (10) and write

$$\Lambda_\sigma(\tau) = \frac{1}{T} \int_{-T/2}^{T/2} w(t+\tau)\lambda_\sigma(t+\delta T)dt, \quad \delta \in [0,1] \tag{12}$$

The parameter δ defines a 'shift' which provides control over the extent to which the evolving time series is 'zero padded'. This parameter provides a critical difference in the accuracy of the results obtained through back-testing in terms of the reliability and accuracy of the function $z_c(t)$ to indicate the nature and duration of a trend. The most successful results occur when the 'shift' is increasingly closer to 0.5. However, in the context of an analysis of a financial time series (as discussed in the following section) this introduces a need for a trader to 'hold their position' until the position of a zero crossing has 'stabilised' (subject to the filter applied), thereby retaining a fixed position in the interim. This requirement introduces a 'Trading Delay Factor' given by δT. Thus an optimum value for the parameter δ is the smallest value that yields a maximum back-testing accuracy.

5.1 Back-testing Results

Back-testing algorithms are designed to 'gauge' the accuracy of the results in terms of trend predictions and are usually but not exclusively related to testing a strategy for forecasting the behaviour of financial time series. In this context, the procedure operates on the basis that the price difference should be positive if the interval between the start and end points of a predicted trend are correct. In those cases where this occurs throughout the duration of the time series considered, the predicted entry and exit points are taken to be correct, else, they are taken to be incorrect. The accuracy associated with the back-test is then computed as a percentage in terms of the predictions being correct and only correct.

The algorithm for the back-tests used is compounded in the following procedure as predicated on the computation of $z_c(t_i)$ - equation (9) - using the filtered index $\Lambda_\sigma(t_i)$, *a priori* where R_+^+, R_+^-, R_-^- and R_-^+ denote a correct positive (up-ward) trend prediction, an incorrect positive trend prediction, a correct negative (down-ward) trend prediction and an incorrect negative trend prediction, respectively, in regard to the discrete time series data $R(t_i)$ used.

$R_+^+ = 0;\ \ R_+^- = 0;\ \ R_-^- = 0;\ \ R_-^+ = 0;$

$k = 0;\ \ \forall i,$
if $z_c(t_i) < 0$, then $\hat{\Lambda}_\sigma(t_k) = \Lambda_\sigma(t_i)$ & $\hat{z}_c(t_k) = z_c(t_i);\ \ k = k + 1;$
if $z_c(t_i) > 0$, then $\hat{\Lambda}_\sigma(t_k) = \Lambda_\sigma(t_i)$ & $\hat{z}_c(t_k) = z_c(t_i);\ \ k = k + 1;$

$\forall j,$
if $\hat{z}_c(t_j) - \hat{z}_c(t_{j+1}) > 0$ & $\hat{\Lambda}_\sigma(t_j) - \hat{\Lambda}_\sigma(t_{j+1}) < 0$, then $R_-^+ = R_-^+ + 1;$
if $\hat{z}_c(t_j) - \hat{z}_c(t_{j+1}) > 0$ & $\hat{\Lambda}_\sigma(t_j) - \hat{\Lambda}_\sigma(t_{j+1}) > 0$, then $R_-^- = R_-^- + 1;$
else
if $\hat{z}_c(t_j) - \hat{z}_c(t_{j+1}) < 0$ & $\hat{\Lambda}_\sigma(t_j) - \hat{\Lambda}_\sigma(t_{j+1}) < 0$, then $R_+^+ = R_+^+ + 1;$
if $\hat{z}_c(t_j) - \hat{z}_c(t_{j+1}) < 0$ & $\hat{\Lambda}_\sigma(t_j) - \hat{\Lambda}_\sigma(t_{j+1}) > 0$, then $R_+^- = R_+^- + 1;$

if $R_-^+ + R_-^- > 0$ then $S = 100R_-^-/(R_-^+ + R_-^-)$ else $S = 0;$
if $R_+^+ + R_+^- > 0$ then $L = 100R_+^+/(R_+^+ + R_+^-)$ else $L = 0;$

where L and S denote the percentage accuracy of going 'Long' (e.g. making an investment predicated on the forecast of an increasing price trend) and going 'Short' (e.g. selling an investment predicated on the forecast of a decreasing price trend). Note that this procedure (at least for $\delta > 0$) is applied to the filtered time series data and represents the accuracy associated a 'Delayed Call'.

5.2 Example Result using Energy Commodities Time Series Data

Figure 1 shows the output associated with a normalised sample of the energy commodity *Brent Crude Daily*, i.e. 1000 samples of Brent Crude Oil price sampled on a daily basis, which exhibits a range of short and long time-scale trends. The plot also shows the filtered time series $\Lambda_\sigma(t_i)$. The period used to compute $\lambda_\sigma(t_i)$ and $\Lambda_\sigma(t_i)$ is 25, with $\delta = 0.4$ yielding a Trading Delay Factor of floor$(\delta T) = 10$. For this case, $S = 50\%$ and $L = 44\%$ assuming a zero Trading Delay Factor which yields and 'Instantaneous Call'. Using a 'Delayed Call', $S = 78\%$ and $L = 70\%$ thereby providing a higher combined percentage accuracy. Figure 2 quantifies the difference between the two cases illustrating the shift that occurs in regard to applying a 'Delayed Call'.

In comparison, Figure 3 shows the effect of applying exactly the same computational procedure to the parameter α_λ using the least squares estimate given in Section 2.4. In this case, the 'Instantaneous Call' yields $S = 67\%$ and $L = 50\%$ but the 'Delayed Call' yields $S = 67\%$ and $L = 64\%$. In general, application of the (filtered) λ_σ-index yields a better performance over application of the (filtered) α_σ-index. This is despite the expense of the higher computational overheads required to implement regression analysis. On the other hand, regression analysis generates a smoother time signature which may be of benefit in forecasting the indicator. It is also clear from Figure 3 that, in comparison to the (filtered) λ_σ-index the (filtered) α_σ-index has a greater dynamic range. This is of value in regard to assessing the stability of a trend when the amplitude of α_σ increases. From the point-of-view of assessing the potential for increasing an existing investment, for example, this amplitude provides a confidence measure whose quantification lies beyond the scope of this work.

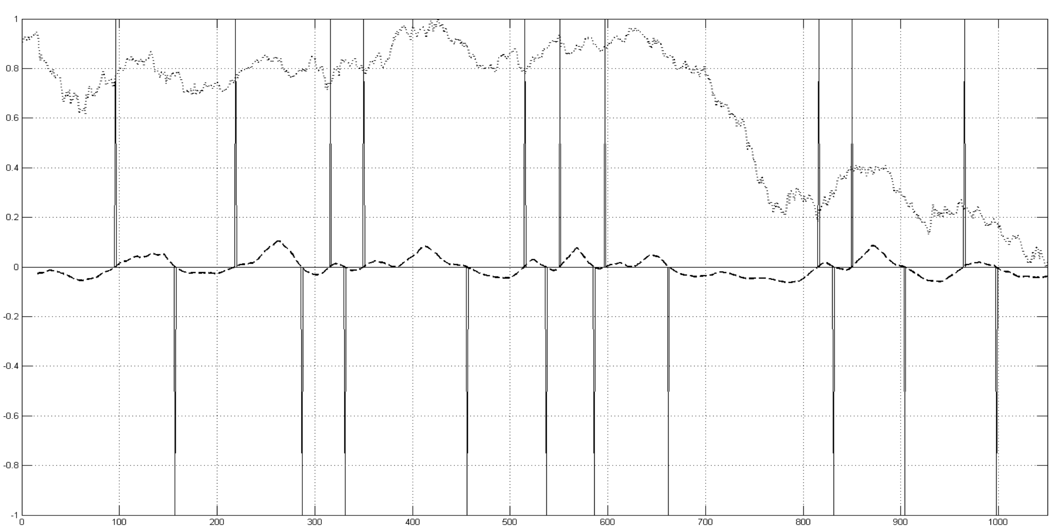

Figure 1: Normalised Brent Crude Daily time series data (Dotted), $\Lambda_\sigma(t_i)$ (Dashed) and the associated Zero Crossings Indicator (Solid).

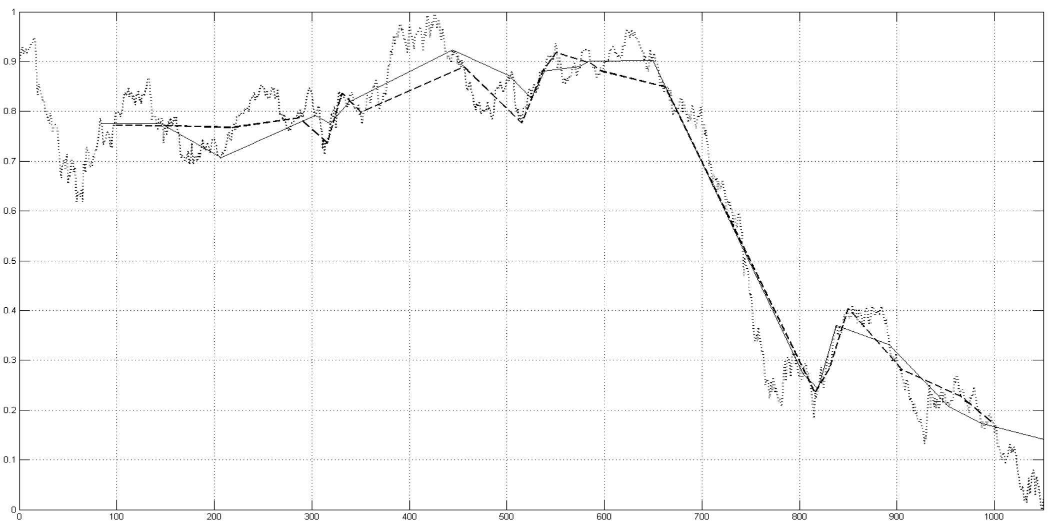

Figure 2: Normalised Brent Crude Daily time series data (Dotted), the predicted trend profile associated with an 'Instantaneous Call' (Solid) and the predicted trend profile associated with a 'Delayed Call' (Dashed).

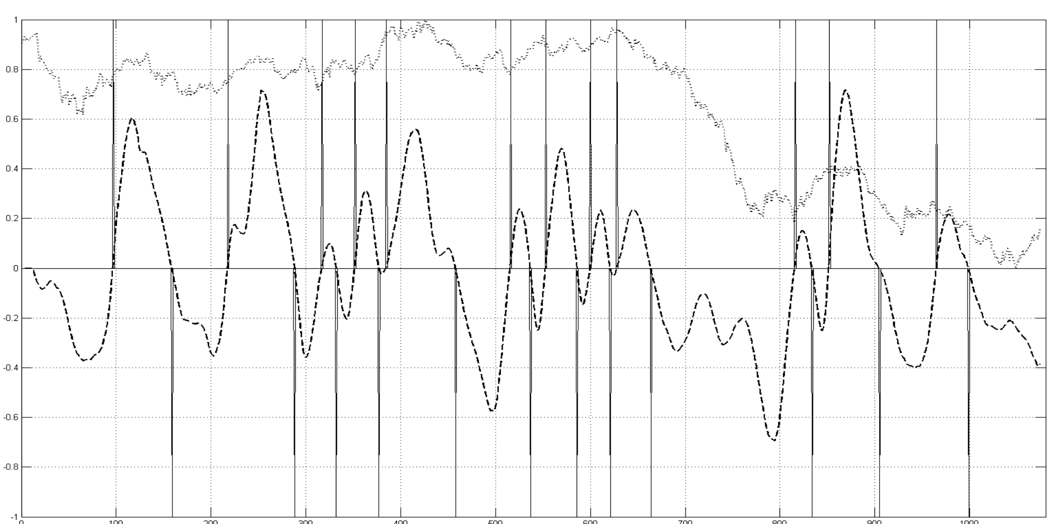

Figure 3: Normalised Brent Crude Daily time series data (Dotted), filtered $\alpha_\sigma(t_i)$ (Dashed) and the associated Zero Crossings Indicator (Solid).

6 Conclusions and Discussion

The purpose of this paper has been to introduce a time series model based on Continuous Time Random Walk Models as derived from Einstein evolution equation - equation (1). In Section 2, we have shown that the first order temporal IRF associated with the Green's function solution to the Generalised KFE for a Memory Function of the type $1/t^{\beta}$ is given by $1/t^{1-\beta}, \quad t > 0 \ \ \beta > 0$ which is independent of the PDF. This result is compatible with a Lévy Distributed PDF for a $\delta(t)$ Memory Function under the asymptotic condition $x \to 0$ (as discussed in Section 2.5). Within the unifying context of equation (1) we have also considered the Lyapunov Exponent and identified the respective numerical ranges associated with differentiating between the persistent and anti-persistent behaviour of a time series (upward and downward trends, respectively) as provided in Section 3 and quantified in Table 1.

The application of the filtering operation discussed in Section 5 is crucial to the success of the 'predictive power' of the indices considered. Without application of this filter the accuracy of determining the correct zero crossings for going Long or Short remains relatively poor as predicated on the back-testing procedure considered in Section 5.1. Moreover, to yield the required accuracy necessary for most trading purposes, the filter must be applied with an appropriate shift. For the case of $\delta > 0$, the filter generates a result that is analogous to a forward-error-correction scheme. Without filtering, back-testing shows that the Long/Short accuracy is $\sim 50\%$. With the implementation of the filter (for $\delta \in [0.4, 0.7]$), the accuracy is typically $\sim 70\%$++. The price that is paid for this accuracy is the delay required before application of a Long/Short Call. While this delay does not reduce accuracy in the prediction of a trend, its implementation yields a lower price difference between the entry and exit points. This dictates that smaller windows are used in the computation of an index subject to the back-testing accuracy obtained for a given time series.

In addition to the applications associated with financial time series analysis, which has been a focus of the results presented in this paper, the indices considered can be used for forecasting the behaviour of any evolving time series. Scaling by the Volatility yields a measure of stability, and, in this sense, the indices can be used to partition a time series into 'stable' and 'non-stable' regions. In the former case, short-time forecasting techniques may be used with increasing confidence such as those based on the application of evolutionary computing methods, for example, to predict the future value(s) of a time series as opposed to the predicting the trend alone. Finally, it is noted that the 'predictive power' of the λ_{σ}-index is slightly superior to that of the α_{σ}-index in terms of the results presented in this paper and those that have been studied by the authors to date. Further, the computational overheads required to compute the λ_{σ}-index are less and thus, for the development of applications on mobile trading devices, for example, the λ_σ-index may be better suited.

Acknowledgements

The author would like to acknowledge the support of the Science Foundation Ireland and the National Digital Research Centre of the Republic of Ireland.

References

[1] A. Einstein, *On the Motion of Small Particles Suspended in Liquids at Rest Required by the Molecular-Kinetic Theory of Heat*, Annalen der Physik, Vol. 17, 549-560, 1905.

[2] A. N. Kolmogorov, *On Analytic Methods in Probability Theory*, Selected Works of A. N. Kolmogorov, Volume II: Probability Theory and Mathematical Statistics (Ed. A. N. Shiryaev), Kluwer, Dordrecht, 61-108, 1992 (From the Original: Uber die Analytischen Methoden in der Wahrscheinlichkeitsrechnung, (1931). *Math. Ann.* 104: 415-458).

[3] W. Feller, *On Boundaries and Lateral Conditions for the Kolmogorov Differential Equations*, The Annals of Mathematics, Second Series, Vol. 65, No. 3, 527-570, 1957.

[4] F. W. Olver and L. C. Maximon, *Mittag-Leffler function*, Handbook of Mathematical Functions in Olver, (Eds. W. J. Frank et al.), NIST, Cambridge University Press, 2010.

[5] H. G. Schuster, *Deterministic Chaos: An Introduction*, VCH, Weinheim, 1988.

[6] V. I. Oseledec, *A Multiplicative Ergodic Theorem: Lyapunov Charcteristic Numbers for Dynamical Systems*, Trans. Mosc. Math. Soc., Vol. 19, 197-231, 1968.

[7] Y. G. Sinai, *Introduction to Ergodic Theory*, Princeton University Press, 1976.

[8] G. Boffetta, M. Cencini, M. Falcioni, and A. Vulpiani, *Predictability: A Way to Characterize Complexity*, 2001. `http://www.unifr.ch/econophysics/`.