CHAPTER 1
INTRODUCTION

“If farming were to be organised like the stock market, a farmer would sell his farm in the morning when it was raining, only to buy it back in the afternoon when the sun came out”

- John Maynard Keynes, economist

"Clearly the price considered most likely by the market is the true current price: if the market judged otherwise, it would quote not this price, but another price higher or lower."

- Louis Bachelier, mathematician

“When the weather changes, nobody believes the laws of physics have changed. Similarly, I don’t believe that when the stock market goes into terrible gyrations its rules have changed.”

- Benoit Mandelbrot, measure scientist

The present work is an empirical study of financial markets using non-traditional methods of time series analysis. Analysing a time series is a common area of study pervading multiple premises of science and has been addressed for a long time by statisticians. Understanding the dynamics of a financial series as a stock market index is still, however, a complex task having its specific requirements.

There are two main schools of time series analysis. The traditional one, which has a long pedigree in applied statistics, is prevalent among statisticians, social
scientists (especially econometricians) and engineers. The more recent school, developed essentially since the 1970s, comes out of physics and nonlinear dynamics. The first views time series as samples from a stochastic process, and applies a mixture of traditional statistical tools and assumptions (linear regression, the properties of Gaussian distributions) and the analysis of the Fourier spectrum. The second school views time series as distorted or noisy measurements of an underlying dynamical system, which it aims to reconstruct.

While the conventional methods have been very useful in modelling and analysing time series, these newer methods are capable of bringing out some other aspects specially related to the non-linear dynamics of the time series. There is a long history of established research work on time series analysis of financial markets using the traditional econometric methods. These works have already established all facets of the time series that can be captured by using these methods. On the other hand, very little work is found, especially in India, which uses the newer methods to capture some other aspects of the dynamics of the financial time series. Our effort is to bridge the gap in this direction. In the process we have used a mix of traditional and non-traditional methods to throw important lights on the validity of a financial model and have been able to compare its applicability vis-à-vis an alternative model.

We have put one of the most celebrated and epoch making models of finance under the scanner. The acceptability of a model is however, subject to empirical
validation. The importance of empirical validation in establishing a model cannot be overestimated. The present work is an endeavour in the same direction. Using the state of the art tools of empirical testing and data from the Indian Financial Market we have examined the validity of the same and compared the results with an alternative model. Since this study adopts a rather non-conventional approach, some discussion on the different models may be useful to justify our approach to investigate the financial time series patterns.

The day-to-day behaviour of the stock indices in a particular economy points to the basic stability of that particular market. There have been several attempts to have an approximate idea about the future behaviour of the market. One major attempt to model the stock price behaviour was by the French mathematician Louis Bachelier, who in his thesis on “Theory of Speculation” in 1901 (Bachelier 1901)\(^1\) presented a model that pre-empted the Brownian Motion. This was further modified by Samuelson to Geometric Brownian motion\(^2\) (Samuelson 1965)\(^3\). One of the corollaries of the same is that the stock price follows a log normal distribution. This hypothesis is one of the foundations of modern financial economic theories, one of the most famous of which is the Option Pricing model of (Black and Scholes 1973)\(^4\) and (R. Merton 1973)\(^5\). Our endeavour was to

---


\(^2\) For a detailed discussion on this please refer to chapter 3


simulate a Geometric Brownian Motion (GBM) based on data from stock indices of India and examine its validity with respect to actual price. To evaluate the efficacy of the model we decided to choose an alternative model and compare the relative performance of the two models, using specific tools. The choice of alternative model was critical. Our selection was based on one major consideration. Since the Black-Scholes-Merton option pricing model gives a closed end solution which makes it so easy to comprehend and use, the model alternative to the Geometric Brownian Motion model must result in an Option pricing mode which gives a closed end solution. This was done, keeping in view the utility of the model for the investor.

That the distributions of empirical returns of stock do not follow the lognormal distribution upon which many models of finance, especially the Option pricing models, are based, have been challenged by many empirical findings. While the derivations are of great importance and widely used, such theoretical option prices do not quite match the observed ones. In particular, the Black-Scholes model underestimates the prices of away-from-the-money options.

This means that the implied volatilities of options of various strike prices form a convex function, rather than the expected flat line. This is known as the “volatility smile”. Indeed, there have been several modifications to the standard models in an attempt to correct these discrepancies. One approach is to introduce a

---

stochastic model for the volatility of the stock price, as was done by Hull and White (1987), or via a Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model of volatility. A review and references like Bollerslev (1986), Engle and Merzrich (1996), Engle and Merzrich (1995), Nelson (1990) can be found in Hull and White (1987).

Another class of models includes a Poisson jump diffusion term (Merton 1976) which can describe extreme price movements. The Deterministic Volatility Function (DVF) approach by (Dupire 1994) as well as combinations of some of these different approaches like that of (Andersen 1996) have also been studied.

A quite different line of thought is offered in (Bouchaud J.-P. 1996, Bouchaud and Potters 2000), where it is argued that heavy non-Gaussian tails and finite hedging time make it necessary to go beyond the notion of risk-free option prices. They obtain non-unique prices, associated with a given level of risk. More

---

recently, other techniques along the lines of (Eberlein, Keller and K. Prause 1988) lead to option prices based on an underlying hyperbolic distribution.

To our knowledge, none of the above models result in manageable closed form solutions, which is a useful result of the Black and Scholes approach. However, Borland (2002) succeeds in obtaining closed form solutions for European options. Their approach is based on a new class of stochastic processes, which allow for statistical feedback as a model of the underlying stock returns. Their stochastic model derives from a class of processes (Borland 1998) which have been recently developed within the framework of statistical physics, namely within the very active field of Tsallis non-extensive thermo statistics (Tsallis 1988) (Curado and Tsallis 1991).

The next important decision was to select proper tools of empirical investigation which will clearly bring out the characteristics of the time series to be compared and evaluated.

There are two distinct yet broadly equivalent modes of traditional time-series analysis which may be pursued. On the one hand there are the time-domain

---

methods having their origin in the classical theory of correlation. Such methods deal predominantly with the auto-covariance functions and the cross-covariance functions of the series, and they lead inevitably towards the construction of structural or parametric models of the autoregressive moving-average type for single series and of the transfer-function type for two or more causally related series. Many of the methods which are used to estimate the parameters of these models can be viewed as sophisticated variants of the method of linear regression.

On the other hand are the frequency-domain methods of spectral analysis. These are based on an extension of the methods of Fourier analysis which originate in the idea that, over a finite interval, any analytic function can be approximated, to whatever degree of accuracy is desired, by taking a weighted sum of sine and cosine functions of harmonically increasing frequencies.

Over the last quarter of a century, apart from these two traditional schools, a new school has emerged due to influence of the paradigm of deterministic chaos. As a mathematical framework it shows rich and powerful structures. Most appealing part about deterministic chaos for the applied sciences researchers has been its ability to provide a remarkable explanation for irregular behaviour and anomalies in systems. The most direct link between the chaos theory and real world problems is the analysis of time series of real systems in terms of nonlinear dynamics. On one hand phenomenal progress in empirical techniques and data
analysis has made it possible to observe the most fundamental properties of nonlinear dynamical systems. On the other hand considerable progress have been made in exploiting ideas from chaos theory in cases where the system is not necessarily deterministic but the data analysis displays more features than can be captured by traditional methods. Problems of this kind are typical in, among many other systems, financial systems.

In this field, even simple models, whether microscopic or phenomenological, can lead to dynamics of high order of complexity. The pertinent question in the research of most financial studies is how best a model can capture the inherent dynamic properties of the time pattern of movements of financial data series. More often than not, one has to analyse the system from observations made on single time series, which is the case for most non-laboratory systems like financial time series in particular. The analytical framework of nonlinear dynamics provides new tools of qualitative and quantitative characterization of irregular (in comparison to absolutely ordered time series like a controlled electrical signal) time series. The scope of these methods range from extension of the spectral analysis like the Empirical Mode Decomposition to phase space reconstruction based tools like the Recurrence Plot and its quantification.

In the next few paragraphs we try to understand the utility and implication of these methodologies in the process explaining our reason for choice of particular methods. First we look at the idea of nonlinearity from the standard literature of
time series econometrics on one hand and nonlinear dynamics and chaos on the other. Next we summarize the nuances of some major tools of nonlinear time series analysis from statistical and nonlinear dynamics perspective. After a brief discussion on the relative advantages and disadvantages of the methods we explain our choice of methodology.

In the literature, there is no generally agreed definition for ‘non-linearity’. From the definition given by (De Grauwe 1993)\(^{20}\), a system \( X_t = h(\Omega_t, \alpha) \) is called a non-linear system if it is not possible to regenerate \( X_t \) by one linear model:

\[
X_t = \sum_{i=0}^{\infty} Y_i \varepsilon_{t-i} \quad \text{and} \quad \varepsilon \text{ is white noise and} \\
\sum_{i=0}^{\infty} Y_i \text{ is such that} \sum_{i=0}^{\infty} |Y_i| < \infty
\]

According to (De Grauwe 1993)\(^{21}\), the definition of non-linearity stems from the negation of linearity. This leaves a lot of other possibilities open for a so-called non-linear system. For example, (Hsieh 1989)\(^{22}\) divided the realm of non-linear dependencies into three categories. Additive non-linearity, also known as non-linear-in-mean, enters a process through its mean or expected value, so that each element in the sequence can be expressed as the sum of zero-mean random element and a non-linear function of past elements\(^1\). With multiplicative nonlinearity, or non-linear-in-variance, each element can be expressed as the


\(^{21}\) Ibid

product of a zero-mean random element and a non-linear function of past elements, so that the non-linearity affects the process through its variance. The final category is known as hybrid dependence, in which non-linearity enters through both the mean and the variance.

(Tsay 2005)23 has considered nonlinearity in the following way.

Let us consider a univariate time series \( x_t \), which, for simplicity, is observed at equally spaced time intervals. We denote the observations by \( \{x_t \mid t = 1 \ldots T\} \), where \( T \) is the sample size. A purely stochastic time series \( x_t \) is said to be linear if it can be written as

\[
x_t = \mu + \sum_{i=0}^{\infty} \psi_i a_{t-i}
\]

[1.1]

where \( \mu \) is a constant, \( \psi_i \) are real numbers with \( \psi_0 = 1 \), and \( \{a_t\} \) is a sequence of independent and identically distributed (IID) random variables with a well defined distribution function. We assume that the distribution of \( a_t \) is continuous and \( E(a_t) = 0 \). In many cases, we further assume that \( \text{Var}(a_t) = \sigma_a^2 \) or, even stronger, that \( a_t \) is Gaussian. If \( \sigma_a^2 \sum_{i=0}^{\infty} \psi_i^2 < \infty \), then \( x_t \) is weakly stationary (i.e., the first two moments of \( x_t \) are time-invariant). Any stochastic process that does not satisfy the condition of Eq. (1.1) is said to be nonlinear. The prior definition of nonlinearity is for purely stochastic time series. One may extend the definition by allowing the mean of \( x_t \) to be a linear function of some exogenous variables, including the time index and some periodic functions.

---

For a definition and understanding of nonlinear systems from the nonlinear
dynamics and chaos theory perspective, we will refer to (Hilborn 2004)\(^\text{24}\) as
under. A nonlinear system is a system whose time evolution equations are
nonlinear; that is, the dynamical variables describing the properties of the system
(for example, position, velocity, etc.) appear in the equation in a nonlinear form.
We can say that a system is linear if, and only if, the following condition holds:
Let us suppose that \(g(x,t)\) and \(h(x,t)\) are linearly independent solutions of the time
evolution equation for the system; then \(cg(x,t) + dh(x,t)\) is also a solution, where \(c\)
and \(d\) are any numbers. For example a system is described by the following
equation is linear

\[
\frac{d^2x}{dt^2} = kx
\]

[1.2]

But if the system is described by the following time evolution equation, it is
nonlinear,

\[
\frac{d^2x}{dt^2} = kx^2
\]

[1.3]

Looking at the understanding of nonlinearity from both the perspectives one can
discern the following. While the econometric approach is based on statistical
properties of a data set the dynamics approach is based on dynamic behaviour
of the time evolution. The latter study revolves around pattern recognition and
characterization of systems through dynamic properties.

\(^{24}\) Hilborn, Robert C. *Chaos and Nonlinear Dynamics*. Oxford University Press, 2004.:5
The approach is evident in the way time series analysis has evolved from the two schools. As brought out by Campbell et al.\textsuperscript{(2007)}\textsuperscript{25}, there is growing literature in econometrics on nonlinear financial time series. A vast majority of the same is devoted to building time series models that capture specific nonlinearities of the observed data set. A typical time-series model relates an underlying sequence of shocks $\varepsilon_t$. In linear time-series analysis the shocks are assumed to be uncorrelated but are not necessarily assumed to be IID. In nonlinear time series analysis the underlying shocks are typically assumed to be IID, but we seek a possibly nonlinear function relating the series $x_t$ to the history of the shocks. Most models used in practice fall into a class that can be written as

$$x_t = g(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots) + \varepsilon_{th} h(\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots),$$

where $g(.)$ is the conditional mean and $h(.)$ is the conditional variance. This equation leads to a natural division in the nonlinear time series literature between models of conditional mean and models of the conditional variance. Based on these lines various parametric as well as nonparametric models have been developed. There are a large number of nonlinear models even in the subset of parametric models. Priestley (1988), Teräsvirta, Tjøstheim, and Granger (1994)\textsuperscript{26}, Tong (1990)\textsuperscript{27}, Gao (2007)\textsuperscript{28}, and Fan and Yao(2003)\textsuperscript{29} provide excellent coverage of many of the most popular non-linear time-series models.


Among these are the Polynomial Models which represents the conditional mean function through its Taylor series expansion yielding a discrete-time Voltera (Voltera [1959]$^{30}$) series. This class of models can also be written in autoregressive as well as mixed Autoregressive/Moving Average representations, the nonlinear equivalent of ARMA models. Then there are the piecewise-Linear models which includes the first-order Threshold Autoregression(TAR) model. This model can be generalized to higher orders and multiple thresholds, as explained in Tong(1990)$^{31}$. This has led to some specialized models like Self-Extracting Threshold Autoregression(SETAR), Amplitude-Dependant Exponential Autoregression (EXPAR), and State-Dependant Models (SDM). A closely related family is the Markov-switching model of Hamilton (1989$^{32}$, 1990$^{33}$, 1993$^{34}$) and Sclove (1983a$^{35}$, 1983b$^{36}$). The key difference is that the changes in regime are determined not by the level of the process, but by an unobserved state variable which is typically modelled as a Markov chain. Then there are the models of changing volatility. To capture the serial correlation of volatility, Engle (1982)$^{37}$ proposed the class of Autoregressive Conditionally Heteroskedastic, or ARCH models. As a way to

---

$^{31}$ Tong, H., 1990, op cit
model persistent movements in volatility without estimating a very large number of coefficients in a high-order polynomial, Bollerslev (1986)\textsuperscript{38} suggested the Generalized Autoregressive Conditionally Heteroskedastic (GARCH) models. These models have been extended to multivariate forms like the VECH and BEKK models. There is a huge body of excellent literature on the entire GARCH family of models. New members are being added to this family to capture various aspects of nonstationary and nonlinear data and the family is ever expanding.

All these parametric models capture specific aspects of nonlinearity quite powerfully. But their main purpose is to test observed data against the estimated model. Since our purpose is to test the validity of a given model based on the Geometric Brownian Motion, these models discussed above, despite being very powerful will not serve our purpose. We propose to test the validity of the model by comparing the dynamic characteristics of the observed data and model generated series. Our interest lies in methods of characterizing observed data so that the same may be compared with the characteristics of the model generated data. We therefore shift our attention towards the methods of testing and characterising nonlinear data series.

A common approach to understanding characterizing nonlinearity in a time series is to establish test statistics for testing nonlinearity. There are a host of tests for linear models of time series. Out of these testing for autocorrelation and cross-correlation of data sets provide a clue towards degree of nonlinear association of

\textsuperscript{38} Bollerslev, T. 1986, \textit{op cit}
data sets. The autocorrelation function (ACF) is basically a statistical tool for finding repeating patterns, such as the presence of a periodic signal, which has been buried under noise, or identifying the missing fundamental frequency in a signal implied by its harmonic frequencies. It is used frequently in signal processing for analysing functions or series of values, such as time domain signals. In other words it is a statistical representation of the degree of similarity between a given time series and a lagged version of itself over successive time intervals. Since we are interested in nonlinear relationships in the data, a useful diagnostic tool is to examine the ACF of the squares or cubed values of a series (Enders [2004])\textsuperscript{39}. This has been illustrated by Granger and Teräsvirta (1993)\textsuperscript{40}. In signal processing, cross-correlation is a measure of similarity of two waveforms as a function of a time lag applied to one of them. This is also known as a sliding dot product or inner product. This gives a clue to the degree of nonlinear relationship between two data sets. Of the frequency domain methods the most common method of spectral analysis through Fast Fourier Transform can give simple clues towards degree of nonlinearity between two systems. The Power spectrum analysis is an effective tool to quantify chaos. Any data series, which is in chaotic state shows irregular behaviour and lots of peaks, arise corresponding to the spectrum analysis. Also a data set in either steady or periodic state will converge to its corresponding number of peaks in Power spectrum diagram.

There are some other tests for establishing nonlinearity in a data. The Regression Error Specification Test (RESET) posits the null hypothesis of linearity against a general alternative hypothesis of nonlinearity. For the standard discrete time models there is also the Lagrange Multiplier (LM) test. Specialized tests for TAR and related models have also been developed. Campbell et al. (2007)\textsuperscript{41} discusses the utility of higher moments of nonlinear models as a basis for a statistical test of nonlinearity Hsieh (1989)\textsuperscript{42} defines a scaled third moment and establishes a test statistic for IID acceptance/rejection. Grassberger & Procaccia(1983)\textsuperscript{43} introduced a dimension based on the behavior of the so called correlation sum or correlation integral. This approach requires large amount of data for which Brock, Dechert, and Scheinkman (1996)\textsuperscript{44} have developed an alternative approach that is better suited to the finite data set. They have proposed their BDS test statistic. However, the BDS (Brock, et al., 1996)\textsuperscript{45} test does not provide a direct test for non-linearity because the sampling distribution of the BDS test statistic is not known, either in finite samples or asymptotically, under the null hypothesis of non-linearity. The rejection of the null of independent and identical distribution (IID) in the BDS test can be due to non-white linear and non-white non-linear dependence in the data. Thus, the effects of linear serial dependencies have to be filtered out by fitting the best possible linear model before the BDS test can be applied to detect any non-linear departure from the

\textsuperscript{41} Campbell et al.(2007), \textit{op cit}
\textsuperscript{42} Hsieh, D.A. (1989) \textit{op cit}
\textsuperscript{45} Ibid
IID null. However, there is always the concern that the rejection of the null by the BDS test could be due to the possibility of imperfect pre-whitening. This concern is well directed since much of the Monte Carlo research that has been published on the BDS test (see example, Brock et al., 1991)\(^{46}\) has emphasized the pre-testing issue and the potential dependence of the properties of the test on the prior linear filter. Some of the test’s sensitivity to non-linearity could be a result of remaining linear dynamics in the data. Another popular non-linear test is the Hinich bispectrum test (Hinich, 1982)\(^{47}\), which involves estimating the bispectrum of the observed time series (see, for example, De Grauwe et al., 1993; Abhyankar et al., 1995\(^{48}\); Brooks, 1996\(^{49}\); Vilasuso and Cunningham, 1996\(^{50}\)). Unlike the BDS test, the Hinich bispectrum test provides a direct test for a non-linear generating mechanism, irrespective of any linear serial dependencies that might be present. Thus, pre-whitening is not necessary in using the Hinich approach. Even if pre-whitening is done anyway, the adequacy of the pre-whitening is irrelevant to the validity of the test. Ashley et al. (1986)\(^{51}\) presented an equivalence theorem to prove that the Hinich linearity test statistic is invariant to linear filtering of the data, even if the filter is estimated. Thus, the linearity test can be applied to the original returns series, or to the residuals of a linear model


with no loss of power. While the rejection of the null of linearity in the Hinich bispectrum test provides a strong support for the presence of non-linearity (Barnett et al., 1997)\textsuperscript{52}, the evidence gives researchers little clue as to what the appropriate functional form for the resultant non-linear model should be. This is basically due to its portmanteau or general nature, that is, it does not have a specific alternative hypothesis. Lukkonen et al. (1988)\textsuperscript{53} developed a linearity test, Lukkonen-Saikkonen-Teräsvirta (LST) test to test for the null of no non-linearity against the alternative of Smooth Transition Autoregressive (STAR) type non-linearity.

All these methods are based on IID acceptance /rejection. None of these methods bring out the dynamics of the time series that will enable us to compare between time series which are almost similar in nature (both may be rejected or accepted by the test statistic, thereby not helping in understanding the differences in time evolution of different time series all of which may be nonlinear in nature and thereby giving a test statistic value above the critical value.)

Keeping our purpose in mind, we had to use established methods which bring out the qualitative and quantitative aspects of the time series data that help us in direct comparison of both. Over the past few years due to rapid progress in computational ability, it has been possible to capture various aspects of a


nonlinear time series through powerful algorithms that render visualisation of the
dynamics of time evolution of signals.

Historically, Fourier spectral analysis has provided a general method for
examining the global energy-frequency distributions. As a result, the term
`spectrum' has become almost synonymous with the Fourier transform of the
data. Partially because of its prowess and partially because of its simplicity,
Fourier analysis has dominated the data analysis efforts since soon after its
introduction, and has been applied to all kinds of data. Although the Fourier
transform is valid under extremely general conditions, there are some crucial
restrictions of the Fourier spectral analysis: the system must be linear; and the
data must be strictly periodic or stationary; otherwise, the resulting spectrum will
make little physical sense. Several other methods like the spectrogram, wavelet
analysis, Wigner Ville distribution, Evolutionary spectrum, Empirical Orthogonal
Function expansion (EOF) etc. are designed to modify the global representation
of the Fourier analysis but none succeed to address the limitations of the Fourier
analysis fully.

Huang et al. (1998)\textsuperscript{54} summarize the necessary conditions for the basis to
represent a nonlinear and non-stationary time series: (a) complete; (b)
orthogonal; (c) local; and (d) adaptive.

Mode Decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis."
The first condition guarantees the degree of precision of the expansion; the second condition guarantees positivity of energy and avoids leakage. They are the standard requirements for all the linear expansion methods. For nonlinear expansions, the orthogonality condition needs to be modified. The details will be discussed later. But even these basic conditions are not satisfied by some of the above mentioned methods.

The additional conditions are particular to the nonlinear and non-stationary data. The requirement for locality is the most crucial for non-stationarity, for in such data there is no time scale; therefore, all events have to be identified by the time of their occurrences. Consequently, we require either the amplitude (or energy) and the frequency to be functions of time. The requirement for adaptivity is also crucial for both nonlinear and non-stationary data, for only by adapting to the local variations of the data can the decomposition fully account for the underlying dynamics of the processes and not just to fulfill the mathematical requirements for fitting the data. This is especially important for the nonlinear phenomena, for a manifestation of nonlinearity is the `harmonic distortion' in the Fourier analysis. The degree of distortion depends on the severity of nonlinearity; therefore, one cannot expect a predetermined basis to fit all the phenomena. An easy way to generate the necessary adaptive basis is to derive the basis from the data.

In the same paper, they introduce a general method which requires two steps in analysing the data as follows. The first step is to preprocess the data by the
empirical mode decomposition method, with which the data are decomposed into a number of intrinsic mode function components. Thus, we will expand the data in a basis derived from the data. The second step is to apply the Hilbert transform to the decomposed IMFs and construct the energy-frequency-time distribution, designated as the Hilbert spectrum, from which the time localities of events will be preserved. In other words, we need the instantaneous frequency and energy rather than the global frequency and energy defined by the Fourier spectral analysis. This method has been designated by the Patent Office as the Hilbert-Huang Transformation (HHT) Method\textsuperscript{55}, a name initiated by Professor Theodore Wu of Caltech as an alternative for Fast-Fourier Transform (FFT).

The Empirical Mode Decomposition results in the intrinsic mode functions (IMF). Through repeated Hilbert Transform of the IMFs one can get the distribution of instantaneous amplitudes and phases of the same. One can then perform a comparative quantitative analysis on different time series and can comment on their similarity/dissimilarity. The Hilbert Spectrum on the other hand, graphically illustrates the energy-frequency-time distribution where qualitatively one can directly compare time series.

An important concept in understanding dynamic systems has been the concept of phase space. A phase space, introduced by Willard Gibbs in 1901, is a space in which all possible states of a system are represented, with each possible state of the system corresponding to one unique point in the phase space. In a phase

\textsuperscript{55} We have discussed the method in detail in section 4
space, every degree of freedom or parameter of the system is represented as an axis of a multidimensional space. For every possible state of the system, or allowed combination of values of the system's parameters, a point is plotted in the multidimensional space. Often this succession of plotted points is analogous to the system's state evolving over time. In the end, the phase diagram represents all that the system can be, and its shape can easily elucidate qualities of the system that might not be obvious otherwise. A phase space may contain very many dimensions. Considering the time series to be representative of the dynamic economic system one can capture the dynamics of the time evolution of the system.

Eckmann et al. (1987)\textsuperscript{56} have introduced a tool which can visualize the recurrence of states $x_i$ in a phase space\textsuperscript{57}. Usually, a phase space does not have a dimension (two or three) which allows it to be pictured. Higher dimensional phase spaces can only be visualized by projection into the two or three dimensional sub-spaces. However, Eckmann’s tool enables us to investigate the m-dimensional phase space trajectory through a two-dimensional representation of its recurrences. Such recurrence of a state at time $i$ at a different time $j$ is marked within a two-dimensional squared matrix with ones and zeros dots (black and white dots in the plot), where both axes are time axes. This representation is called recurrence plot (RP). Through a careful observation one can easily detect the similarity/dissimilarity in dynamics of the time series under study.

\textsuperscript{57} For a detailed discussion please refer to Chapter 4
A quantification of recurrence plots was developed by Zbilut and Webber Jr. (1992) and extended with new measures of complexity by Marwan et al. (2002). This is known as Recurrence Quantification Analysis (RQA). The RQA measures have been quite useful for the analysis of a variety of data. Yet, in order to not only detect qualitative changes in a system’s dynamics but to be able to judge their significance or to compare two univariate time series, it is necessary to derive a quantitative judgment such as a confidence interval. However, Schinkel et al. (2009) have recently shown that for recurrence-based complexity measures those intervals can be estimated using a resampling paradigm. Thus a quantitative comparison of different time series is possible by this method.

We now explain the basis of selection of the tools for our research purpose. We simulate two distinct time series based on initial data from Indian stock market. One is based on the classical Gaussian model where stock price follows Geometric Brownian Motion. The other is based on the Non-Gaussian model based on Tsallis distribution as proposed by Borland (2002). Using proper tools we intend to examine the underlying dynamic characteristics of both the simulated time series and compare them with the characteristics of actual data. We find out which one better captures the non linear dynamics of the index.

---


returns. After having reviewed the different methods of analysing nonlinear time series, we arrived at the methods which will be useful for our purpose.

Empirical mode decomposition can decompose a time series into constituent functions and plot them in frequency domain. This tool thus enables us to directly compare two time series by comparing the similarity/dissimilarity of these constituent functions. Recurrence Plot, on the other hand is an effective tool for detecting regime changes and critical phases in a time series. It can also detect similarity/dissimilarity between two time series in terms of their nonlinear dynamics by comparing the characteristics of the two time series as they go through various phases.

Firstly, we perform the different nonlinearity tests on the three data sets to find out whether the simulated data sets are capturing the nonlinearity of the actual data set. We chose Cross-Correlation studies for giving initial clue towards degree of linear relationship between two time series. We then use FFT to draw the Power Spectrum Diagram and compare the time series based on the peaks and troughs of the spectra. Then we use the EMD analysis to compare the underlying periodicity/aperiodicity of the time series. Finally we do a recurrence Plot and Quantification Analysis to understand and compare the time series based on their underlying nonlinear dynamics.
Motivated by this, we simulate two distinct time series based on initial data from Indian stock market. One is based on the classical Gaussian model where stock price follows Geometric Brownian Motion. The other is based on the Non-Gaussian model based on Tsallis distribution as proposed by Borland (2002). Using techniques of Non-linear dynamics we examine the underlying dynamic characteristics of both the simulated time series and compare them with the characteristics of actual data. We find out whether as per claim the Tsallis distribution better captures the non linear dynamics of the index returns.

The rest of the work is presented in the following manner. Chapter 2 presents our survey of existing literature in the major research directions. In chapter 3 we have discussed the theoretical framework of the models under study viz., the Geometric Brownian Motion and the alternative model based on Tsallis distribution. In chapter 4 we have outlined our research methodology explaining in the process the theoretical premise of each of the methods used. In the next three chapters we present our empirical work based on data from stock market of India and abroad. In chapter 5 we have presented a study using Empirical Mode Description and established its utility in detecting and comparing dynamics of time series. In chapter 6 we have presented our work focusing on critical regimes of financial time series. We have used the Recurrence Plot analysis to a good effect to detect regime change in the financial time series. In chapter 7, we proceed with the analysis of the models under study using the methodology explained in chapter 4. Finally in chapter 8 we present our conclusions. In the

\[\text{Borland, Lisa (2002), op cit}\]
next few paragraphs we explain in brief the main features of some major chapters.

As stated above, in Chapter 4 we have explained in detail our methodology wherein we have elaborated on the theoretical and implementation aspects of all the tools used, viz., the nonlinearity tests (BDS, Hinich Bi-Spectrum, Keenan's test, White's and Terasvirta Neural Network Test) Cross correlation analysis, Power spectrum analysis, Empirical Mode Decomposition (EMD), Recurrence analysis. We have also explained our simulation techniques in this chapter.

In the next two chapters, in establishing the utility of the tools we have first used the same on various data sets from stock exchanges from India and abroad. These works clearly indicate the ability of the tools to throw light on the dynamics of a time signal.

In chapter 5 we discuss an empirical work where we have used the EMD technique to analyse two different financial time series, viz., the daily movement of NIFTY index value of National Stock Exchange, India, and that of Hong Kong AOI, Hong Kong Stock Exchange from July 1990 to January 2006. The returns of the two indices are shown to have strikingly similar probability distribution. The IMF phase and amplitude probability distribution of the two indices also reveal striking similarity. This indicates a remarkable similarity of trading behaviour in the two markets. Considering the geographical and political separation of the
two, this indeed is an important discovery. Our study also clearly demonstrates the efficacy of EMD as a dynamic data analytical tool.

In chapter 6, we have done an extensive study on extreme behavior of stock prices, i.e., bubbles and crashes, using Recurrence plot (RP) and Recurrence quantification analysis (RQA). RP and RQA are good at working with non-stationarity and noisy data, in detecting changes in data behavior, in particular in detecting breaks, like a phase transition and in informing about other dynamic properties of a time series. Endogenous Stock Market Crashes have been modelled as phase changes in recent times. Motivated by this, we have used RP and RQA techniques for detecting critical regimes preceding an endogenous crash seen as a phase transition and hence give an estimation of the initial bubble time. We have used a new method for computing RQA measures with confidence intervals. We have also used the techniques on a known exogenous crash to see if the RP reveals a different story or not. The analysis is made on Nifty, Hong Kong AOI and Dow Jones Industrial Average, taken over a time span of about three years for the endogenous crashes. Then the RPs of all time series have been observed, compared and discussed. Our studies reveal that, in the case of an endogenous crash, we have been able to identify the bubble, while in the case of exogenous crashes the plots do not show any such pattern, thus helping us in identifying such crashes. This study also clearly demonstrates the utility of recurrence analysis in bringing out the underlying nonlinear dynamics of a time series.
Having fully explored these two techniques from Nonlinear dynamics (EMD and recurrence analysis) with data from Indian as well as other markets, in chapter 7 we have proceeded with using the same methodology to empirically test the Geometric Brownian Motion for validity and also compare the same with the model based on Tsallis distribution. We first simulate distributions based on the alternative models (Borland and BSM) and performed the tests on both series and the actual data series. For these test we have specifically used data from Indian stock exchange, as our purpose was to test the model in Indian context. First we test for nonlinearity in all the three sets of data using Brock-Dechert-Scheinkman test (BDS). Additionally we also perform the Keenan’s test for nonlinearity. Another popular non-linear test is the Hinich Bi-spectrum test, which involves estimating the bispectrum of the observed time series. We also use this test to find out whether it detects nonlinearity in these time series. To reinforce our findings we also conduct the White’s neural Network tests on the same data set. Another linearity test for time series was introduced based on concepts from the theory of neural networks. Teräsvirta et al. developed its power fully. We use this Terasvirta Neural Network test as a final reinforcement of our findings. To analyse the dynamics of the time movement of the indices and the simulated data series we have used Empirical Mode Decomposition, Recurrence Plot analysis, Power Spectrum analyses, Delay based cross correlation function. Our findings enable us to comment on the efficacy of the Borland (2002) proposed model over the Black-Scholes model. The EMD analysis clearly shows that the
underlying periodicity of the return series based on the Borland model is much closer to the periodicity of actual return series of NIFTY than that of the simulation based on BSM model. The closer inspection of the phase and amplitude distribution of the 1st two IMFs also reveal the same picture. Our Recurrence Plot analysis showed that both the models capture the essential dynamics of the original data set. But a closer inspection of the Recurrence Quantification Analysis revealed a better picture. Our findings indicate that at micro epoch level the pricing model using Tsallis distribution captures the time series dynamics of the actual data set much better than the simulation using Geometric Brownian Motion assumption. The Power Spectrum Analysis shows the Borland model to be closer to actual data set in terms of distribution of peaks and troughs than the Geometric Brownian Motion Model. Finally the cross-correlation study gives us a clear verdict in favour of the Borland model as we find that for all values of the delay factor the correlation for Borland and NIFTY remains higher than that of GBM and NIFTY. This acts as an empirical validation of the underlying theoretical framework of the option-pricing model. The Borland model being a major closed end solution to Option Pricing can become a replacement of the BS model as a benchmark tool for the investors. The result of our findings establishes that the Borland model captures the non-linear dynamics of the time series representing the underlying better than the BS model.

We summarize our findings and explain our conclusions in the final chapter, i.e., chapter 8. Our results help us in establishing the Borland model as an alternative
to BSM model, which has far reaching implications for the derivatives market given the critical role of derivatives in global financial world and India being a key emerging market.

Our study thus focuses on techniques developed from physics and non-linear dynamics to compare and analyse two alternative models of stock price movements using data from Indian Stock market. In the process we have also used the techniques on two other markets viz., the US market and the Hong Kong market which further revealed interesting features of the dynamics of respective time series. We chose our methods mainly for two reasons. On one hand, the ability of these methods to pictorially capture the entire time evolution of the series, render these methods very useful as a comparative analysis tool. On the other hand, while a plethora of work exists which use the traditional methods, very few work has been done using these new methods. Our effort is thus to bridge this gap in existing work and bring out hitherto unexplored facets of the time series in the light of newer methods. We now proceed to present a survey of the existing literature on the models and the methods.