POLITECNICO DI TORINO

Master's Degree in Physics of Complex Systems

Master's Degree Thesis

Complex Systems Methods to describe Financial Time Series

Supervisors:

Prof. Luca Dall'Asta

Dr. Elio Stocchi

Candidate:

Emanuele Ricco

S261955

Summary

Financial markets are a perfect example of a complex system due to the behaviour of millions of investors who try to gain money each second, evolving as a chaotic environment very difficult to predict. In the last 20 years physicists and economists have tried to explain the price dynamics using tools from Statistical Mechanics, Theory of Turbulence and even Quantum Mechanics, combining concepts from financial world and theoretical physics. In the first part of the thesis we will expound these models both in a mathematical and in a historical point of view, retracing the development of Econophysics. In the second part of thesis we will compare these mathematical models with the aim of describing real historical data very different from each other; from financial indices to single stocks, from commodities to cryptocurrencies like Bitcoin, also analyzing the trend of Forex. We will also employ ideas from natural selection with Genetic Algorithms to evolve sets of parameters in order to better describe the real world data, solving optimization problems writing code on Python. For each time series we will also discuss the results in order to comprehend the accuracy of the algorithm, matching the outcomes of each time series.

Dedication

Dedico questa tesi a mio nonno Vincenzo, che veglia su di me durante ogni giorno della mia esistenza.

Acknowledgements

Ringrazio in primis i miei genitori Annapina e Ciro, mia sorella Anita, e le mie zie Marisa, Vanna e Carmen, per avermi sempre supportato durante il mio percorso di crescita personale ed universitario, senza i quali non avrei raggiunto questo traguardo.

Ringrazio il mio relatore Luca Dall'Asta, il quale mi ha accompagnato in questa tesi non convenzionale, offrendomi un importante supporto alla sua stesura.

Ringrazio il Dr. Elio Stocchi ed il Dr. Antonio Simeone, referenti di Euklid, che mi hanno dato la possibilità di trattare l'argomento filo conduttore del mio intero percorso universitario.

Ringrazio Niccolò, amico e collega prezioso, onnipresente durante la mia esperienza torinese.

Ringrazio infine me stesso per tutte le volte che sono caduto e mi sono rialzato, rimanendo sempre focalizzato sull'obiettivo finale senza mai perdere la rotta.

E il mio maestro mi insegnò com'è difficile trovare l'alba dentro l'imbrunire.

Contents

1 Introduction to Econophysics **6**
- 1.1 What is Econophysics? 6
- 1.2 GBM and Financial Markets 10
- 1.3 Quantum Mechanics and Financial Markets 14
- 1.4 Mathematical Model of the Thesis 22

2 Financial Time Series Results **27**
- 2.1 Introduction 27
- 2.2 NASDAQ 28
- 2.3 FTSE MIB 33
- 2.4 S&P500 36
- 2.5 EURO STOXX 50 40
- 2.6 DOW JONES 44
- 2.7 Gold 49
- 2.8 Petroleum 52
- 2.9 Amazon 56
- 2.10 Tesla 59
- 2.11 Forex EUR - USD 62
- 2.12 Bitcoin 65
- 2.13 Conclusion 70

3 Thesis codes **71**
- 3.1 Entropy 71
- 3.2 ECDF 71
- 3.3 Quantum Algorithm 72
  - 3.3.1 Quantum PDF 72
  - 3.3.2 Quantum Fitness of a vector: 72
  - 3.3.3 Quantum Fitness of a matrix: 72

3.3.4 Quantum Selection 72
3.3.5 Quantum Crossover 73
3.3.6 Quantum Genetic Algorithm 73
3.4 GBM Algorithm 75
3.4.1 GBM PDF 75
3.4.2 GBM Fitness of a vector: 75
3.4.3 GBM Fitness of a matrix: 75
3.4.4 GBM Selection 75
3.4.5 GBM Crossover 76
3.4.6 GBM Genetic Algorithm 76
3.5 PDF Plots 77

Chapter 1

Introduction to Econophysics

In this chapter we will discuss about the development and the theoretical concepts of the Econophysics, a new branch of the Theoretical Physics that was born in the early 2000's. We will also analyze in detail two of the most meaningful physical models to describe financial time series, Geometric Brownian Motion (GBM) and Quantum Harmonic Oscillator. We will explore in depth why these two methods are very common in the scientific community, explaining them both from a physical and financial point of view.

1.1 What is Econophysics?

Econophysics is a research field whose position is in the middle between many branches of Theoretical Physics, like Statistical Mechanics, Turbulence, Chaos Theory and Fractals, whose goal is to solve problems in economics. Financial markets, in fact, follow non linear dynamics enabling physicists to hypothesize their evolution like models already known, as the kinetic theory of gas or fluid dynamics. Econophysics was introduced for the first time by the American physicist Eugene Stanley, co-author of 'Introduction to Econophysics [24], giving rise to a new age of the finance description. This interdisciplinary field employs concepts from Probability Theory and Statistical Physics [25] to describe the quantitative properties of complex economical systems, made up of a huge amounts of humans who interact purchasing and selling financial products like stocks, commodities, ETF's and cryptocurrencies.

The need of new tools from physics that could overcome classical models from statistics and econometrics began in the middle 80's when the amount of financial data became very huge, and traditional methods were insufficient. The birth of Econophysics, instead, was not the first time in the history of science in which physics were involved in the economics or financial troubleshooting. Daniel Bernoulli published in 1738 'Specimen theoriae novae de mensura sortis', discussing about theories for the measurement of risk explaining the St. Petersburg Paradox as a basis for Game Theory concepts like risk adversion and utility function. Pierre - Simon Laplace, furthermore, pointed out the simplification of many random and unpredictable phenomena in his 'A Philosophical Essay on Probabilities' written in 1812. The first mathematician to apply physical concepts to financial markets was Luis Bachelier, who worked out on the Brownian motion modeling it as a stochastic process in 1900, 5 years before Albert Einstein. Bachelier anticipated by 70 years the famous publication of 'Black and Scholes' in 1973 regarding the pricing of derivatives in his economics PhD thesis. In 1969 Jan Tinbergen was the first physicist to win the Nobel Prize for Economics for having developed and applied dynamic models for the analysis of economic processes. Subsequently Benoit Mandelbrot, famous for his work on fractals, gave a very important contribution to the definition of the Modern Portfolio Theory.

The Polish mathematician analyzed the stock market variations, in particular the fluctuations in all the time scales of prices, researching the multiscaling laws that could describe the evolution of a rate. His analysis started from the leptokurtic distributions as a way to model return distributions, corroborated from the experimental results in this thesis, and also to the absence of short memory on returns that has also been confirmed by our work. In 1970 it was developed the 'Efficient Market Hypothesis' or EMH, according to which share prices reflect all information about the external market having an independence between prices of two subsequent days, confirming their evolution as a random walk. This theory assumes that when a new information comes into the market, with natural disaster like 2004 Tsunami in Asia or natural diseases like COVID-19 Pandemic, there is no way for an investor to gain more than a benchmark made up of randomly selected stocks [25]. EMH is the base of the technical analysis which is used by millions of traders each day to forecast stock prices. There have been created a lot of technical indicators that make traders understand which direction the

market is taking, without considering concepts from fundamental analysis or behavioural finance.
Fundamental analysis, in fact, exploits a more general approach to analyze the trend of a company [26], establishing its correct price both on the financial conditions of the company and on the geopolitical features of the market in which it is inserted. Behavioural finance, moreover, is a theory that deals with the relationship between cognitive psychology and the lack of rationality of economic agents, portraying why many investors don't follow the assumptions of game theory. Going back to the technical analysis, one of the indicators applied to financial time series in the thesis was the Hurst coefficient.

The starting point is always the EMH, so the assumption that in financial markets prices do not have trends, behaving casually in an independent way without presenting short or long term memory [27]. This thesis has been denied by Mandelbrot because in the real markets prices often follow trends, bullish or bearish, that depend on economic phases like inflation, recession and others, with non periodic cycles. For example the average italian annual inflation, or the general price increase, was larger than 3% in 2008, 2012 and maybe in 2022 according to [28], not following an aforethought scheme. We can therefore recognize these trends in detail analyzing historical time series of prices and returns, trying to detect long-term memory. These kind of testings are very important both for the portfolio optimization and for its prediction, being often used by algorithmic traders to device a financial strategy to invest.

Harold Edwin Hurst was a British hydrologist who tried to explain why on Nile river strong flood waves were followed by strong flood waves while light flood waves were followed by light flood waves in different time periods. So he tried to spot and describe these cyclical trends creating the Hurst index that would be used in the analysis of financial markets.
Given a historical time series $X_1, ..., X_n$ of length n, we have to:

- Calculate the mean of the data $m_n = \frac{1}{n}\sum_{i=1}^{n} X_i$
- Obtain the deviation of observation from the mean: $Y_t = X_t - m_n$ for $t = 1, 2, .., n$

- Compute the cumulative deviation for each observation: $Z_t = \sum_{i=1}^{t} Y_t$ for $t = 1, 2, .., n$
- Take the range of this cumulative: $R_n = max(Z_1, ..., Z_n) - min(Z_1, ..., Z_n)$
- Estimate the standard deviation of the data: $S_n = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2}$
- Standardize the range: $\frac{R_n}{S_n}$
- Fit the power law $\frac{R_n}{S_n} = CN^H$ with H the Hurst coefficient and $C \in R$

Benoit Mandelbrot in 1972 [29] demonstrated that the Hurst coefficient could assume values from 0 to 1 with the following meaning:

- $H < 0.5$: historical data are dependent and mean-reversing, so there is a negative correlation between subsequent events. A bearing phase will be followed by a bullish phase and viceversa, having a long-term correlation of the series.
- $H = 0.5$: historical data are not dependent and their behaviour is like a random walk. This is the only case in which EMH is valid.
- $H > 0.5$: historical data are dependent and trend - reinforcing, so there is a positive correlation between subsequent. The trend will be maintained and we expect that if it is positive, it will remain positive, if it is negative, it will remain negative.

As we will discuss later, non classical tools from Econophysics provide the methods to intercept interesting features of the historical series of returns like fat-tails and excess of kurtosis. This new area of research, moreover, is not free from criticism because it tries to make physicists and economists agree. The first ones, having a strong quantitative and scientific background, manage to detect some empirical laws from the financial markets describing them by the use of dynamical laws, even not closely linked to the world of finance. As we will show, in fact, there could be shown not only a very interesting connection between Quantum Mechanics and Finance but also with the study of earthquakes and with many different topics from Statistical Physics like the Master Equation and the Ising Model. Economists, instead, can depict many important variables that could guide periods of financial bubbles of crisis, and working with physicists creating models to better control them.

1.2 GBM and Financial Markets

We can observe a Brownian motion if there is a very small particle, with a diameter of the order of $10^{-6}\,m$ hanging on a gas or a liquid without the effect of the gravity force. A little grain of smoke suspended in the air or ink particles absorbed in a bottle of water have a continuous agitated and disordered motion. Each of these particles endures thousands of bumps with the neighbours, resulting in an almost null deviation of its position without a preferential direction of flow. If we want to model the Brownian motion of a set of flecks we need to introduce the Wiener Process, a Stochastic Gaussian process in continuous time which follow the subsequent properties copied from [31]:

'A standard one dimensional Wiener Process is a stochastic process W_t with $t \geq 0+$ having the following properties:

- $W_0 = 0$.
- With probability 1, the function $t \rightarrow W_t$ is continuous in t and a Markov chain.
- The process W_t has independent, stationary increments, so is a Levy Process.
- The increment $W_{t+s} - W_t$ has the $N(0,t)$ distribution.

We can comment each of the properties:

- The Wiener process is a Markov chain, so it depends on the past only through the immediately previous state, not on other past states or external information.
- The term independent increments means that $\forall s_1, t_1, ..., s_n, t_n : 0 \leq s_1 < t_1 \leq s_2 < t_2 \leq ... \leq s_n < t_n < \infty$ the increment random variables $W_{t_1} - W_{s_1}, W_{t_2} - W_{s_2}, ..., W_{t_n} - W_{s_n}$ are jointly independent.
- The term stationary increments means that for any $0 < s, t < \infty$ the distribution of the increment $W_{t+s} - W_s$ has the same distribution as $W_t - W_0 = W_t$.'

Next passages are copied from [32]:

The Brownian Motion has other fundamental properties:

- $E[W_t] = 0$ its expected value is null.
- $Var[W_t] = t$ its variance is equal to t.
- In an infinitesimal time δt we can say that $W_{t+\delta t} - W_t = dW_t = \epsilon_t \sqrt{\delta t}$ where $\epsilon_t \sim N(0, 1)$.
- The variable ϵ_t is uncorrelated because $E[\epsilon_t, \epsilon_s] = 0$ if $t \neq s$
- From the last two properties we can say that $E[dW_t] = 0$ and $Var[dW_t] = E[(dW_t)^2] = \delta t$

We can also generalize the Wiener process to more complex models, introducing the 'Brownian Motion with drift':

$$dX_t = \mu dt + \sigma dW_t \tag{1.1}$$

In which:

- μ is the drift parameter.
- σ is the variance parameter.
- dW_t is the increment of the Wiener process.
- X_t is the increment of the Brownian motion with drift.

It is also very important to itemize the most useful properties for the Brownian motion with drift over a finite time interval Δt :

- Brownian motion with drift is also a Gaussian process with following mean and variance.
- $E[\Delta X] = \mu \Delta t$
- $Var[\Delta X] = \sigma^2 \Delta t$

The Brownian motion with drift (or generalized Wiener process) is not suited to model price stocks both because it can generate negative values making it unfeasible and for its simplification of the phenomena. The mean and the variance, in fact, are constant both over time and the variation of the price, but we know that volatility varies each moment and is also one of the technical indicators used by traders, with CBOE Volatility Index. A way to model the continuous shift over time is a generalization of the Brownian motion with drift, also taken from [32]:

$$dX_t = a(x,t)dt + b(x,t)dW_t \tag{1.2}$$

In which:

- $a(x,t), b(x,t)$, drift and variance are known function of current state (price) and time.

The last one is an Ito Stochastic Differential Equation. As before, calculating mean and variance:

- $E[dX_t] = a(x,t)dt$ instantaneous drift rate
- $Var[dX_t] = b^2(x,t)dt$ instantaneous variance rate

The most interesting application of the equation (1.2) is the **Geometric Brownian Motion**, described by the following formula:

$$dX_t = \mu X_t dt + \sigma X_t dW_t \tag{1.3}$$

The terms have the following meaning in finance:

- dX_t : price variation over time
- μ, σ : constants
- dW_t : variation of a Wiener process over time

To push ahead in our discussion we have to do a step back to the description of the 3 Wiener process features: Markov chain, independent increments and stationary and Gaussian changes. The first two assumptions could be valid for the stock prices, because the Markov property is another way to explain one of the main hypothesis of the EMH about the overall information contained in the price, and it is also reasonable that price increments are independent each other. The Gaussian property is not allowable because the price of a stock could never be negative (except the Petroleum during Covid lockdown but we will discuss later this extreme case), so assume that price changes are log-normally distributed and natural logarithm of price follow the Ito process. We can therefore define a variable $F(t) = log(X_t)$ which is normally distributed, while X_t is log-normally distributed, rewriting the (1.3) in the following way:

$$dF_t = \mu X_t dt + \sigma X_t dW_t \tag{1.4}$$

with:

- $a(x,t) = \mu X_t$
- $b(x,t) = \sigma X_t$

Heralding now the Ito's lemma [32]:
Let's consider a Ito stochastic process X_t, like the one described in (1.2) and a function $S(x(t),t)$ which is 'at least twice differentiable in x and once in t', we can expand writing its total differential:

$$dS = \frac{\partial S}{\partial t}dt + \frac{\partial S}{\partial x}dx + \frac{1}{2}\frac{\partial^2 S}{\partial^2 x}(dx)^2 + ... \tag{1.5}$$

Substituting the results of equation (1.2) and neglecting terms of the order $0(dt)^2$, we obtain the result of Ito's lemma:

$$dS = \left[\frac{\partial S}{\partial t} + a(x,t)\frac{\partial S}{\partial x} + \frac{1}{2}b^2(x,t)\frac{\partial^2 S}{\partial^2 x}\right] dt + b(x,t)\frac{\partial S}{\partial x}dW_t \tag{1.6}$$

Now we will carry this physical concepts into the financial world.

Until now we have consider X_t as the historical series of prices, so $F(t)$ is the historical series of log-prices. We can apply the Ito Lemma to $F(t)$ noticing that at each discrete time $t = 1,..,T$, X and so on $F(X) = log(X)$ do not depend on time, so $\frac{\partial F}{\partial t} = 0$. Applying to (1.6) the terms coming from (1.4) we can obtain the differential of F. Knowing also that $\frac{\partial F}{\partial x} = \frac{1}{x}$ and that $\frac{\partial^2 F}{\partial^2 x} = -\frac{1}{x^2}$ the final result is:

$$dF = \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t \tag{1.7}$$

Equalizing the results coming from (1.7) and (1.4), calling τ as the holding period of the data we can then write that:

$$dF \sim N\left[\left(\mu - \frac{1}{2}\sigma^2\right)\tau, \sigma^2\tau\right] \tag{1.8}$$

Defining the continuously compounded returns:

$$R_t = \frac{252.5}{\tau} ln \left(\frac{x_{t+\tau}}{x_t} \right) \qquad [1] \tag{1.9}$$

with $\tau = 1, 5, 20$ which represents the holding period if data are Daily, Weekly or Monthly, we have demonstrated that the PDF of the continuously compounded returns is:

$$R_t \sim N \left[\left(\mu - \frac{\sigma^2}{2} \right) \tau, \sigma^2 \tau \right] \qquad [1] \tag{1.10}$$

We have then confirmed the hypothesis of the GBM model depicted in [1], which will be the basis for the estimation of μ and σ^2 of the model for all the historical series analyzed.

1.3 Quantum Mechanics and Financial Markets

There is a weird analogy between two topics seemingly far from each other, but with many concepts in common: Quantum Mechanics and Financial Markets.

In recent years, a specific branch of Econophysics has developed, namely quantum finance. The primary concept is the study of the collective actions of a group of particles at scales of very small lengths. Quantum Mechanics describes the behaviour of radiation, matter and their interactions at scales of the order of atomic and subatomic length scales with a probabilistic formalism.
As in classical physics, at macro-scale, we can be sure of the evolution of a system governed by deterministic laws, in the financial markets we can observe the price trend governed by stochastic processes, as explained in the previous chapter. If we wanted to study the behavior of the individual atoms of matter, however, we need particular formalism such as quantum mechanics. Individual atoms, in finance, are individual investors whose individual behavior is unknown, but we can only model the probabilities of buying or selling stocks or commodities. They are called 'Quantum Financial Particles', as described in [33] and their analysis permit both to device many statistical

models to describe the financial trends, like we are doing in this thesis, but also to forecast their future patterns.

One of the main idea of the Quantum Mechanics is the wave-particle duality, according to which matter and electromagnetic radiation have a dual nature, both corpuscolar and waves. This behaviour was the agreement between two features of light that were observed in a distincted way:

- **Corpuscolar:** the photoelectric effect is a physical phenomena which describes the emission of particles from a surface hit by EM radiation. This effect shows the quantum nature of the light, because the energy is distributed in discrete quanta, the photons. Each photon interacts individually with an electron, giving him his energy only if it is greater than the minimum threshold described by Planck's equation.

- **Waves:** the Young experiment in 1801 proved the wave nature of the light explaining the phenomena of diffraction and interference. These features demonstrated that light behaved like elastic electromagnetic radiation, and therefore like a wave.

The most interesting topic about this duality is that, depending on the ways in which we detect the light, sometimes it behaves like a particle and sometimes like a wave. Quantum mechanics, with its Schrödinger equation, manages to describe both of the two dynamics of light. In Financial Markets also, especially visualizing the trend of a stock price, there is a wave - particle duality because, as in the atomic case, it assumes one of the two shapes depending on the way we visualize it.
As we have seen before, traders use the technical analysis to study all the indicators created to predict the price movement. These indicators are various, here are some of the most important:

- **Moving averages:** mean of the last n values of a historical series of data, used to scan the trend of prices. They can be simple or weighted; in the previous case last prices have higher weight than the first ones.

- **Bollinger bands**: using the 20-days simple moving average, we double the standard deviation of the price creating an interval for each timeframe producing two lines, the upper and lower bounds. These bands are used as indicator of volatility, that increases when they are broader, and diminishes when they are more narrow.

- **RSI indicator:** Relative Strength index is a momentum indicator, so it measures the variations of the prices with respect to their effective levels, calculated as the closure price of fixed time intervals before, like five days. It is calculated as:

$$RSI = 100 * \left(\frac{U}{U + D}\right) \qquad [34] \qquad (1.11)$$

in which U is the average of the upward closing differences along the fixed time interval while D is the absolute value of the downward closing differences along the same fixed time interval. RSI is then a percentage which indicates whether a stock is overbought or oversold, giving an indication to the traders about the trends.

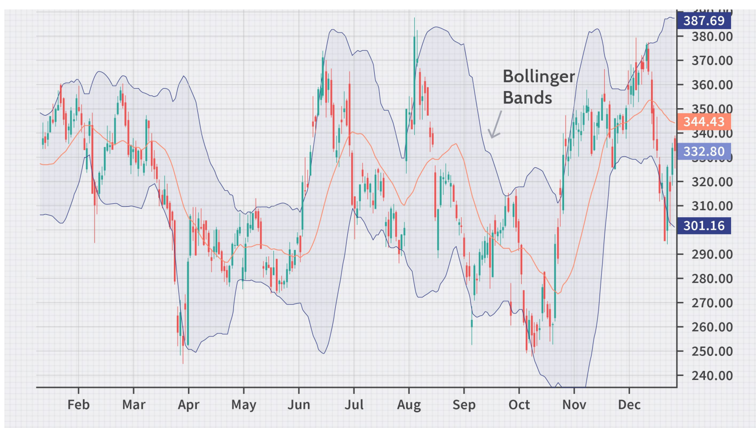

Figure 1.1: Example of Bollinger Bands taken from investopedia.com

Besides, when we observe a chart analysis graph like the one shown in Figure 1.1, studying the entire trend of the stock price and using trading tests to predict its pattern, we are dealing with the waves features of the price, considering its oscillatory motion. Using the technical analysis, furthermore, allows traders to set a price at which to enter the market. They can then decide based on technical analysis that if the price reaches a certain threshold they will buy or sell. In this case we are considering the corpuscular nature of the price, since all the information is contained in a single number that determines the behavior of the investors.

One of the most important blocks in Quantum Mechanics is the Heisenberg's Uncertainty Principle:

$$\Delta x \Delta p \geq \frac{\hbar}{2} \tag{1.12}$$

according to which the product between the uncertainty of two conjugate physical quantities like position and momentum has a lowering value, that derives from the commutation relation $[x, p] = i\hbar$. So we cannot measure both the position and the momentum of a particle in the same moment with an infinite precision. It looks like weird but this principle is also present in the observation of Financial Markets, as broadly discussed in [33] and in [35]. Having a look at the trend of a stock price using the corpuscolar property, in fact, we can imagine that the operator related to the position in Quantum Mechanics coincides with the price operator in Financial Markets because, as written in [35], 'the fluctuation of the stock price can be viewed as the motion of a particle in the space. Moreover, the energy of the stock, which represents the intensity of the price's movement, can be described by the Hamiltonian that simulates the fluctuation of the stock's price'. We can also make a parallel between the Dirac representation of a stationary wave function and the wave property of a stock:

$$|\psi\rangle = \sum_n c_n |\psi_n\rangle \tag{1.13}$$

ψ_n is the state of the stock before making the measure of trading it while $c_n = \langle\psi|\psi_n\rangle$ is its relative coefficient. Each state, as we will see, represents a different energy state of a stock which can be linked to its volatility which will increase with an upper value of n. Before making an operation of purchasing or selling a stock, then, each price will be a superposition of different volatility states with probabilities $|c_n|^2$.

Getting back to the uncertainty principle, as written in [35], there should be another operator $\hat{p}$ corresponding to the momentum. As guidance in quantum theory, the correspondence principle figures out that when the laws within the framework of the micro-world extend to macroscope, the results should be consistent with the outcomes of the classical laws. In the macro system, the momentum can be written as the mass times the first-order time derivative of the position in some special cases, so:

$$\hat{p} = m\frac{d}{dt}\hat{x} \qquad [35] \tag{1.14}$$

m is the 'stock mass' which is a constant of the motion and represent the capitalization of the stock, or the difficult of the price to change. When m increases, instead, so for very large companies like AMAZON with respect to lower companies like Ferrari, it is more difficult for the price to change. Moreover we can observe the uncertainty principle in finance because, as taken from [35], 'at a certain time someone knows nothing but the exact price of a stock. As a result, he certainly does not know the rate of price change at next time and the direction of the price's movement. In other words, the uncertainty of the trend seems to be infinite. However in the real stock market, we know more than the stock price itself at any time. We can always get the information about how many buyers and sellers there are near the current price. It is actually a distribution of the price within a certain range instead of an exact price due to the representation in equation (1.13). As a result, we can evaluate a standard deviation of the price. Thus the trend of the stock price may be partly known via the uncertainty principle. For example, a trader sees the number of buyers is far more than the number of sellers near the current price, he may predict that the price will rise at next time.'

According to the above, we need a quantum model to describe the trend of a financial stock, in particular regarding the distribution of continuously compounded returns, defined in (1.10), that could capture stock features like positive excess kurtosis and negative skewness, as explained in [36]. This quantum model is the one described in [1], with 3 great differences that will be expounded later.
As in the GBM model, we start from a Brownian motion with drift, described in (1.3). In the paper [1] the authors introduce the PDF $\rho(x,t)$, which has an evolution portrayed by the Fokker-Planck equation:

$$\frac{\partial}{\partial t}\rho(x,t) = \frac{\partial^2}{\partial x^2}[D(x,t)\rho(x,t)] + \frac{\partial}{\partial x}\left[\rho(x,t)\frac{\partial V(x,t)}{\partial(x)}\right] \qquad [1] \qquad (1.15)$$

in which $D(x,t) = \frac{\sigma^2}{2}$ [1] is a diffusion coefficient and $V(x,t)$ is the potential. If D is a constant and the potential does not depend on time, we can rewrite (1.15) in the following way:

$$\frac{\partial}{\partial t}\rho(x,t) = \left[\frac{\partial^2 V}{\partial x^2} + \frac{\partial V}{\partial x}\frac{\partial}{\partial x} + D\frac{\partial^2}{\partial x^2}\right]\rho(x,t) \equiv \hat{L}\rho(x,t) \qquad [1] \qquad (1.16)$$

The operator $\hat{L}$ is non - Hermitian due to the second term, which has the prime derivative. This non Hermitian property has non sense in our discussion treating observable quantities like prices and trends, so a possible solution is to transform it to a stock Hamiltonian, discussed in [1] and [35]:

$$i\hbar\frac{\partial\psi}{\partial\tau} = \hat{H}\psi(\hat{x},\tau) \qquad [35] \tag{1.17}$$

which evolution will be described by a Schrödinger equation shown in (1,17). As we can see, $\hat{H}$ will depend on τ because analyzing one historical series at a time, the model was not carried out on the temporal evolution of the stock prices but at various scales like Daily, Weekly and Monthly.

$$\hat{H} = \hat{H}(\hat{x},\hat{p},\tau) \qquad [35] \tag{1.18}$$

Hamiltonian is a function of the price, of the momentum and of the holding period. The most difficult challenge is to create an Hamiltonian because of all the external information contained like the psychology of the traders, the geopolitical information that could affect the price, liberal or statist policies which could determine a free or controlled price trend and many others. Writing as usual the Hamiltonian operator as the sum of the kinetic and the potential part:

$$\hat{H} = \frac{-\hbar^2}{2m}\bigtriangledown^2 + U(\overrightarrow{\mathrm{r}}) \qquad [1] \tag{1.19}$$

The vector $\overrightarrow{\mathrm{r}}$ has n components, each of which corresponds to a different stock. From now on we will consider a 1D Hamiltonian, because in our model we have analyzed one stock at a time.

$$U(x) = U(0) + \frac{1}{2}kx^2 \qquad [1] \tag{1.20}$$

In the paper [1] the stock Hamiltonian obtained is an harmonic oscillator with equation (1.20) describing its potential. The mathematical steps have not been entered in the thesis, but it is very important to clarify the reason for the harmonic oscillator potential. It has been chosen this quantum model to describe the **mean reversion** of a stock price at very long times, or the price tendency to always return to an equilibrium value, neglecting terms of high order then considering small values of oscillations, with:

$$k = d^2U(x)/dx^2|_0 \tag{1.21}$$

the elastic constant of the oscillator which is a firm constant describing the speed of the mean reversion, depending also on the external condition.

$$F(x) = -\frac{dU(x)}{dx} = -kx \tag{1.22}$$

F is the Hooke force that recalls the price towards the average.

$$\omega = \sqrt{\frac{k}{m}} \tag{1.23}$$

The angular frequency ω, explains the price fluctuations around an equilibrium point. We therefore expect lower fluctuations in higher time-frame returns such as Monthly data, and greater fluctuations in Daily data. The angular frequency of oscillation in this model is therefore expected to increase from Daily to Monthly data.

It is very important to recall the solutions of the n^{th} eigenfunctions of the stationary harmonic oscillator:

$$\phi_n(x) = \frac{1}{\sqrt{2^n n!}} \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{4}} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) exp\left(-\frac{m\omega}{2\hbar}x^2\right) \qquad [1] \tag{1.24}$$

As explained before, each one of these N states represent the volatility of a time series, whose energy describes the excitation of the state, with H_n the n^{th} Hermite polynomials. Mentioning also the eigenenergies of a quantum harmonic oscillator,

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega \qquad [1] \tag{1.25}$$

A time series described by an upper value of n will be more energetic and then more volatile then another series with a lower value of energy.
As described in [1], we can write down the final solution of the Fokker - Planck equation:

$$\rho(x,\tau) = \sum_{n=0}^{\infty} \frac{A_n}{\sqrt{2^n n!}} \left(\frac{m\omega}{\pi\hbar}\right)^{\frac{1}{2}} exp(-E_n\tau) H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) exp\left(-\frac{m\omega}{2\hbar}x^2\right) \qquad [1] \tag{1.26}$$

Which can be rewritten in the following way, as a linear combination of harmonic oscillators:

$$\rho(x,\tau) = \sum_{n=0}^{N} C_n(\tau)\ \rho_n(x) \qquad [1] \tag{1.27}$$

In which:

- $\rho(x, \tau)$: theoretical distribution probability, output of the model to obtain
- N: number of harmonic oscillators of the model. We have considered $N = 5$ in order not to increase the complexity of the model too much.
- $C_n(\tau)$: coefficients that depend on both the n^{th} oscillator and the parameter τ, to be obtained through the optimization algorithm.
- $\rho_n(x) = H_n\left(\frac{m\omega}{\hbar}\right) exp(\frac{m\omega}{\hbar}x^2)$

The estimated coefficients $C_n(\tau)$ are very important because we can obtain the probability of each singular state in the following way:

$$P_n = \frac{|C_n|^2}{\sum_{k=0}^{N} |C_k|^2} \qquad [1] \tag{1.28}$$

We expect that the probability of the $n = 0$ state, the Gaussian one with the least volatility, will be the more probable with a probability of more than 90%.

In [1] the coefficients $C_n(\tau)$ are obtained by minimizing the Cramer Von Mises Test:

$$T(\theta) = \frac{1}{12M} + \sum_{j=1}^{M} \left[F(r_j, \theta) - \frac{j - \frac{1}{2}}{M}\right]^2 \tag{1.29}$$

In which:

- $r_j = R_j - R_{av}$ with R_j the j^{th} compounded sorted return and R_{av} their mean
- M is the number of returns
- $F(r_j, \theta)$ is the cumulative distribution function of the returns

The optimization algorithm used in the paper [1] is the Newton - Raphson one, applying the GBM and Quantum models only to the FTSE ALL Share Index.

1.4 Mathematical Model of the Thesis

The differences in this thesis with respect to [1] are the following two:

- **Optimization function:**
  Instead of the Cramer-Von Mises Test, in this thesis the N parameters $C_n(\tau)$ and the parameter $\frac{m\omega}{\hbar}$ for the Quantum Model and the 2 parameters μ, σ^2 for the GBM model are obtained maximizing the posterior distribution:

  $$p(\theta|x) \sim p(x|\theta) \tag{1.30}$$

  We have then considered each of the $N+1$ parameters of the Quantum model and each of the 2 parameters of the GBM model extracted from a different continuous uniform distribution, whose choice of extreme values will be explained later. So the goal has been to maximize the Maximum Likelihood, that is analogous to minimize the Kullback-Leibler divergence between the empirical distribution function and the theoretical one, as demonstrated in [7]:

  $$KL(\tilde{p} \mid\mid q(.|\theta)) = -\frac{1}{M}\sum_{\mu=1}^{M} log(x^{\mu}|\theta) - S(\tilde{p}) \tag{1.31}$$

  In the last equation:

  $$\tilde{p}(x) = \frac{1}{M}\sum_{\mu=1}^{M} \delta(x, x^{\mu}) \tag{1.32}$$

  is the ECDF (empirical cumulative distribution function) of the continuously compounded returns. The last 2 equations are drawn by [7]

  For the entropy S of a historical time series we have considered the Approximate Entropy introduced in [8] as a measure of complexity of the data that analyzes the repetition of patterns into the time series. If there are repeated patterns in the series and therefore it will be forecastable, the entropy will be lower, if instead the series is random with noise and difficult to predict, the entropy will increase. Describing the formulas written in [8] at page 110:

ApEn is a function that depends on 3 variables:

- U: historical series considered
- m: length of the sequences compared
- r: filter of the data sequences

Having the U series made up of N data points this is a step by step explanation of the formula:

1. Create $N - m + 1$ vectors u of length m, each of them is $x[i] = [u[i]...u[i + m - 1]]$ with $i = 1...N + m - 1$
2. Define $d(x_i, x_j)$ with $i, j = 1...N+m-1$ as the maximum difference in absolute value between their scalar components
3. For each $i = 1...N + m - 1$ define $C_i^m(r)$ as the (number of $j \leq N + m - 1$ that $d(x_i, x_j) \leq r)/(N - m + 1)$
4. Define $\Phi^m(r) = (N - m + 1)^{-1} \sum_{i=1}^{N-m+1} ln\ C_i^m(r)$ in which ln is the natural logarithm
5. Finally $ApEn(U, m, r) = \lim_{N\to\infty}[\Phi^m(r) - \Phi^{m+1}(r)]$

- **Genetic Algorithms**
  Instead of the Newton - Raphson test, the $N + 1$ parameters for Quantum and the 2 parameters for GBM were estimated using Genetic Algorithms.
  The GA steps have been the following:

  - **Initializing population:** [9].

    For both the GBM and the Quantum model the initial population was made up of 1000 chromosomes or hyperparameters. GBM chromosomes were 2 in length, estimating μ, σ^2. Quantum chromosomes were 6 in length, estimating $C_0, C_1, C_2, C_3, C_4, \frac{m\omega}{\hbar}$
    The parameters are greatly influenced by the initial distribution chosen. In particular, the following values were used, assuming a uniform priority for each single parameter:

    * $\mu \sim U(0.5, 1.5)$
    * $\sigma^2 \sim U(3, 4)$

* $C_0 \sim U(-0.1, 0.1)$
* $C_1 \sim U(-0.1, 0.1)$
* $C_2 \sim U(-0.1, 0.1)$
* $C_3 \sim U(-0.1, 0.1)$
* $C_4 \sim U(-0.1, 0.1)$
* $\frac{m\omega}{\hbar} \sim U(0, 0.3)$

These initial values, kept constant also for the following time series, have been chosen heuristically to minimize the possibility of having not physical probability distributions with negative values. It was in fact observed at the beginning of the development of the code that, following the indications of [1], the order of magnitude of the coefficients C_1, C_2, C_3, C_4 of the Quantum algorithm was about 10^{-3} times less than C_0. Initially, therefore, very small values were chosen as extreme values of the uniform distribution of these parameters, $[-0.005, 0.005]$, with a smaller number of generations, 15, a smaller size of initial population, 100 instead of 1000, and a smaller value of hyperparameters selected at each cycle, 50 instead of 100. This resulted in non-physical PDFs and, for the code described in the pseudocode, in a choice of parameters having a fitness equal to 9999999, reducing the decrease in fitness itself and the robustness of the code. The parameters of the NASDAQ COMPOSITE, first data analyzed, were therefore modified first and then these initial conditions were maintained for all subsequent time series.

- **Fitness Evaluation:**

  The Kullback-Leibler equation (2.7) has been the Fitness Function for each hyperparameter, both for the GBM and for the Quantum Model.

- **Selection:**

  The goal is to minimize the Fitness Function filtering the hyperparameters not performing well, classified in two main classes:

  * High Fitness value.
  * Non - Physical PDF.
    In the previous case we have decided to assign at each set of

hyperparameters the default value of 9999999 so as not to be selected by the algorithm.

The size of the initial population was 1000 both for the Quantum Model and for the GBM one. At each generation the best 500 set of coefficients have been selected passing to next step.

- **Crossover:**

  The goal is to construct new chromosomes taking good features from the past generation. The crossover between GBM and Quantum was done in two different ways in the following way:

  * **Quantum Crossover:**
    At each one of the chromosomes it is associated a different integer random number r from 1 to N, different at each generation, to create a crossover between two subsequent hyperparameters. The next chromosome will be obtained by the coefficients from 1 to r in the same row, and by the coefficients from r to N in the next row. The $\frac{m\omega}{\hbar}$, instead, which cannot negative, are scaled from one row to the next row to preserve their physical meanings different to the other coefficients.
  * **GBM Crossover:**
    Two different integer random numbers are generated, one for μ and the other one for σ^2, different at each generation. The two coefficients are then scaled in a various way randomizing the association between the mean and the variance of the distribution.

- **Mutation:**

  At each generation two matrices are created of the same shape of the coefficient matrices, one for GBM and one for Quantum. The element of these mutation matrices, which change at each generation, are random numbers close to 1 in order to modify a little bit the coefficients mutating them.

- **New generation:**

  At each new generation, which is composed by 1000 hyperparameters, 500 will be the selected ones also called the parents, the other 500 will be the mutated ones. With this choice the worst result in term of fitness will be almost the one of the parents, so that the fitness will decrease at each generation.

– **Final choice:**

  At the end of the generations, different for each model and time series, we choose between the hyperparameters with the lowest fitness values the one that fits better the histogram of the compounded returns, creating the plot on Matplotlib that will be shown in the next chapter.

In this thesis we have analyzed the behaviour of a variety of different time series. The GBM and the Quantum algorithm have been applied to financial indices like NASDAQ COMPOSITE, FTSEMIB, S&P500, EURO STOXXX 50 and DOW JONES, to commodities like Gold and Petroleum, to single stocks like Tesla and Amazon, to the forex change between EUR-USD and to cryptocurrencies like Bitcoin. In the next chapter we will show and discuss the results for each price series.

Chapter 2

Financial Time Series Results

2.1 Introduction

In this chapter of the thesis all the financial time series analyzed will be studied in detail. The results of the inference testing will be shown for each historical series for both models, GBM and Quantum, with the explanation of each stock, commodity, index or crypto displayed. All the historical data have been downloaded on Jupyter Notebook from Yahoo Finance and the PDF of their compounded returns are represented by 'density histograms', so the total area of all the bars add to 1. These histograms have been built by the Freedman - Diaconis rule. The goal is to choose the optimal bin of an histogram in order to let it be proportional to the variance of the distribution, and diminishing by the sample size N.

$$\Delta = \frac{2IQR(x)}{\sqrt[3]{N}} \qquad [10] \tag{2.1}$$

In the previous formula:

- IQR(x) is the interquartile range of the data, so the difference between the 75^{th} and the 25^{th} percentiles.
- N is the dimension of the sample.

The data analysis period is variable and will be made explicit in each historical series.

2.2 NASDAQ

NASDAQ, which stands for 'National Association of Securities Dealers Automated Quotation', was the first stock exchange exclusively electronic and was born on February 8th 1971. In 2021, over 3000 different companies are listed on NASDAQ many of which belong to the technology sector. NASDAQ COMPOSITE is in fact the main reference for all securities in this trade area ranking second globally, behind the NYSE, in Daily trading volume. The capitalization is in fact about 9700 billion dollars, according to the information reported in [3].

NASDAQ COMPOSITE is a capitalization-weighted index; its components are stocks whose weight is determined by the market value of its company. The weight of each stock is not a constant and it can be shifted based on the number of the shares exchanged. The total price of the index is then a weighted sum between all the prices of the stocks.

Currently the top 10 constituents of the index are [4]:

- Apple
- Amazon
- Microsoft
- Tesla
- Alphabet
- Paypal
- Intel
- Facebook
- NVIDIA
- Comcast

In this thesis it was considered the behaviour of NASDAQ COMPOSITE from 21st September 1996 to 1st September 2021.

Figure 2.1: NASDAQ COMPOSITE behaviour from 1996 to 2021.

From a preliminary analysis of the compounded returns of the data, the following table was obtained:

	τ	N° of observations	Mean	Std	Skewness	Excess kurtosis	Entropy	Hurst
0	1	6278	0.101619	3.98087	-0.185306	6.270589	0.811658	0.249004
1	5	1302	0.097997	1.643938	-0.961178	7.608045	0.208412	0.312633
2	20	298	0.106978	0.835199	-0.735656	1.934904	0.027778	0.255679

Figure 2.2: Features of NASDAQ COMPOSITE compounded returns.

$\tau = 1$ stands for Daily data, $\tau = 5$ stands for Weekly data and $\tau = 20$ stands for Monthly data. τ represents the interval of consecutive days in which each time series is considered. In fact, on Saturday and Sunday the stock exchanges are closed and therefore a week lasts 5 days and a month lasts 20 days, considering it consists of 4 weeks. This is in fact the reason why the months on the stock exchange are not the same as in the calendar.

As we can observe from NASDAQ table the PDF of compounded returns remains leptokurtic but decreasing its kurtosis despite increasing τ, so the holding period. This behaviour, that as we shall see is not true for all the series analyzed, it indicates that more extreme events will have a greater likelihood in shorter time intervals. The mean remains almost constant between the holding periods, the standard deviations diminishes from Daily to Weekly but increasing itself during the Monthly. This parameter, which is one of the main constants for volatility, indicates how much the compounded returns are dispersed from the mean and could predict the tendency of the markets. The skewness is a little bit negative and increases in absolute value during the holding period, so the compounded returns tend to be positive, especially in the Monthly data. This means that the NASDAQ is bullish, in fact its last price is 12.51 times the first one.

The volatility is one of the main components of the strategy of a trader because, depending on the willingness to take risk, he might decide if opening the position and going long expecting a bull market or rest. The index of NASDAQ Composite is then a benchmark, in fact all investors analyze its data as an index to understand if their strategies are profitable or not. When a trader wants to invest in a stock like Apple, expecting a strong rebound in its returns, he will compare the average volatility of his stock, normally higher, with the NASDAQ, thus deciding the strategy.

The price fluctuations will then be broader for a Daily trader with respect to a Monthly trader, increasing the risk and decreasing the predictability of the series as we can see from the entropy, which is in Daily data 40 times larger than the Monthly one. As explained before, analyzing the fractality of the historical data, instead, we can observe that the 3 series are antipersistent, having an Hurst coefficient lower than 0.5. The data are so exhibiting a strong long - term memory and all the features explained before could maintain themselves in the future, in fact Hurst coefficient is one of the main parameters used to create a forecasting strategy for stock and index traders.

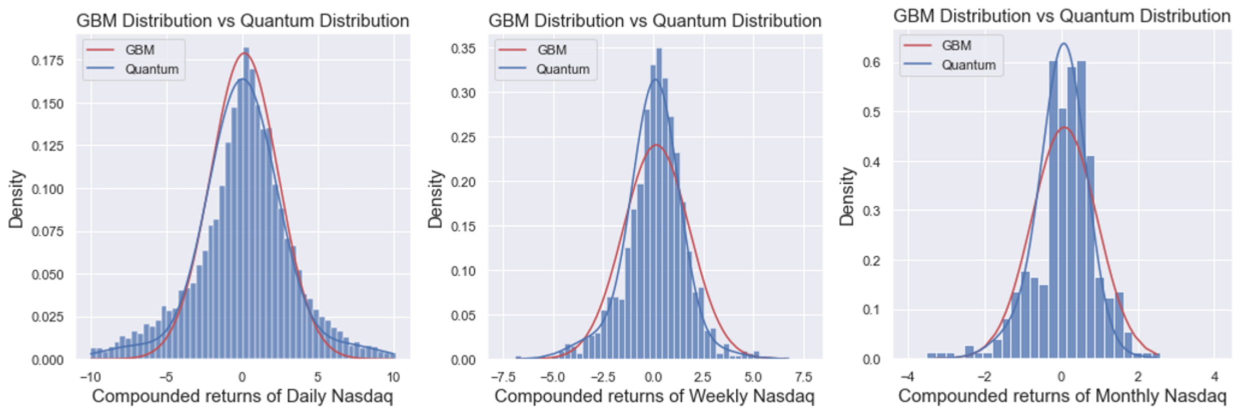

Figure 2.3: GBM vs Quantum NASDAQ COMPOSITE.

As we can see in each of the three cases, the Quantum Model has a good fit of the histogram, expecially in the Weekly data because it manages to recover the kurtosis and the tails of the distribution. The GBM Model overcomes the Quantum just in the Daily data fitting the extreme peak of the PDF. We can therefore combine the information of the Monthly data with that concerning the entropy and the Hurst coefficient by predicting that the upward trend of the NASDAQ will continue in the future.

Running the Quantum Genetic Algorithm for **75 generations** and the GBM Genetic Algorithm for **40 generations**, we have obtained the following fitness results:

Model Fitness	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM Fitness	2.80	1.92	1.24
Quantum Fitness	2.72	1.86	1.20

Figure 2.4: Fitness NASDAQ COMPOSITE.

Quantum Fitness is always less than GBM Fitness, so the Quantum Algorithm manages to better describe the empirical distribution. We can also observe that the fitness decreases by increasing τ, it can therefore be observed that the efficiency of the genetic description algorithm increases for longer time intervals.

Models	Parameters	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM	μ	1.27	0.98	0.51
	σ^2	2.23	1.65	0.85
Quantum	C_0	6.8363655e-01	1.2963708e-01	5.6941934e-01
	C_1	1.1152895e-02	6.8135983e-07	-3.212482e-06
	C_2	4.2946522e-05	-1.262956e-06	-3.445024e-03
	C_3	2.4101291e-05	-2.102043e-03	-1.128959e-02
	C_4	1.3493506e-02	1.8642911e-03	4.2375116e-03
	$m\omega$	5.7075986e-36	2.361275e-35	1.0917404e-34

Figure 2.5: Estimated Parameters NASDAQ COMPOSITE.

As regarding GBM, both the mean and the variance decrease over the holding period, so the parameters will confirm the drop of the kurtosis, resulting in a more flattened distribution. For the Quantum parameters, instead, we can observe that the product between mass and frequency, which must be multiplied to $\hbar$ with respect to the output of the code, increases for the holding period. As explained in the previous chapter the mass represents the firm capitalization or the total market value of its stocks. Being a peculiar characteristic of an index, it remains constant for the holding period, and increases because it is multiplied by ω.

	P_0	P_1	P_2	P_3	P_4
$\tau = 1$	9.99344e-01	2.65974e-04	3.94385e-09	1.24206e-09	3.89326e-04
$\tau = 5$	9.99530e-01	2.76115e-11	9.48668e-11	2.62797e-04	2.06711e-04
$\tau = 20$	9.99515e-01	3.18131e-11	3.65855e-05	3.92899e-04	5.53537e-05

Figure 2.6: Probabilities NASDAQ COMPOSITE.

The probability of being in $n = 0$, so the state less volatility, is the greatest for each holding period as explained in [2]. It can also be observed that the probability of the first state is almost constant during the holding period, so the NASDAQ COMPOSITE tends to remain to the same state of volatility. It is interesting to observe the behaviour of the probability of the higher volatility states, that will be compared to the stocks of Amazon and Apple.

As regarding the odd states, the one with $n = 1$ diminishes in probability instead of $n = 3$ that grows. For even states, the one with $n = 2$ increases in the Weekly data and decreases in the Monthly ones, while the one with $n = 4$ tends to decrease. This information therefore helps to explain the behavior of skewness, maximum in the Weekly for the combination of a decrease in P_2 and an increase in P_3.

2.3 FTSE MIB

FTSE MIB, which stands for 'Financial Times Stock Exchange Milano Indice di Borsa' is the index representing the Italian Stock Exchange trend. In particular, like NASDAQ, FTSE MIB is made up of the stocks of the 40 italian largest companies, even if their registered office is in a foreign country, with the greatest market capitalization. Some of the most important companies represented are:

- Campari
- Poste Italiane
- Exor
- Eni
- Ferrari

In this thesis it was considered the behaviour of FTSE MIB from 21st September 1996 to 1st September 2021.

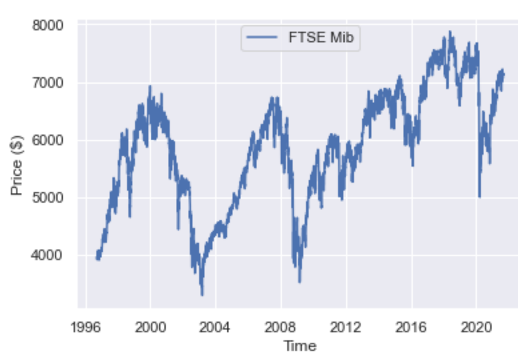

Figure 2.7: FTSE MIB behaviour from 1996 to 2021.

From a preliminary analysis of the compounded returns of the data, the following table was obtained:

	τ	N° of observations	Mean	Std	Skewness	Excess kurtosis	Entropy	Hurst
0	1	6286	0.023522	2.986324	-0.304879	7.149416	0.609121	0.256658
1	5	1302	0.022713	1.239301	-1.137494	11.03137	0.101046	0.323479
2	20	298	0.024649	0.512855	-0.71451	1.111746	7e-05	0.217938

Figure 2.8: Features of FTSE MIB compounded returns.

From FTSE MIB table we can observe that the mean remains almost the same during the holding periods while the standard deviation diminishes, restoring the centrality of the distribution around the mean. The skewness remains negative, so the distribution has a longer left tail, with a peak of its value at the Weekly data in which the kurtosis is also very high. We can also point out that the 3 different distributions are leptokurtic with a very fat-tailed distribution for the Weekly one. The entropy decreases with the holding period, instead, being 10^{-6} times the Daily one for the Monthly data indicating a time series much more predictable. The Hurst coefficient, instead, remains under the value of 0.5 for each of the 3 holding periods pointing out a long switching between very different values in size, so the autocorrelation is low and the fractal dimension is high.

The FTSE MIB is like the NASDAQ for Italy, being a benchmark for the italian traders who want to overcome the market returns. It is also considered as an index for foreign investors regarding the Italian economic situation. Looking at the figure (4.7) we can observe a very steep price increase, precisely of the 32%, from the begin of 2020.

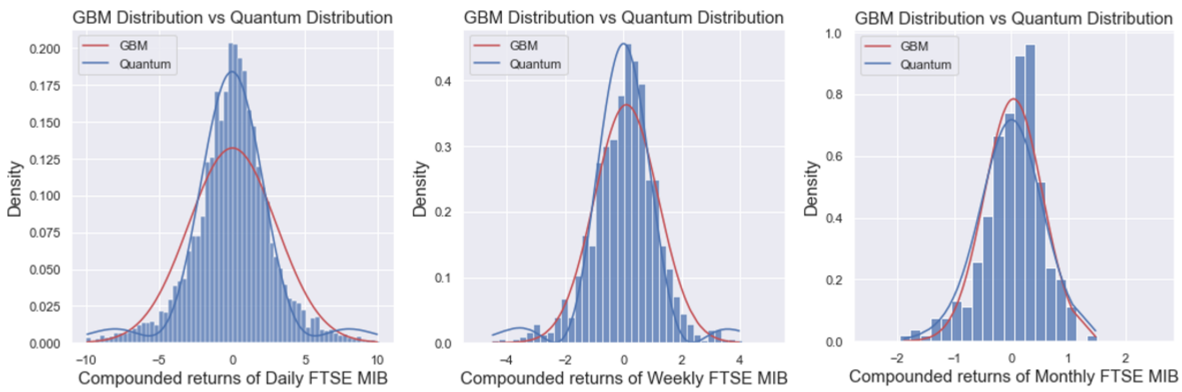

Figure 2.9: GBM vs Quantum FTSE MIB.

Looking at the three graphs we can observe that for Daily and for Weekly data the Quantum model recovers a better fit than the GBM, being more able to describe both the excess of the kurtosis and the fat tails of the distributions. In the Monthly graph, instead, neither of them succeeds to recover the features of the distribution, missing the skewness and the peaks.

Running the Quantum Genetic Algorithm for **70 generations** and the GBM Genetic Algorithm for **30 generations**, we have obtained the following fitness results:

Model Fitness	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM Fitness	2.51	1.63	0.75
Quantum Fitness	2.45	1.56	0.71

Figure 2.10: Fitness FTSE MIB.

We can behold that the GBM Fitness is always greater than the Quantum one, resulting instead in a better fitting along the Monthly Data. Furthermore both of the values diminish for the increasing of the holding period, so the description of the model becomes better both for the models.

Models	Parameters	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM	μ	1.54	0.66	0.29
	σ^2	3.02	1.10	0.51
Quantum	C_0	2.95574263e-01	1.85527616e-01	7.34796740e-01
	C_1	-6.08212208e-09	-1.62360788e-05	-1.90384712e-04
	C_2	2.88381983e-07	-2.79653139e-03	7.74827727e-03
	C_3	1.00442719e-05	-6.18980284e-04	-5.80268105e-04
	C_4	1.01458532e-02	7.21915335e-03	3.47466577e-03
	$m\omega$	5.43739597e-36	2.94920022e-35	1.56446937e-34

Figure 2.11: Estimated Parameters FTSE MIB.

The parameter C_0 has a minimum in the Weekly data, then increasing in the Monthly data. The parameter C_1 remains negative but increases in the absolute value, C_2 becomes negative in the Weekly remaining positive in Daily and Monthly, C_3 passes from a positive value to negative ones, C_4 remains positive and increases like obviously the product between mass and frequency.

	P_0	P_1	P_2	P_3	P_4
$\tau = 1$	9.988231e-01	4.229272e-16	9.508053e-13	1.153432e-09	1.176880e-03
$\tau = 5$	9.982506e-01	7.645128e-09	2.268098e-04	1.111158e-05	1.511454e-03
$\tau = 20$	9.998657e-01	6.712300e-08	1.111776e-04	6.235404e-07	2.235801e-05

Figure 2.12: Probabilities FTSE MIB.

The probability of the lowest volatility state remains almost constant during the holding period. There is an explanation about the higher kurtosis in the Weekly data observing the P_2 at $\tau = 5$, which is the maximum probability of the table except for the P_0. Moreover the odd probabilities in the Weekly data are the greatest one, evincing its excess kurtosis with respect to the other time scales like the Monthly ones. We can also point out that the standard deviation, so the volatility, of the Monthly data is the fewest, feature that could be outspread by the P_0 for $\tau = 20$, which is the major.

2.4 S&P500

The S&P500 is an American stock market index comprehensive of the 500 best American companies with the greatest market capitalization. Its name come from Standard & Poor, the credit risk agency which created it in 1957, and only companies that are exchanged at NYSE, at AMEX or at NASDAQ can be part of S&P500. In the end of June 2021, the total weight of the largest 10 companies in the index was about the 26,6% [11]. Some of the most important companies listed at S&P500 are:

- Amazon
- Moderna
- Tesla
- PayPal
- Google
- Facebook

Like NASDAQ COMPOSITE and FTSE MIB, S&P500 is a capitalization-weighted index, so not all the companies have the same weight that can be modified over time. In this thesis it was considered the behaviour of S&P500 from 21st September 1996 to 1st September 2021.

Figure 2.13: S&P500 behaviour from 1996 to 2021.

From a preliminary analysis of the compounded returns of the data, the following table was obtained:

	τ	N° of observations	Mean	Std	Skewness	Excess kurtosis	Entropy	Hurst
0	1	6278	0.075794	3.111397	-0.407285	10.384141	0.59434	0.266432
1	5	1302	0.073092	1.260043	-0.885556	7.109434	0.102376	0.292052
2	20	298	0.078727	0.562821	-0.794656	1.511342	0.001148	0.246224

Figure 2.14: Features of S&P500 compounded returns.

As we can observe from the previous figure the mean of the price remains almost constant increasing the holding periods, instead the standard deviation decreases, so the compounded returns become less volatile. As for the skewness, however, its value remains negative with a peak in absolute value compared to the Weekly data which have a more asymmetrical probability peak than the other time-frame data. The excess kurtosis remains more than 0, with a very leptokurtic distribution regarding the Daily data, diminishing to Weekly and Monthly ones, more flattened. The entropy of the series is also getting down from Daily to Monthly data, being about 600 times bigger denoting a very sharp enhancement in the possibility of forecasting of the Monthly returns. The Hurst coefficient, instead, remains almost the same in a range between 0.24 and 0.30 showing a good long-term correlation of the returns. It can also be seen that the GBM model excessively exaggerates the PDF tails fitting an excess of the probability for extreme events.

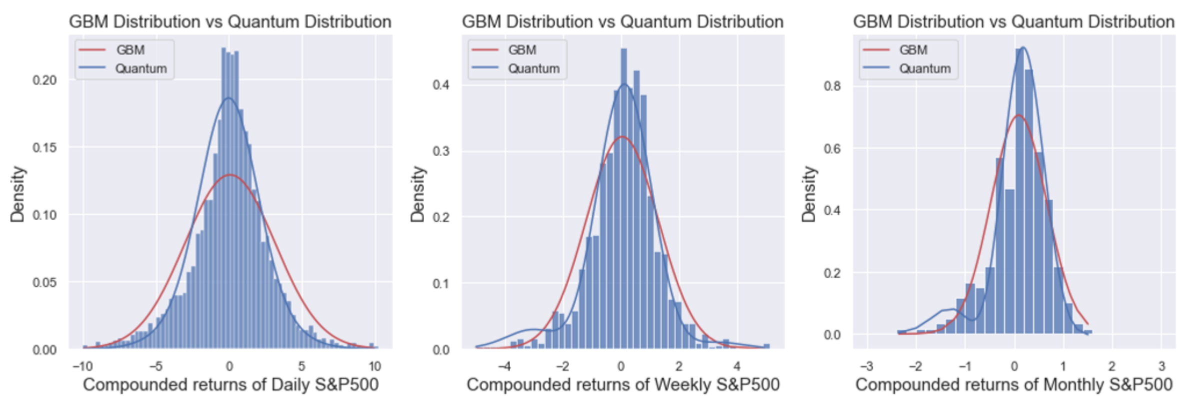

Figure 2.15: GBM vs Quantum S&P500.

In each of the three figures we can observe a good fitting for both of the

2 models. In particular the Quantum one manages to reach the peaks of distribution in the Weekly data, but above all in the Monthly ones where the fit is almost perfect. For Daily data the model is able to retrieve information on skewness not on kurtosis, still obtaining an excellent fit.
Running the Quantum Genetic Algorithm for **70 generations** and the GBM Genetic Algorithm for **40 generations**, we have obtained the following fitness results:

Model Fitness	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM Fitness	2.55	1.65	0.84
Quantum Fitness	2.46	1.58	0.80

Figure 2.16: Fitness S&P500.

The Fitness function, or KL divergence, is decreasing augmenting the holding period due to the obvious lowering of the number of data. We can furthermore observe that Quantum Fitness is always less than GBM Fitness, resulting in a better description of returns.

Models	**Parameters**	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM	μ	1.64	0.67	0.38
	σ^2	3.09	1.24	0.56
Quantum	C_0	1.8376933e-01	3.0561478e-01	5.3361438e-01
	C_1	-4.833199e-04	2.7362455e-05	5.7716634e-02
	C_2	1.6827885e-05	-2.426895e-05	-2.240663e-04
	C_3	-2.159198e-05	-8.870327e-03	-3.825931e-02
	C_4	1.0196674e-03	7.4584687e-03	8.5400184e-03
	$m\omega$	1.0055428e-35	3.1011486e-35	1.6166916e-34

Figure 2.17: Estimated Parameters S&P500.

Analyzing the GBM parameters we can observe a decreasing both of the μ and σ^2, For the Quantum parameters, instead, C_0 has a minimum in the Daily parameters and increases along the holding periods, C_1 is negative for the Daily, but becomes positive augmenting his value, C_2 follows an opposite direction, being positive for Daily data and then becoming negative, C_3 remains negative with a minimum in Weekly data, C_4 remains positive increasing its value. The product between mass and angular frequency increases as always.

	P_0	P_1	P_2	P_3	P_4
$\tau = 1$	9.999623e-01	6.916827e-06	8.384878e-09	1.380456e-08	3.078609e-05
$\tau = 5$	9.985640e-01	8.004557e-09	6.296934e-09	8.412139e-04	5.947391e-04
$\tau = 20$	9.831914e-01	1.150230e-02	1.733550e-07	5.054258e-03	2.518260e-04

Figure 2.18: Probabilities S&P500.

The probability of the first oscillator with $n = 0$, the gaussian one, decreases over time in favour of the other odds. The one linked at the state with $n = 1$, instead, increases very much from Daily to Monthly passing from 10^-6 to 10^-2 falling into a minimum in Weekly, which is also repeated in the state with $n = 2$. The state with $n = 3$, instead, increases its probability during the holding periods unlike the state with $n = 4$ which also becomes more probable passing from Daily to Monthly but having a local maximum in probability at Weekly data.

S&P500 has a maximum value in absolute value of skewness during the Weekly data but this is not explained by the table of probabilities. Unlike FTSE MIB, in fact, the state with maximum odd sum of probabilities does not correspond to that with maximum skewness.

2.5 EURO STOXX 50

EURO STOXX 50 is an European stock market index made up of the companies which are leader in their own sector. Like the other indices described, its composition is not fixed and varies to have the largest fifty companies in terms of capitalization and liquidity. Financial institutions like ECB (European Central Bank), FED (Federal Reserve Bank) esteem the EURO STOXX 50 as a benchmark to track the trend of the economic health of the most wealthy countries in the EU.
Updated at March 2021 the percentage composition by country is [12]:

- France $17/50 = 34\%$
- Germany $16/50 = 32\%$
- Netherlands $6/50 = 12\%$

- Spain 4/50 = 8%
- Italy 3/50 = 6%
- Ireland 2/50 = 4%
- Belgium 1/50 = 2%
- Finland 1/50 = 2%

In this thesis it was considered the behaviour of EURO STOXX 50 from 30th March 2007 to 1st September 2021.

Figure 2.19: EURO STOXX 50 behaviour from 2007 to 2021.

From a preliminary analysis of the compounded returns of the data, the following table was obtained:

	τ	N° of observations	Mean	Std	Skewness	Excess kurtosis	Entropy	Hurst
0	1	3608	0.000257	3.668635	-0.322798	7.656829	0.758533	0.270547
1	5	753	0.000738	1.597938	-1.370636	9.389112	0.183729	0.30746
2	20	173	-0.003233	0.666522	-0.518803	1.108777	0.002948	0.279615

Figure 2.20: Features of EURO STOXX 50 compounded returns

Being data available on Yahoo Finance since 2007, the number of observations in EURO STOXX 50 is less than the number of the other series analyzed before. We can see that the mean is a very little value, positive for

Daily and Weekly data but negative for Monthly data, so unlike the other indices the distribution will be more centered to the zero value, so to a constant value of the price. The standard values of the series get down increasing the holding period, decreasing the volatility of the data as the other indices. Regarding the skewness, instead, it has a maximum in its absolute value during the Weekly data with a minimum in the Monthly data. This behaviour is the same as the excess kurtosis which recommends about three leptokurtic distributions with a maximum in the Weekly kurtosis data. The entropy diminishes of 375 times between Daily and Monthly data, denoting a more repetitive pattern in Monthly data. Moreover, the Hurst coefficient remains almost constant between 0,27 and 0,30, indicating long term correlations about the data.

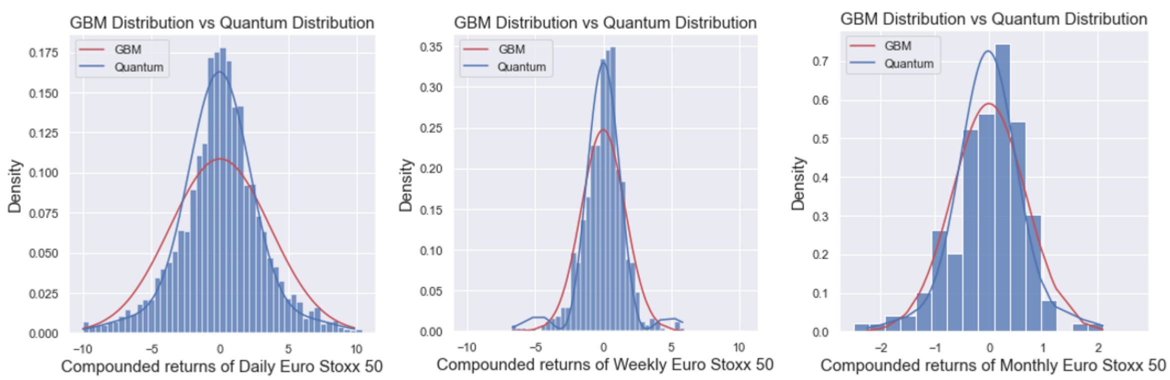

Figure 2.21: GBM vs Quantum STOXX 50.

As we can observe from the fits, the Quantum Model is always better than the GBM one, expect in the Daily data in which the skewness is the lowest one. Regarding the kurtosis, instead, the blue line manages to describe the features of the leptokurtic distribution in each of the three case, which GBM does not.
Running the Quantum Genetic Algorithm for **70 generations** and the GBM Genetic Algorithm for **40 generations**, we have obtained the following fitness results:

Model Fitness	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM Fitness	2.72	1.89	1.01
Quantum Fitness	2.63	1.82	0.97

Figure 2.22: Fitness EURO STOXX 50.

The better result of the Quantum fit is confirmed by the values of Quantum Fitness which are always lower than the GBM one.

Models	Parameters	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM	μ	1.64	0.67	0.38
	σ^2	3.09	1.24	0.56
Quantum	C_0	2.125545e-01	2.19104e-01	1.779147e-02
	C_1	-1.02965e-03	-2.6392e-04	-2.67309e-04
	C_2	3.065739e-06	2.80382e-04	-8.21091e-05
	C_3	1.090638e-07	-6.6355e-04	7.520871e-05
	C_4	3.359154e-03	8.67406e-03	2.721682e-04
	$m\omega$	6.1391464e-36	1.614293e-35	1.215743e-34

Figure 2.23: Estimated Parameters EURO STOXX 50.

The parameters show a decreasing both of the mean and of the standard deviation of the GBM model. It confirms the diminishing of the volatility which was observed in the table before. Regarding Quantum model parameters we can notice that the Gaussian parameter C_0 increases from Daily to Monthly data, and then decreases touching a minimum in Monthly data, C_1 remains negative lowering its absolute value getting nearest to 0 while C_2 gets positive with a minimum in Daily data and a maximum in Monthly. For the last parameters we can observe that C_3 has a peculiar behaviour changing sign from Daily to Weekly and coming back positive at Monthly, C_4 follows the trend of C_2 having a maximum in the Weekly data, and the product between mass and harmonic oscillator increases.

	P_0	P_1	P_2	P_3	P_4
$\tau = 1$	9.997268e-01	2.345990e-05	2.079750e-10	2.632100e-13	2.496896e-04
$\tau = 5$	9.984229e-01	1.448656e-06	1.634968e-06	9.157239e-06	1.564792e-03
$\tau = 20$	9.995013e-01	2.256255e-04	2.128839e-05	1.786058e-05	2.339021e-04

Figure 2.24: Probabilities EURO STOXX 50.

The probabilities can explain the excess kurtosis because in the Weekly data we can point out that the sum of the odds of the even states are at the peak, contributing to the PDF to be leptokurtic. The probability of the state corresponding to $n = 0$ remains almost constant having a minimum in the Weekly data, which is the same behaviour of both P_1 and P_2 with a maximum in the Monthly data. P_3 in the Daily data, instead, is negligible being of the order of 10^{-13} and increasing until reaching a maximum in the Monthly data while P_4 has a maximum in Weekly data and a minimum in the Monthly data.

2.6 DOW JONES

The DOW JONES is the second-oldest stock market index of US stocks, being founded in 1896, and the most known index at New York Stock Exchange. It was created by Charles Dow, founder of 'The Wall Street Journal' and Edward Jones, an American statistician with a different composition as compared to the other indices analyzed in this thesis. NASDAQ or S&P500 are capitalization-weighted indices, while DOW JONES price is obtained by the 30 main titles in Wall Street, including:

- Apple
- Boeing
- Coca-Cola
- Nike
- Visa
- Walt Disney

As written in [13], DOW JONES is a really democratic index giving the same importance to each of the stocks inside. To reflect the overall health of the market in the most useful weight, there is a choice of 30 different companies that change in time both due to the market trend and also to the selection of the editors of 'The Wall Street Journal'. The calculation of the DOW JONES price involves the use of the divisor, avoiding high value fluctuations of the index when a stock is added, delisted, split of a company or mergers. To understand let's make a practical example on Daily prices:

Today there are 3 companies listed on DOW JONES; A,B,C respectively with prices 30\$,25\$, 45\$. If yesterday there were no changes in the index, today the price will be the average of the individual companies, so:

$$DJIA = \frac{\sum_{i=0}^{n} P_i}{n} \tag{2.2}$$

with:

- DJIA: Daily price of DOW JONES
- P_i: price of each single stock
- $n = 3$: number of the stocks

Hence today $DJIA = 33,3\$$. Tomorrow the company D will join the DOW JONES index with a price of 15\$. According to [13] the new index value will be:

$$DJIA_{new} = \frac{\sum_{i=0}^{n_{new}} P_i}{D} \tag{2.3}$$

with:

- $DJIA_{new}$: new Daily price of DOW JONES
- P_i: price of each single stock
- n_{new}: updated number of the stocks
- $D = \frac{\sum_{i=0}^{n_{new}} P_i}{DJIA}$:

Thus $D = 115\$/33,3\$ = 3,45$ and the new Daily price of DOW JONES will be: $DJIA_{new} = 115\$/3,45 = 33,3\$$ so its value will remain unchanged for the first day and then the divisor of 3,45 will be constant until the number of the stock is equal to 4.
In this thesis it was considered the behaviour of DOW JONES from 21st September 1996 to 1st September 2021.

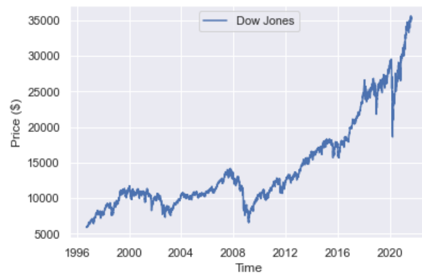

Figure 2.25: DOW JONES behaviour from 1996 to 2021.

From a preliminary analysis of the compounded returns of the data, the following table was obtained:

	τ	N° of observations	Mean	Std	Skewness	Excess kurtosis	Entropy	Hurst
0	1	6278	0.072098	2.99616	-0.402272	12.28345	0.564587	0.268398
1	5	1302	0.069529	1.250499	-1.046096	9.023872	0.097569	0.301511
2	20	298	0.074943	0.554853	-0.7455	1.549256	0.000634	0.207858

Figure 2.26: Features of DOW JONES compounded returns.

As we can observe from the table, the mean remains almost equal between Daily, Weekly and Monthly data with a slight minimum in the Weekly data. The standard deviation diminishes as the other indices, the skewness remains negative with a peak in the Weekly data and an absolute value minimum in Daily. The kurtosis, instead, has the same behaviour of S&P500 reaching a maximum in Daily data and a minimum in Monthly data, the entropy falls from 0.56 to a value 934 times lower, while the Hurst coefficient remains between 0.20 and 0.30. It is slight less than the other indices having a similar meaning, so long-term correlation between historical returns.

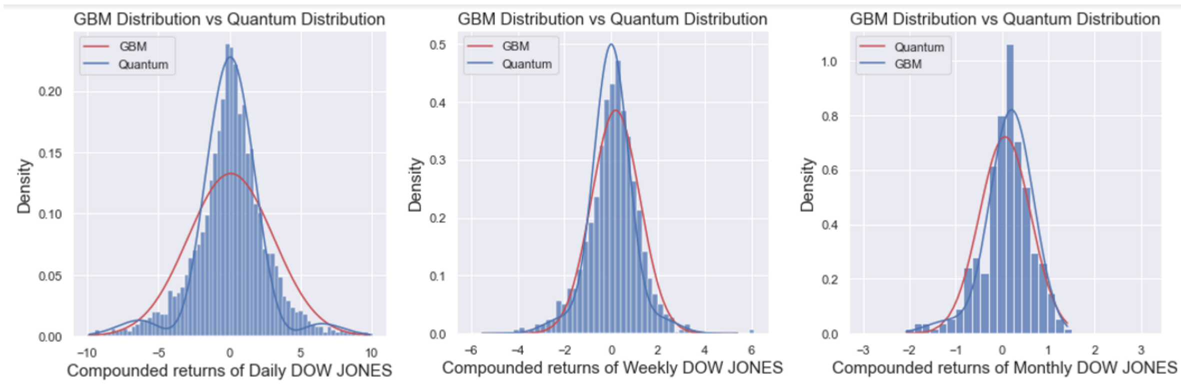

Figure 2.27: GBM vs Quantum DOW JONES.

In the Daily data histograms the Quantum PDF manages to reach both the kurtosis and the skewness of the data describing the peak in a very good way, in the Weekly data both the GBM and the Quantum model do their job, but GBM manages to catch the skewness. In the Monthly data, instead, Quantum model is the better fitting, not peaking the PDF but understanding its features.

Running the Quantum Genetic Algorithm for **75 generations** and the GBM Genetic Algorithm for **40 generations**, we have obtained the following fitness results:

Model Fitness	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM Fitness	2.52	1.64	0.83
Quantum Fitness	2.42	1.59	0.78

Figure 2.28: Fitness DOW JONES.

The better result of the Quantum fit is confirmed by the values of Quantum Fitness which are always lower than the GBM one.

Models	Parameters	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM	μ	1.60	0.71	0.34
	σ^2	3.01	1.03	0.55
Quantum	C_0	0.37029435	1.575012e-01	1.7747324e-01
	C_1	-0.00062383	-2.744753e-04	3.0744038e-02
	C_2	-0.0004072	9.1274296e-07	-2.729167e-03
	C_3	-0.00321457	7.878584e-07	-7.312144e-03
	C_4	0.0127007	2.9149426e-03	8.8104952e-04
	$m\omega$	8.5429690e-36	5.5327356e-35	1.5341070e-34

Figure 2.29: Estimated Parameters DOW JONES.

The parameters show a decreasing both of the mean and of the standard deviation of the GBM model. It confirms the diminishing of the volatility which was observed in the figure before. The positive parameter C_0 goes down from Daily to Weekly reaching a minimum and then growing at Monthly data, C_1 is negative both for Daily and Weekly getting down and then becoming positive at Monthly data, C_2 alternates its sign being negative at Daily data, then positive at Weekly and so negative at Monthly. C_3 follows the same behaviour of C_2, while C_4 remains positive but diminishing from the order of 10^{-2} at Daily data to the order of 10^{-4} at Monthly data while the product between mass and angular frequency increases due to ω.

	P_0	P_1	P_2	P_3	P_4
$\tau = 1$	9.987457e-01	2.834575e-06	1.207730e-06	7.526755e-05	1.174940e-03
$\tau = 5$	9.996545e-01	3.035906e-06	3.357211e-11	2.501370e-11	3.424062e-04
$\tau = 20$	9.690223e-01	2.907969e-02	2.291544e-04	1.644968e-03	2.388188e-05

Figure 2.30: Probabilities DOW JONES.

The probability of the first state P_0 has its minimum in Monthly data, so it is more probable to have volatility states there. P_1 is almost the same between Daily and Weekly data, becoming 10^4 times larger at Monthly data while P_2 has a minimum at Weekly data and a maximum at Monthly. For the last probabilities P_3 has a strong minimum at Weekly data, being of the order of 10^{-11} while P_4 diminishes in a regular way from Daily to Monthly. These probabilities manages to explain the excess of kurtosis in Daily data, being $P_4 \sim 10^{-3}$ and its falling trend but not the skewness.

2.7 Gold

Gold is a chemical element with symbol Au and 79 as atomic number. Gold is the safest financial investment to achieve because being in a limited quantity does not have large swings of price, being also a shield for inflation, the continuous increase of prices. It is a commodity, its price is determined by the continue trades on financial markets and its exchange rate is expressed in dollars per ounce, which corresponds to about 31 grams. One of the most important features is the gold purity when is in alloy with other metals like platinum or copper. This property is measured by percentage with respect to 24 carats which is the pure gold, so the green gold which is composed of 75% gold, 12,5% silver and 12,5% copper is 18 carats.

As described in [14], gold price is fixed by 'gold fixing', valuating instruments like rates and price of other precious metals like iridium, palladium and platinum. The quotation determines the value of the price each day twice by Iba (Ice Benchmark Administation). In this thesis we have considered gold prices from Yahoo Finance, which exhibits its price by 'COMEX Delayed Price' with currency in USD, so it is exposed to fluctuations of dollars. As written in [15], COMEX is the primary futures and options market for trading metals such as gold, silver, copper, and aluminum and is the division responsible for metal trading.
In this thesis it was considered the behaviour of Gold from 30th August 2000 to 1st September 2021.

Figure 2.31: Gold behaviour from 2000 to 2021.

From a preliminary analysis of the compounded returns of the data, the following table was obtained:

	τ	N° of observations	Mean	Std	Skewness	Excess kurtosis	Entropy	Hurst
0	1	5196	0.091897	2.822271	-0.29565	5.401066	0.608922	0.268054
1	5	1096	0.086616	1.215252	-0.329882	1.956338	0.104812	0.22904
2	20	213	0.112075	0.636192	-0.313444	0.933476	0.003022	0.249389

Figure 2.32: Features of Gold compounded returns.

The behaviour of Gold is almost the same as the financial indices analyzed before with a mean that does not have so many variations between Daily, Weekly and Monthly data with a thin peak in Monthly ones. The standard deviation diminishes during the holding period and is smaller than the indices one, about 77% of the EURO STOXX 50 and 90% of the S&P500. The skewness remains negative with a peak in the Weekly data like in the indices, while the excess kurtosis gets down from Daily to Monthly data like the DOW JONES. Analyzing the last 2 features, the entropy of Daily data is almost 203 times the one of Monthly data while the Hurst coefficient is slightly the same from Daily to Monthly, in a range between 0,22 and 0,26.

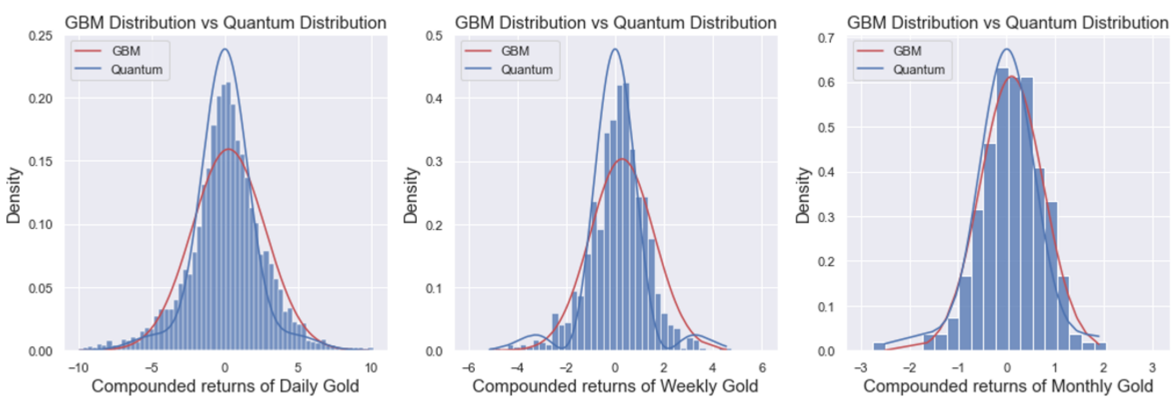

Figure 2.33: GBM vs Quantum Gold.

As we can see from the plots, the Quantum model exceeds the kurtosis for all the 3 holding periods, expecially in the Daily data. The GBM model, instead, captures the skewness of the data better than the Quantum model except for the Monthly ones, in which Quantum model provides a very good visual result.

Running the Quantum Genetic Algorithm for **70 generations** and the GBM Genetic Algorithm for **40 generations**, we have obtained the following fitness results:

Model Fitness	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM Fitness	2.46	1.61	0.962
Quantum Fitness	2.40	1.59	0.967

Figure 2.34: Fitness Gold.

As in the other indices we can observe that both of the two Fitness functions diminishes over the holding periods, but differently from them the Monthly ones are very similar, confirming the curves that almost overlap.

Models	Parameters	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM	μ	1.51	0.94	0.43
	σ^2	2.51	1.31	0.65
Quantum	C_0	4.25122107e-01	7.510037205e-03	4.06112201e-01
	C_1	5.48210718e-04	-4.77807357e-06	-1.92021397e-06
	C_2	7.51594246e-08	-3.18247507e-05	-1.78112809e-05
	C_3	7.77378612e-08	5.76941101e-08	-1.06403988e-03
	C_4	8.98896621e-03	2.92926931e-04	8.88848212e-03
	$m\omega$	1.19738514e-35	3.44123880e-35	8.98169993e-35

Figure 2.35: Estimated Parameters Gold.

Excluding NASDAQ, the standard deviations of the GBM models are smaller than the ones of the indices, having a diminution with the mean from Daily to Monthly data. Discussing the parameters, instead, we can observe that C_0 increases from Daily to Weekly but having a global minimum in Monthly data, C_1 is positive at Daily becoming negative in Weekly and Monthly, diminishing also in absolute value. C_2 also has a weird behaviour being positive at Daily but negative at Weekly and Monthly, increasing its absolute value. C_3, instead, is positive in Daily and Weekly but turns negative in Monthly being also 10^5 times larger while C_4 has a maximum in Daily data and a minimum in Weekly, increasing a little bit of a 10 power in Monthly. The product between mass and angular frequency as always increases during the holding periods.

	P_0	P_1	P_2	P_3	P_4
$\tau = 1$	9.995514e-01	1.662159e-06	3.124242e-14	3.342280e-14	4.468865e-04
$\tau = 5$	9.984626e-01	4.041599e-07	1.792989e-05	5.892649e-11	1.519029e-03
$\tau = 20$	9.995143e-01	2.234581e-11	1.922592e-09	6.861400e-06	4.787978e-04

Figure 2.36: Probabilities Gold.

The probability of the less volatility state, P_0, is almost the same between the different time scales with a little minimum in Weekly, while P_1 diminishes of 10^5 times from Daily to Monthly, P_2 is quite negligible in Daily data but increases in Weekly of 10^9 times decreasing in Monthly. P_3 is very similar to P_2 in Daily increasing over the holding periods, having then a warring trend respect to P_1, while $P_4$is almost the same with a peak in Weekly data. The probabilities don't manage to describe neither the trend of the kurtosis or the skewness of the data.

2.8 Petroleum

Petroleum is a fossil fuel that occurs naturally beneath the earth's surface, which is very important due to its social and economical geopolitical consequences, being in the same time cause of desire and war. It is the most traded commodity all over the world, becoming a product of speculation for investors. All over the years there were a lot of fluctuations of its price, giving rise to periods of very large volatility as referred in [16], basically increasing the supply except particular events like 11th September 2001 of COVID -19 pandemic, when petroleum has followed the collapse of the market. The prices analyzed in this thesis have been downloaded from Yahoo Finance, which follow the trend of crude oil futures, which is not the price it would cost to buy oil at that time. Like the other commodities, the wholesale manifacturers need to fix a price before the delivery to the final customer, bonding for a future transaction
In this thesis it was considered the behaviour of Petroleum from 30th August 2000 to 1st September 2021.

Figure 2.37: Petroleum behaviour from 2000 to 2021.

As we can see from the figure, on 21st April 2020 the price of the petroleum became negative, touching -37 dollars per barrel. In the historical time series we have decided to avoid that price except for the forward plot. The main cause of this particular phenomena is due to the sudden fall of the world consumption of gas and oil due to the restrictions, not followed by a drop of the world production. The principal effect was the payment from the producers to the oil tankers to preserve the barrels to avoid the interruption of production, which would have resulted in serious economic damages.

From a preliminary analysis of the compounded returns of the data, the following table was obtained:

	τ	N° of observations	Mean	Std	Skewness	Excess kurtosis	Entropy	Hurst
0	1	5245	0.036565	6.963114	-1.889848	57.711675	1.261566	0.308511
1	5	1097	0.034965	2.74601	-0.794858	5.598421	0.509294	0.265501
2	20	213	0.0518	1.561866	-1.242175	13.293254	0.139173	0.366205

Figure 2.38: Features of Petroleum compounded returns.

We can observe that the properties of petroleum compounded returns are very different to each of the prices analyzed before. The mean is almost the same like the other time series, but the standard deviation is very high, 1,77 times more than the NASDAQ one, the most volatile compounded returns index analyzed until now. The skewness is also much elevated in absolute value touching a Daily data value of -1.89, with a minimum in Weekly data. The most shocking element of the table is surely the excess kurtosis for Daily data of 57 together with its trend, going to an ordinary minimum for Weekly data and increasing again in the Monthly data. Very high kurtosis is due to the presence of many outliers in the data, so days when returns far from the average were high. This is a volatility index of returns and therefore also of prices, also confirmed by the high standard deviation. Furthermore the entropy is higher than the other indices, being 1.26 for Daily data and lowering to 9.6 times less in the Monthly data, slight fall respect to the other series. The Hurst coefficient, instead, is under 0.5 but in average higher than the other time series.

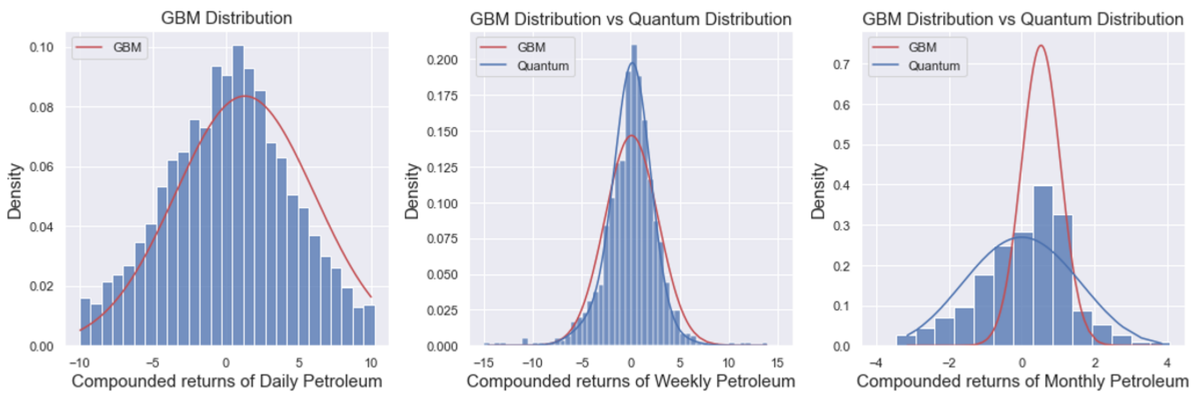

Figure 2.39: GBM vs Quantum Petroleum.

As we can observe from the fitting, the parameters of the Quantum model for the Daily data were not obtained even by decreasing the number of generations in the Genetic Algorithm. A possible cause is excessive kurtosis due to which the algorithm had excessive uncertainty in estimating the parameters giving an error. The fitting is already very good for the Weekly data, the ones with a lower kurtosis and skewness, getting back in error for the Monthly data with GBM model exaggerating the peak of the distribution and the Quantum model not fitting the right skewness.

Running the Quantum Genetic Algorithm for **75 generations** and the GBM Genetic Algorithm for **30 generations**, we have obtained the following fitness results:

Model Fitness	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM Fitness	3.35	2.42	1.86
Quantum Fitness	\	2.37	1.75

Figure 2.40: Fitness Petroleum.

As always the GBM Fitness values are larger than the Quantum ones, except for the Daily data we cannot discuss nothing about. We can then observe that Fitness values are larger for Petroleum with respect to the other indices.

Models	Parameters	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM	μ	3.72	1.45	0.80
	σ^2	4.78	2.71	0.53
Quantum	C_0	\	2.10822021e-01	3.70164278e-01
	C_1	\	4.05881203e-06	1.41707005e-04
	C_2	\	1.34347018e-04	6.58119529e-04
	C_3	\	-2.95608454e-03	3.09827118e-06
	C_4	\	1.86978612e-03	1.42397106e-02
	$m\omega$	\	1.049955136e-35	9.274642265e-36

Figure 2.41: Estimated Parameters Petroleum.

For the GBM model we can observe a weird behaviour of the Monthly parameters with μ larger than σ^2 for the first time. Discussing the Quantum model instead, we don't have the Daily parameters and we can observe a similar trend with respect to the other indices of C_0, C_1, C_2 and C_4, increasing from Weekly to Monthly and C_3 changing sign from negative to positive. The product between mass and angular frequency is getting down between Weekly and Monthly data being an exception to the theory for the first time.

	P_0	P_1	P_2	P_3	P_4
$\tau = 1$	\	\	\	\	\
$\tau = 5$	9.997244e-01	3.705494e-10	4.059797e-07	1.965540e-04	7.863795e-05
$\tau = 20$	9.985190e-01	1.463355e-07	3.156290e-06	6.995297e-11	1.477643e-03

Figure 2.42: Probabilities Petroleum.

The table of probabilities, despite the lack of Daily values, explains the excess of kurtosis because the sum of even probabilities, except P_0 is more at Monthly data than at Weekly data, being of the order of 10^{-3}, highest even probability met until this series.

2.9 Amazon

Amazon, which is the first stock analyzed in the thesis, is the biggest internet company of the world founded from Jeff Bezos in 1995 and entering into S&P500 in 2005. It started as an online bookseller, but then also CD's, DVD, videogames, software, electronic products, comic books, clothing, furnishing and many other objects were available on it.
In this thesis it was considered the behaviour of Amazon from 30th August 2000 to 1st September 2021.

Figure 2.43: Amazon behaviour from 2000 to 2021.

Unlike other time series like Petroleum when data were not available before 2000, Amazon prices were present on Yahoo Finance since 1997. We decided to start our analysis with data starting on 1st January 2020 because, as described in [17], Amazon has split its stocks three times in the history:

Amazon stock split history

	Split ratio	Price before split
2 June 1998	2:1	$85.68 (1 June 1998)
5 January 1999	3:1	$354.96 (4 January 1999)
2 September 1999	2:1	$119.06 (1 September 1999)

Figure 2.44: Amazon stock splits, image taken from [17].

A stock split like Amazon one consists in stepping up its number of shares decreasing its price of the same number. So an investor who had 100 shares of Amazon on 1 June 1998 pricing $\$85,68$, next day he would have had 200 shares of value $\$42,84$. We therefore decided to analyze time series after the splits because importing from Yahoo Finance prices were displayed as 'not a number'. In particular this trouble was not present in Daily Data, but in Weekly and Monthly, because Amazon stock split dates coincide with weekend and month-end, making analysis useless.

From a preliminary analysis of the compounded returns of the data, the following table was obtained:

	τ	N° of observations	Mean	Std	Skewness	Excess kurtosis	Entropy	Hurst
0	1	5451	0.176938	8.02729	0.46491	12.850953	1.3102	0.268247
1	5	1131	0.170555	3.405626	0.056242	8.813533	0.665802	0.296232
2	20	259	0.194225	1.648128	-0.519183	3.016752	0.218853	0.283709

Figure 2.45: Features of Amazon compounded returns.

As we can observe the mean remains almost equal with a peak in Monthly data, having values bigger than the other time series. Furthermore the standard deviation is basically higher than indices, because stocks have more volatility. For the first time we see also data with positive skewness, so many compounded returns have value less than the mean in Daily and Weekly time frame, turning into negative for Monthly. The kurtosis goes down from Daily to Monthly with a very high value in Daily data while the entropy, being also very large unlike indices, goes down of 5.9 times from Daily to Monthly. The Hurst coefficient, instead, remains almost equal between the holding periods.

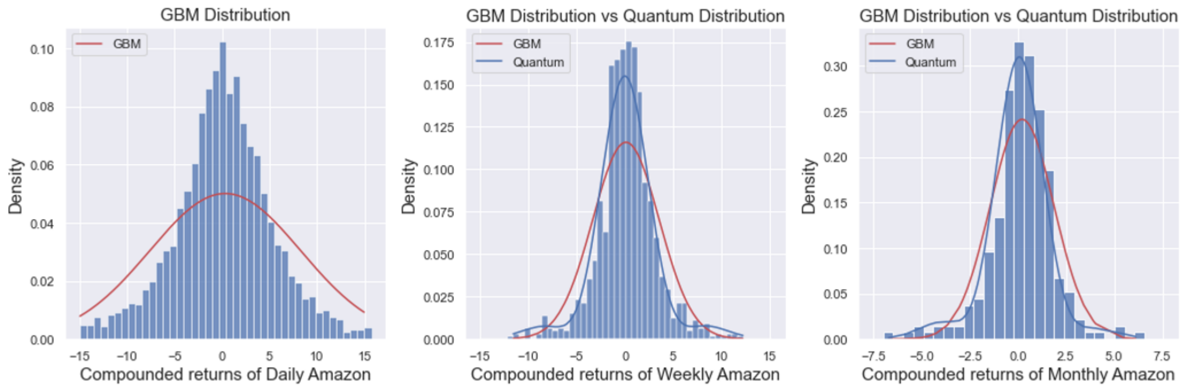

Figure 2.46: GBM vs Quantum Amazon.

Like Petroleum, the parameters of the Quantum model for the Daily data were not obtained maybe due to the excessive standard deviation.
For Daily data we can observe a good fit in terms of the skewness for the GBM Model, while for Weekly and Monthly data the Quantum Model performs visually better than GBM. Running the Quantum Genetic Algorithm for **75 generations** for the three holding periods and the GBM Genetic Algorithm for **40 generations** only for Weekly and Monthly data, we have obtained the following fitness results:

Model Fitness	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM Fitness	3.50	2.64	1.91
Quantum Fitness	\	2.54	1.82

Figure 2.47: Fitness Amazon.

Like the other series the GBM fitness is always higher than the Quantum one, except for Daily Data due to the troubles explained before.

	P_0	P_1	P_2	P_3	P_4
$\tau = 1$	\	\	\	\	\
$\tau = 5$	9.992833e-01	3.487904e-09	1.311684e-12	7.091133e-07	7.159245e-04
$\tau = 20$	9.992088e-01	7.536861e-06	2.450270e-07	1.974066e-04	5.860052e-04

Figure 2.48: Estimated Parameters Amazon.

The GBM distribution is more flattened in Amazon than in the other time series, as we can see comparing their standard deviations. The Quantum coefficients, instead, follow a regular trend with an increase of C_0 and $m\omega$, while C_1 and C_3 turn from negative to positive growing its absolute value, C_2 from positive to negative and C_4 decreasing a little bit.

	P_0	P_1	P_2	P_3	P_4
$\tau = 1$	\	\	\	\	\
$\tau = 5$	9.992833e-01	3.487904e-09	1.311684e-12	7.091133e-07	7.159245e-04
$\tau = 20$	9.992088e-01	7.536861e-06	2.450270e-07	1.974066e-04	5.860052e-04

Figure 2.49: Probabilities Amazon.

Obviously the probabilities for Daily data were not available. We can see that the even probabilities, except the gaussian one, decrease over the holding periods explaining the fall of kurtosis from Weekly to Monthly data. The odd probabilities for Monthly data, furthermore, are larger than the one of Weekly data, possibly explaining the increase in absolute value of the skewness but not its change of sign.

2.10 Tesla

Tesla is an American luxury electric vehicle company whose name is a tribute to Serbian inventor Nikola Tesla. This company is also part of both S&P500 with a weight of 1,5% and NASDAQ COMPOSITE with a weight of 4,25% at September 2021.
In this thesis it was considered the behaviour of Tesla from 29th June 2010, data of the IPO (Initial Public Offering) on NASDAQ to 1st September 2021.

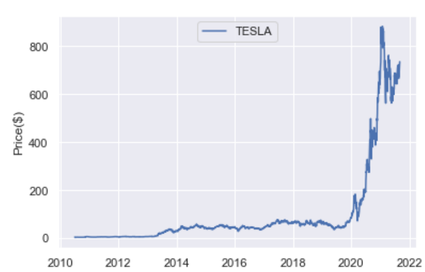

Figure 2.50: Tesla behaviour from 2000 to 2021.

From a preliminary analysis of the compounded returns of the data, the following table was obtained:

	τ	N° of observations	Mean	Std	Skewness	Excess kurtosis	Entropy	Hurst
0	1	2813	0.452115	8.903208	0.00628	5.909066	1.509334	0.248937
1	5	583	0.455226	3.724624	0.182544	1.873744	0.841507	0.248535
2	20	132	0.513352	2.005248	0.791189	1.235925	0.339639	0.250985

Figure 2.51: Features of Tesla compounded returns.

We have deleted from the dataset downloaded from Yahoo Finance the price referred to 31st August 2020 because, as described in [18], there have been five-for-one split of Tesla's common stock in the form of a stock dividend. The price of that day was then a NaN value causing the same problem of Amazon but for Monthly data.

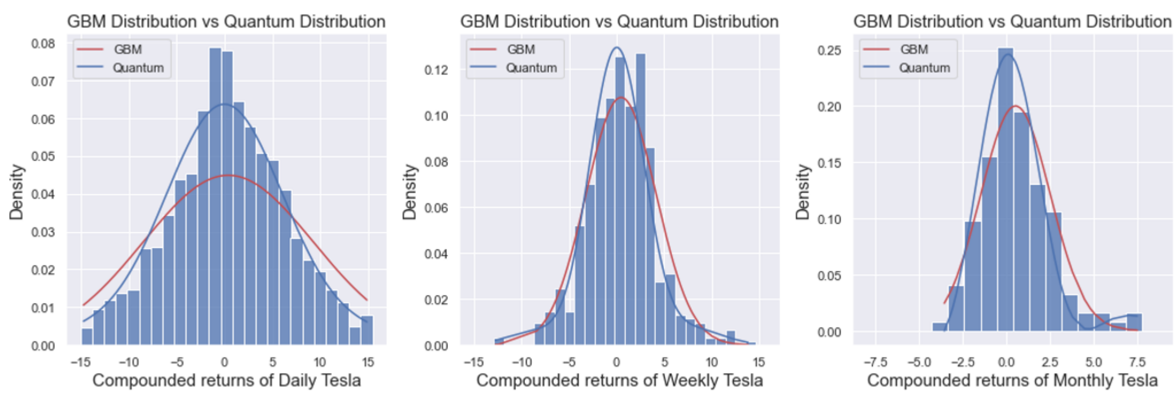

Figure 2.52: GBM vs Quantum Tesla.

As we can observe from the fitting, Quantum Model performs better than GBM Model in each of the holding periods data. In particular in Weekly and Monthly data it manages to recover the kurtosis of PDF, even if in Weekly it does not recognize the bimodal distribution. In Daily data, instead, it fits the skewness while not the kurtosis.

Running the Quantum Genetic Algorithm for **70 generations** and the

GBM Genetic Algorithm for **40 generations**, we have obtained the following fitness results:

Model Fitness	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM Fitness	3.60	2.73	2.11
Quantum Fitness	3.51	2.71	2.04

Figure 2.53: Fitness Tesla.

Like the other series the GBM fitness is always higher than the Quantum one, except for Daily Data due to the troubles explained before.

Models	Parameters	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM	μ	4.85	2.34	1.51
	σ^2	8.90	3.71	1.99
Quantum	C_0	9.24186005e-01	3.00983843e-01	0.55172816
	C_1	1.16885764e-03	5.57334615e-03	0.03859749
	C_2	3.98416609e-06	-4.33681930e-05	-0.0033051
	C_3	1.74820600e-05	7.60979666e-05	-0.00225803
	C_4	9.12538454e-03	5.58122474e-03	0.02539573
	$m\omega$	9.599448846e-37	3.661275815e-36	7.458598348e-36

Figure 2.54: Estimated Parameters Tesla.

The GBM parameters for Daily data are very large with respect to the other time series and following their trend the mean is always less than the standard deviation. In particular for Daily data the σ^2 is greater with respect to other time series, endorsing the thesis that stocks are much more volatile than time indices. Discussing about Quantum parameters, instead, we can observe that C_0 has a maximum in Daily data and a minimum in Weekly data, C_1 increases of 10^2 from Daily to Monthly, C_2 turns negative from Daily to Weekly increasing its absolute value, like C_3 which becomes negative in Monthly. C_4 and the product between mass and angular frequency, instead, remains positive and increase.

	P_0	P_1	P_2	P_3	P_4
$\tau = 1$	9.999009e-01	1.599415e-06	1.858288e-11	3.577860e-10	9.748561e-05
$\tau = 5$	9.993136e-01	3.426474e-04	2.074713e-08	6.387950e-08	3.436168e-04
$\tau = 20$	9.929841e-01	4.859708e-03	3.563373e-05	1.663228e-05	2.103844e-03

Figure 2.55: Probabilities Tesla.

We can observe that the probabilities of the state with $n = 0$ remain almost the same with a slight diminution in Monthly data. All the other probabilities, instead, increase over the holding period. In particular the sum of the odd probabilities is maximum in Monthly data, explaining the peak of skewness while the sum of even probabilities is not the maximum in Daily data, not explaining the excess of kurtosis.

2.11 Forex EUR - USD

As explained in [19], foreign exchange, also known as forex or FX, is the exchange of different currencies on a decentralized global market. It is one of the largest and most liquid financial markets in the world. Forex trading involves the simultaneous buying and selling of the world's currencies on this market.
It plays a very important role in foreign trade and business, as products or services purchased in a foreign country must be paid for using that country's currency.

Forex is one of the most traded markets in the world, with total average Daily turnover of over $ 5 trillion per day. The forex market is not based in a central or exchange location, and is open 24 hours a day, Sunday evening through Friday evening. Forex works traded in currency pairs. In this thesis we have analyzed the behaviour of EUR / USD (pound against US dollar), which was the most traded forex pair with 24% of the total market in 2019 according to the Bank of International Settlements (BIS) triennial survey.

It speculates whether the price of one country's currency will rise or fall against another country's currency and takes a position accordingly. Looking at the EUR / USD currency pair, the first currency (EUR) is called the "base currency" and the second currency (USD) is known as the "counter currency". Trading on forex means speculating that the price of the base currency will rise or fall relative to the counter currency [19]. So, in EUR / USD, if you think the pound will go up against the US dollar, and so that the price will rise, you buy the currency pair going long. If you instead think that the EUR will fall against the USD (or the USD will rise against the EUR) and the price will go down, you go short selling the currency pair.

In this thesis it was considered the behaviour of Forex EUR - USD from 12th January 2003 to to 1st September 2021.

Figure 2.56: Forex EUR - USD behaviour from 2003 to 2021.

From a preliminary analysis of the compounded returns of the data, the following table was obtained:

	τ	N° of observations	Mean	Std	Skewness	Excess kurtosis	Entropy	Hurst
0	1	4582	-0.00078	1.868586	0.557512	96.832389	0.200083	0.390544
1	5	921	-0.001711	0.661146	-0.300576	1.661437	0.00634	0.25063
2	20	212	-0.003875	0.336982	-0.346939	2.048526	0.0	0.277165

Figure 2.57: Features of Forex EUR - USD compounded returns.

As we can observe the mean is already unchanged between the 3 holding periods while the standard deviation or volatility is slightly going down from Daily to Monthly data. The skewness instead, turns from positive to negative showing a shift to the left of the center of the PDF, while the kurtosis, which has the most interesting behaviour, going from the highest result obtained in all the time series in Daily data down to an ordinary value in the other holding periods. The entropy, instead, has a very low values with respect to the other series going to 0 in the Monthly data, so it means that the price fluctuations will repeat very much in that holding period and that this series is optimum to forecast by algotrading.

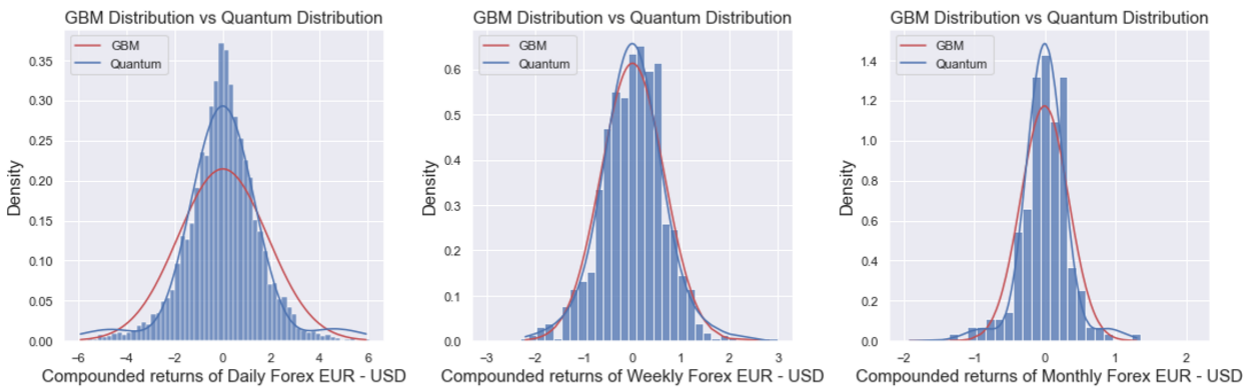

Figure 2.58: GBM vs Quantum Forex EUR - USD.

As we can observe from the fits, the Quantum Model is always better than the GBM one, managing to take the peak of distribution in Weekly and Monthly but not in Daily, maybe due to the very high kurtosis of the data.

Running the Quantum Genetic Algorithm for **70 generations** and the GBM Genetic Algorithm for **40 generations**, we have obtained the following fitness results:

Model Fitness	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM Fitness	2.04	1.00	0.33
Quantum Fitness	1.95	0.98	0.27

Figure 2.59: Fitness Forex EUR - USD.

The Fitness of the Quantum Model is always higher than the GBM one, and is lower than the other time series because the Forex EUR-USD data analyzed were available only since 2003, while for time indices like NASDAQ we started from 2003.

	P_0	P_1	P_2	P_3	P_4
$\tau = 1$	9.991576e-01	4.051694e-12	3.866445e-07	2.819243e-06	8.391211e-04
$\tau = 5$	9.999058e-01	2.999326e-09	1.653664e-07	9.680840e-08	9.389654e-05
$\tau = 20$	9.990680e-01	8.809665e-08	3.409670e-04	2.284579e-08	5.908630e-04

Figure 2.60: Estimated Parameters Forex EUR - USD.

The GBM parameters follow a very similar trend to the other time series with the mean less than the standard deviation, and both of the two parameters decreasing from Daily to Monthly data. In the Quantum Model, instead, the parameters C_0 and C_4 remain positive with a maximum in Monthly data, while C_1 turns positive from Weekly to Monthly and C_2 negative from Weekly to Monthly. $m\omega$, instead, increases along the holding periods.

	P_0	P_1	P_2	P_3	P_4
$\tau = 1$	9.991576e-01	4.051694e-12	3.866445e-07	2.819243e-06	8.391211e-04
$\tau = 5$	9.999058e-01	2.999326e-09	1.653664e-07	9.680840e-08	9.389654e-05
$\tau = 20$	9.990680e-01	8.809665e-08	3.409670e-04	2.284579e-08	5.908630e-04

Figure 2.61: Probabilities Forex EUR - USD

The probabilities in Forex EUR-USD confirm both the very high excess kurtosis in the Daily data, because the sum of P_2 and P_4 is maximum there, and also the maximum skewness in absolute value because also the sum of P_1 and P_3 is maximum.

2.12 Bitcoin

Bitcoin is a digital decentralized cryptocurrency invented in 2009 by Satoshi Nakamoto, a misterious and anonymous Japanese man who created this technology. The exchange of Bitcoins, digital money which does not have physical counterpart, is kept sicure by the use of cryptography, writing each transaction to a public ledger called blockchain. There is no need of a central control, like the one of banks or governments, keeping it non legal in many states like China, but being very popular becoming the first one and also the most famous of the cryptocurrencies, as explained in [19].
Despite its born in 2008, data were available on Yahoo Finance since 2019, so in this thesis it was considered the behaviour of Bitcoin from 18th September 2019 to to 1st September 2021.

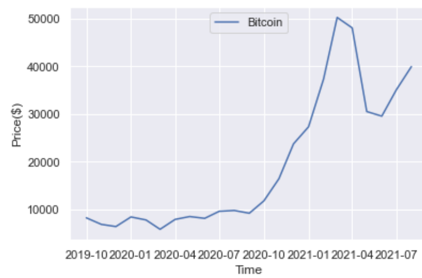

Figure 2.62: Bitcoin behaviour from 2019 to 2021.

As we can see, bitcoin is very volatile, passing to a price of 10k dollars in 2019 to a maximum of 50k dollars in April 2021, to a local minimum of 30k dollars in June 2021, with dropping of over the 30% in 24 hours breaking more than 2000$. There are many causes for this great uncertainty of bitcoin price like its limited supply, because there is a complex procedure called bitcoin mining to create new tokens, and this quantity entered into the market is halving each 4 years. In fact currently there are about 18 millions of bitcoin in circulation, but the maximum threshold will be 21 million in 2140, making bitcoin a very good investment to contrast inflation and very apt to variations.

Another very important cause of such a volatility is the lack of a central bank, as explained in [20], because its behaviour is so much affected by overall news. In 11^{th} May 2021, for example, Elon Musk announced on Twitter that Tesla would have not accepted anymore bitcoin payments to purchase electric cars. In a day bitcoin has lost 144 billions of dollars of capitalization according to [21] falling of 13% in 2 hours, passing from the exchange value of 54500 $ to 46980$.

From a preliminary analysis of the compounded returns of the data, the following table was obtained:

	τ	N° of observations	Mean	Std	Skewness	Excess kurtosis	Entropy	Hurst
0	1	709	0.521054	10.278073	-1.796037	24.163128	1.382873	0.378651
1	5	102	0.730159	5.245185	-0.918976	1.974489	0.911239	0.26425
2	20	22	0.905355	2.676114	-0.592423	-0.230043	0.515338	\

Figure 2.63: Features of Bitcoin compounded returns.

Starting in the begin of 2019, the number of observation is lower than the other time series. We can observe that the mean is not equal during the holding periods having an increase from Daily data to Monthly data while the standard deviation drops from a high value of 10.27 to a lower value of 2.67. The skewness remains negative diminishing in absolute value while the excess kurtosis drops from a very high value of 24.16 to a negative low value of -0.23, becoming a light platykurtic distribution having then fewer extreme positive or negative events. The very high volatility of Bitcoin is also evident from the entropy, because that values are among the highest, second only to Tesla values, meaning that this series are not very predictable. The Hurst coefficient, instead, is as usual under 0.5, but for the Monthly data is not shown because the number of data is very low, under 100, and this information would not be so meaningful.

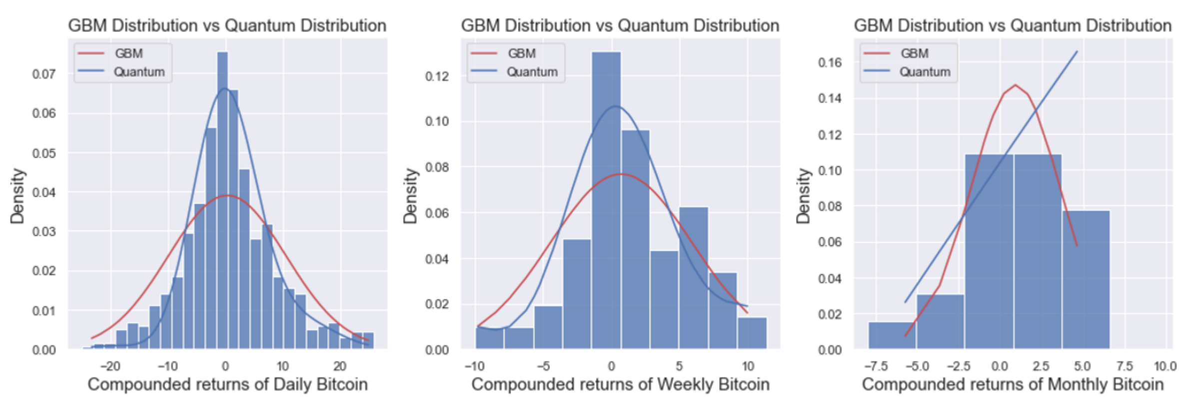

Figure 2.64: GBM vs Quantum Bitcoin.

We can observe how good is the Quantum model to describe Bitcoin PDF's for Daily and Weekly data, in particular regarding data with the lowest holding period, despite the failure of intercepting the kurtosis. For Monthly data, instead, the Quantum model isn't meaningful due to the little number of data while the GBM model manages to fit better the distribution.

Running the Quantum Genetic Algorithm for **75 generations** and the GBM Genetic Algorithm for **40 generations**, we have obtained the following fitness results:

Model Fitness	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM Fitness	3.74	3.07	2.40
Quantum Fitness	3.68	2.97	2.21

Figure 2.65: Fitness Bitcoin.

The Fitness values are higher with respect to the other time series, so the goodness of the algorithms is the lowest one, and they diminish over the holding periods as usual due to the drop of the number of data.

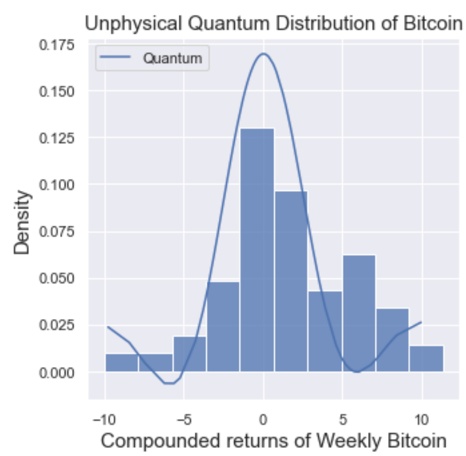

Figure 2.66: Not physical Bitcoin Weekly PDF.

We have also wanted to show an output of the algorithm which is not physical having values of the theoretical PDF less than 0. Despite the selection during the various generations, in fact, it may happen that some PDFs with negative values have the lowest fitness. This shows that fitness is not the only parameter to evaluate the goodness of a fitting, but also the theoretical PDF values must be greater than zero.

Models	Parameters	$\tau = 1$	$\tau = 5$	$\tau = 20$
GBM	μ	5.43	3.34	2.27
	σ^2	10.23	5.20	2.71
Quantum	C_0	4.81155078e-02	3.98145543e-02	8.29584812e-04
	C_1	5.52969396e-03	4.51868658e-03	4.84236638e-03
	C_2	-1.82056753e-05	6.34650861e-03	-7.69747791e-03
	C_3	1.03336203e-03	1.83276278e-05	-4.49326371e-02
	C_4	7.20499710e-04	1.56094478e-03	1.72551491e-03
	$m\omega$	1.028577340e-36	1.99775362e-36	8.017535728e-39

Figure 2.67: Estimated Parameters Bitcoin.

The first observation about the GBM parameters is that the standard deviations are the highest of all the time series for the 3 holding periods confirming the big volatility of Bitcoin returns. Analyzing the Quantum parameters, instead, we can notice that C_0 and C_1 are always positive with a minimum in Weekly data, C_2 is negative in Daily and Monthly having a positive maximum in Weekly, while C_4 remains positive increasing during the holding periods and C_3 turns negative from Weekly to Monthly. The parameter $m\omega$, instead, shows a weird behaviour being minimum in Monthly data against the theoretical foresight, confirming that the Quantum model for Monthly data is not so meaningful.

	P_0	P_1	P_2	P_3	P_4
$\tau = 1$	9.862969e-01	1.302685e-02	1.412053e-07	4.549278e-04	2.211592e-04
$\tau = 5$	9.616985e-01	1.238738e-02	2.443571e-02	2.037829e-07	1.478190e-03
$\tau = 20$	3.268933e-04	1.113781e-02	2.814371e-02	9.589773e-01	1.414236e-03

Figure 2.68: Probabilities Bitcoin.

The high volatility of Bitcoin prices is confirmed by its probabilities, because P_0 has the lowest values of all the time series, reaching the value of 0.96 in Weekly data and $3.26 * 10^{-4}$ for the Monthly data, even if they are not meaningful. The higher volatility probabilities, then, are larger with respect to the other time series, confirming the accuracy of the model noticing in particular the increase of the even probabilities which raise of 10^5 and of 10 respectively for P_2 and P_4.

2.13 Conclusion

The results have confirmed the goodness of the models and in particular the thesis described in paper [1] for almost all the historical series considered. The series analyzed have a very similar behaviour, with Quantum Model outperforming GBM Model almost always except in very rare case like Monthly FTSE MIB. In general the result is satisfactory with the Quantum Model that manages to recover the excess kurtosis of the data also in series like Weekly Amazon or Weekly Petroleum, but not being obtained in series with large standard deviation like Daily Petroleum and Daily Amazon. The Quantum Fitness is always smaller than the GBM one, but the Quantum Model almost never confirms the theory according to which the even probabilites contribute to the kurtosis while the odd probabilities to the skewness. The model, however, confirms in all series except that Petroleum and Bitcoin, that the product between the capitalization and the angular frequency increases. Moreover, perhaps the most important aspect, all the coefficients are of the same order of those obtained in [1] with even better fits, so we have confirmed the thesis that Financial Markets can be described by Complex Systems like GBM and Quantum Mechanics.

Chapter 3

Thesis codes

3.1 Entropy

```
def ApEn(U, m, r):
    U = np.array(U)
    N = U.shape[0]


    def phi(m):
        z = N - m + 1.0
        x = np.array([U[i:i+m] for i in range(int(z))])
        X = np.repeat(x[:, np.newaxis], 1, axis=2)
        C = np.sum(np.absolute(x - X).max(axis=2) ≤ r, axis=0)/z
        return np.log(C).sum()/z
     return abs(phi(m + 1) - phi(m))
```

The Entropy code has been taken from a github profile, linked at [23].

3.2 ECDF

```
def ecdf(data):
    xaxis = np.sort(data)
    yaxis = np.arange(1,len(data)+1)/len(data)
    return xaxis, yaxis
```

3.3 Quantum Algorithm

3.3.1 Quantum PDF

```
def psisquare(C,y):
    N = len(C)-1
    P = np.zeros((N,len(y)))
    y = np.sort(y)
    for i in range(N):
        P[i,:]=C[i]*np.exp(-C[-1]*y**2)*hermite(i)(np.sqrt(C[-1])*y)
    return P.sum(axis=0)/np.trapz(P.sum(axis=0),y)
```

3.3.2 Quantum Fitness of a vector:

```
def Fitness(x,C):
    M = len(x)
    return -(M)**(-1)*sum(np.log(psisquare(C,x))) - ApEn(ecdf(x)[1],
m=2, r=3)
```

3.3.3 Quantum Fitness of a matrix:

```
def fitnessvector(data,C):
    T = np.array([Fitness(data,C[i,:]) for i in range(C.shape[0])])
    for i in range(len(T)):
        if math.isnan(T[i])==True:
            T[i] = 9999999
    return T
```

3.3.4 Quantum Selection

```
def selection(fitness,C,number):
    C = initial population of the coefficients
    number = number of the rows of new vector
    fitness = fitnessvector(data,C)
    li = np.argsort(fitness)[:number]
    return np.array([C[b,:] for b in li])
```

3.3.5 Quantum Crossover

```
def crossover(C):
    offspring = np.zeros(shape=C.shape)
    crossoverpoints = np.empty(C.shape[0],dtype=int)
    for k in range(C.shape[0]):
        crossoverpoints[k] = random.randint(1,N)
        parent1idx = k%C.shape[0]
        parent2idx = (k+1)%C.shape[0]
        offspring[k, 0:crossoverpoints[k]] = C[parent1idx, 0:crossover-
points[k]]
        offspring[k, crossoverpoints[k]:N] = C[parent2idx, crossover-
points[k]:N]
        offspring[k,N] = C[:,-1][parent2idx]
    return offspring
```

3.3.6 Quantum Genetic Algorithm

As explained in the theoretical part, the mutation part of the algorithm has been computed at each generation.

Inputs of the equation.

data

Number of the weights we are looking to optimize.

numweights = N + 1

Defining the population size.

numpop = 1000
popsize = np.array([numpop,numweights])

Initializing populations

trypopulation = np.zeros((numpop,numweights))

C_0 coefficients.

trypopulation[:,0] = np.random.uniform(low=0, high=0.1, size = numpop)

C_1,C_N coefficients.

trypopulation[:,1:N] = np.random.uniform(low=-0.1, high=0.1, size=(numpop,N-1))

Oscillator masses m.

trypopulation[:,-1] = np.random.uniform(low=0, high=0.3, size = numpop)

Number of the generations decided.
numgenerations = 75
Number of the parents selected.
numparents = 500

Coefficientmatrix = pd.DataFrame(columns=['Generation','C_0', 'C_1', 'C_2','C_3','C_4','m','QuantumFitness']).
entropy = ApEn(ecdf(data)[1], m=2, r=3)
for generation in range(numgenerations):
Measuring the fitness of each chromosome in the population.
fitness = fitnessvector(data,trypopulation)
Selecting the best parents in the population.
parents = selection(fitness,trypopulation,numparents)
Generating next generation using crossover.
cross = crossover(parents)
Adding some variations to the offspring using mutation.
row = cross.shape[0]
column = cross.shape[1]
mutationmatrix = np.random.rand(row,column)*np.array([[random.randint(-1,1) for j in range(column)] for i in range(row)])
mutationmatrix+= np.ones((row,column))
mutant = mutationmatrix*cross
Creating the new population based on the parents and offspring.
trypopulation[0:parents.shape[0], :] = parents
trypopulation[parents.shape[0]:, :] = mutant
np.random.shuffle(trypopulation)
The best result in the current iteration.
gen = np.insert(trypopulation[np.where(fitness==np.min(fitness))], 0,int(generation + 1))
gen = np.insert(gen,len(gen),np.min(fitness)-entropy)
Coefficientmatrix.loc[generation] = gen

3.4 GBM Algorithm

3.4.1 GBM PDF

```
def GBM(D,x,t):
    mu = (D[0]-D[1]/2)*t
    sigma = D[1]*t
    return norm(mu, sigma).pdf(x)
```

3.4.2 GBM Fitness of a vector:

```
def GBMFitness(x,D,t):
    M = len(x)
    return-(M)**(-1)*sum(np.log(GBM(D,x,t))) - ApEn(ecdf(x)[1],
m=2, r=3)
```

3.4.3 GBM Fitness of a matrix:

```
def GBMfitnessvector(x,D,t):
    T = np.array([GBMFitness(x,D[i,:],t) for i in range(D.shape[0])])
    for i in range(len(T)):
        if math.isnan(T[i])==True:
            T[i] = 9999999
    return T
```

3.4.4 GBM Selection

```
def GBMselection(fitness,D,number):
    C = initial population of the coefficients
    number = number of the rows of new vector
    fitness = fitnessvector(data,C)
    li = np.argsort(fitness)[:number]
    return np.array([D[b,:] for b in li])
```

3.4.5 GBM Crossover

```
def GBMcrossover(D):
    offspring = np.zeros(shape=D.shape)
    randommu = random.randint(1,D.shape[0])
    randomsigma = random.randint(1,D.shape[0])
    for k in range(D.shape[0]):
        shiftmu = (k+randommu)%D.shape[0]
        shiftsigma = (k+randomsigma)%D.shape[0]
        offspring[k,0] = D[shiftmu,0]
        offspring[k,1] = D[shiftsigma,1]
    return offspring
```

3.4.6 GBM Genetic Algorithm

As in Quantum Model in the theoretical part, the mutation part of the algorithm has been computed at each generation.

Inputs of the equation.

data

Number of the weights we are looking to optimize.

numweights = 2

Defining the population size.

numpop = 1000

popsize = np.array([numpop,numweights])

Initializing populations

Dpopulation = np.zeros((numpop,numweights))

μ coefficients.

Dpopulation[:,0] = np.random.uniform(low=0.5, high=1.5, size = numpop)

σ^2 coefficients.

Dpopulation[:,1] = np.random.uniform(low=3, high=4, size=numpop)

Number of the generations decided.

numgenerations = 40

Number of the parents selected.

numparents = 500

```
CoefficientmatrixGBM = pd.DataFrame(columns=[’Generation’,’GBMmu’,
’GBMsigma’,’GBMFitness’])
entropy = ApEn(ecdf(data)[1], m=2, r=3)
for generation in range(numgenerations):
```

Measuring the fitness of each chromosome in the population.

```
GBMfitness = GBMfitnessvector(data,Dpopulation,t=1)
```

Selecting the best parents in the population.

```
GBMparents = GBMselection(GBMfitness,Dpopulation,GBMnumparents)
```

Generating next generation using crossover.

```
GBMcross = GBMcrossover(parents)
```

Adding some variations to the offspring using mutation.

```
GBMrow = GBMcross.shape[0]
GBMcolumn = GBMcross.shape[1]
GBMmutationmatrix = np.random.rand(GBMrow,GBMcolumn)*np.array([[random.r
1,1) for j in range(GBMcolumn)] for i in range(GBMrow)])
GBMmutationmatrix+= np.ones((GBMrow,GBMcolumn))
GBMmutant = GBMmutationmatrix*GBMcross
```

Creating the new population based on the parents and offspring.

```
Dpopulation[0:GBMnumparents, :] = GBMparents
Dpopulation[GBMnumparents:, :] = GBMmutant
np.random.shuffle(Dpopulation)
```

The best result in the current iteration.

```
GBMgen = np.insert(Dpopulation[np.where(GBMfitness==np.min(GBMfitness))]
GBMgen = np.insert(GBMgen,len(GBMgen),np.min(GBMfitness)-
entropy)
CoefficientmatrixGBM.loc[generation] = GBMgen
```

3.5 PDF Plots

```
plt.figure(figsize=(5,5))
plt.title(’GBM Distribution vs Quantum Distribution’, fontsize=’15’)
plt.xlabel(’Compounded returns of Data’, fontsize=’15’)
plt.ylabel(’Density’, fontsize=’15’)
GBMdatad = np.array(CoefficientmatrixGBM)[-1,1:-1]
Quantumdatad = np.array(Coefficientmatrix)[-1,1:-1]
datarg=15 (depending on data if Daily, Weekly or Monthly)
```

```
datat = np.sort([data[i] for i in range(len(data)) if -datarg¡=data[i]¡=datarg])
width = stats.freedmanbinwidth((data), returnbins=False)
histplot = sns.histplot(data,stat=”density”, binwidth=width,binrange=(-
datarg,datarg))
plt.plot(datat,GBM(GBMdatad,datat,1),color = ’r’,label = ’GBM’)
plt.plot(datat,psisquare(Quantumdatad,datat),color = ’b’,label =
’Quantum’)
plt.legend(loc=”upper left”)
```

Bibliography

[1] K. Ahnyz, M.Y. Choi, B. Daiy, S. Sohnz and B. Yangz, (2017) Modeling stock return distributions with a quantum harmonic oscillator, 9.

[2] Engle, R. and Patton, A., (2018) EPL (Europhysics Letters), Volume 120, Number 3 What good is a volatility model, (2001) Quantitative Finance, 237-245.

[3] https://www.borsaitaliana.it/notizie/sotto-la-lente/nasdaq.htm

[4] https://www.nasdaq.com/

[5] Luca Dall'Asta, Lecture Notes on Statistical Field Theory, (2020), 143

[6] https://en.wikipedia.org/wiki/Geometric Brownian motion

[7] Alfredo Braunstein, (2020), Algorithms for Optimization, Inference and Learning, 95

[8] Steve Pincus, (1995),Chaos: An Interdisciplinary Journal of Nonlinear Science, Approximate entropy (ApEn) as a complexity measure

[9] Elisa Ficarra, (2020), Slides for the course of Inference in Biological Systems

[10] http://faculty.washington.edu/ivezic/Teaching/Astr323/lec10.pdf

[11] https://www.investopedia.com/ask/answers/08/find-stocks-in-sp500.asp

[12] https://www.investopedia.com/terms/d/dowjoneseurostoxx50.asp

[13] https://investopedia.com/articles/investing/082714/what-dow-means-and-why-we-calculate-it-way-we-do.asp

[14] https://www.apesodoro.com/oro-quotazione/chi-e-cosa-determina-la-quotazione-dell-oro

[15] https://www.investopedia.com/terms/c/comex.asp

[16] https://www.plus500.it/Instruments/CL/What-Influences-the-Price-of-Oil 1

[17] https://www.ig.com/au/trading-strategies/amazon-stock-split-history–what-you-need-to-know-190725

[18] https://ir.tesla.com/press-release/tesla-announces-five-one-stock-split

[19] https://www.cmcmarkets.com/it-it/formazione-forex/cosa-e-il-forex

[20] https://www.investopedia.com/terms/b/bitcoin.asp

[21] https://www.cnbc.com/2021/05/19/why-is-bitcoin-so-volatile.html

[22] https://www.wired.it/economia/finanza/2021/05/13/musk-tesla-bitcoin-cripto/

[23] https://gist.github.com/DustinAlandzes/a835909ffd15b9927820d175a48dee41

[24] Eugene Stanley, Rosario Mantegna, (2000), Cambridge University Press, An Introduction to Econophysics: Correlations and Complexity in Finance

[25] https://www.econopoly.ilsole24ore.com/2020/01/11/mercati-finanziari-fisica/

[26] CFA Institute Journal Review,(November 2003), Volume 33, Issue 4, The Efficient Market Hypothesis and Its Critics (Summary)

[27] http://dspace.unive.it/bitstream/handle/10579/1672/818364-100978.pdf?sequence=2, 27 - 30.

[28] https://www.rivaluta.it/serie-inflazione-media.asp

[29] Mandelbrot, B.B. (1972) Statistical Methodology for Nonperiodic Cycles: From the Covariance to R/S Analysis. Annals of Economic and Social Measurement, 1, 259-290.

[30] Jianbo GaoJing Hu, Financial Crisis, Omori's Law, and Negative Entropy Flow, Abstract

[31] https://galton.uchicago.edu/ lalley/Courses/313/BrownianMotionCurrent.pdf

[32] http://www.sfu.ca/ kkasa/Dixit/Chpt3.pdf

[33] Raymond S.T.Lee, Quantum Finance: Intelligent Forecast and Trading Systems 1st ed. 2020 Edition, Chapter 1, 4.

[34] https://www.investopedia.com/investing/momentum-and-relative-strength-index/

[35] C Zhang, Lu Huang, A quantum model for the stock market, Science Direct, 2

[36] Ataullah, A., Davidson, I. and Tippett, M., A wave function for stock market returns. Physica A: Statistical Mechanics and its Applications, 2009, 388(4), 455-461.

List of Figures

1.1 Example of Bollinger Bands taken from investopedia.com . . . 16

2.1 NASDAQ COMPOSITE behaviour from 1996 to 2021. 29
2.2 Features of NASDAQ COMPOSITE compounded returns. . . 29
2.3 GBM vs Quantum NASDAQ COMPOSITE. 31
2.4 Fitness NASDAQ COMPOSITE. 31
2.5 Estimated Parameters NASDAQ COMPOSITE. 32
2.6 Probabilities NASDAQ COMPOSITE. 32
2.7 FTSE MIB behaviour from 1996 to 2021. 33
2.8 Features of FTSE MIB compounded returns. 34
2.9 GBM vs Quantum FTSE MIB. 35
2.10 Fitness FTSE MIB. 35
2.11 Estimated Parameters FTSE MIB. 36
2.12 Probabilities FTSE MIB. 36
2.13 S&P500 behaviour from 1996 to 2021. 37
2.14 Features of S&P500 compounded returns. 38
2.15 GBM vs Quantum S&P500. 38
2.16 Fitness S&P500. 39
2.17 Estimated Parameters S&P500. 39
2.18 Probabilities S&P500. 40
2.19 EURO STOXX 50 behaviour from 2007 to 2021. 41
2.20 Features of EURO STOXX 50 compounded returns 41
2.21 GBM vs Quantum STOXX 50. 42
2.22 Fitness EURO STOXX 50. 43
2.23 Estimated Parameters EURO STOXX 50. 43
2.24 Probabilities EURO STOXX 50. 43
2.25 DOW JONES behaviour from 1996 to 2021. 46
2.26 Features of DOW JONES compounded returns. 46
2.27 GBM vs Quantum DOW JONES. 47

2.28 Fitness DOW JONES. 47
2.29 Estimated Parameters DOW JONES. 48
2.30 Probabilities DOW JONES. 48
2.31 Gold behaviour from 2000 to 2021. 49
2.32 Features of Gold compounded returns. 50
2.33 GBM vs Quantum Gold. 50
2.34 Fitness Gold. 51
2.35 Estimated Parameters Gold. 51
2.36 Probabilities Gold. 52
2.37 Petroleum behaviour from 2000 to 2021. 53
2.38 Features of Petroleum compounded returns. 53
2.39 GBM vs Quantum Petroleum. 54
2.40 Fitness Petroleum. 55
2.41 Estimated Parameters Petroleum. 55
2.42 Probabilities Petroleum. 55
2.43 Amazon behaviour from 2000 to 2021. 56
2.44 Amazon stock splits, image taken from [17]. 56
2.45 Features of Amazon compounded returns. 57
2.46 GBM vs Quantum Amazon. 58
2.47 Fitness Amazon. 58
2.48 Estimated Parameters Amazon. 58
2.49 Probabilities Amazon. 59
2.50 Tesla behaviour from 2000 to 2021. 59
2.51 Features of Tesla compounded returns. 60
2.52 GBM vs Quantum Tesla. 60
2.53 Fitness Tesla. 61
2.54 Estimated Parameters Tesla. 61
2.55 Probabilities Tesla. 61
2.56 Forex EUR - USD behaviour from 2003 to 2021. 63
2.57 Features of Forex EUR - USD compounded returns. 63
2.58 GBM vs Quantum Forex EUR - USD. 64
2.59 Fitness Forex EUR - USD. 64
2.60 Estimated Parameters Forex EUR - USD. 64
2.61 Probabilities Forex EUR - USD 65
2.62 Bitcoin behaviour from 2019 to 2021. 66
2.63 Features of Bitcoin compounded returns. 67
2.64 GBM vs Quantum Bitcoin. 67
2.65 Fitness Bitcoin. 68

2.66 Not physical Bitcoin Weekly PDF. 68
2.67 Estimated Parameters Bitcoin. 69
2.68 Probabilities Bitcoin. 69