

Article

A Novel Methodology to Calculate the Probability of Volatility Clusters in Financial Series: An Application to Cryptocurrency Markets

Venelina Nikolova ^{1,†}, Juan E. Trinidad Segovia ^{1,*,†} , Manuel Fernández-Martínez ^{2,†} 
and Miguel Angel Sánchez-Granero ^{3,†} 

¹ Department of Accounting and Finance, Faculty of Economics and Business, Universidad de Almería, 04120 Almería, Spain; vdn088@inlumine.ual.es

² University Centre of Defence at the Spanish Air Force Academy, MDE-UPCT, 30720 Santiago de la Ribera, Región de Murcia, Spain; manuel.fernandez-martinez@ud.upct.es

³ Department of Mathematics, Faculty of Science, Universidad de Almería, 04120 Almería, Spain; misanche@ual.es

* Correspondence: jetrini@ual.es; Tel.: +34-950-015-817

† These authors contributed equally to this work.

Received: 25 May 2020; Accepted: 22 July 2020; Published: 24 July 2020



Abstract: One of the main characteristics of cryptocurrencies is the high volatility of their exchange rates. In a previous work, the authors found that a process with volatility clusters displays a volatility series with a high Hurst exponent. In this paper, we provide a novel methodology to calculate the probability of volatility clusters with a special emphasis on cryptocurrencies. With this aim, we calculate the Hurst exponent of a volatility series by means of the FD4 approach. An explicit criterion to computationally determine whether there exist volatility clusters of a fixed size is described. We found that the probabilities of volatility clusters of an index (S&P500) and a stock (Apple) showed a similar profile, whereas the probability of volatility clusters of a forex pair (Euro/USD) became quite lower. On the other hand, a similar profile appeared for Bitcoin/USD, Ethereum/USD, and Ripple/USD cryptocurrencies, with the probabilities of volatility clusters of all such cryptocurrencies being much greater than the ones of the three traditional assets. Our results suggest that the volatility in cryptocurrencies changes faster than in traditional assets, and much faster than in forex pairs.

Keywords: volatility cluster; Hurst exponent; FD4 approach; volatility series; probability of volatility cluster; S&P500; Bitcoin; Ethereum; Ripple

1. Introduction

It is easy to observe that large fluctuations in stock market prices are followed by large ones, whereas small fluctuations in prices are more likely to be followed by small ones. This property is known as volatility clustering. Recent works, such as [1,2], have shown that while large fluctuations tend to be more clustered than small ones, large losses tend to lump together more severely than large gains. The financial literature is interested in modeling volatility clustering since the latter is considered as a key indicator of market risk. In fact, the trading volume of some assets, such as derivatives, increases over time, making volatility their most important pricing factor.

It is worth mentioning that both, high and low volatilities, seem to be a relevant factor for stock market crises according to Danielsson et al. [3]. They also found that the relation between unexpected volatility and the incidence of crises became stronger in the last few decades. In the same

line, Valentine et al. [4] showed that market instability is not only the result of large volatility, but also of small volatility.

The classical approach for volatility clusters lies in nonlinear models, based on heteroskedastic conditional variance. They include ARCH [5], GARCH [6–8], IGARCH [9], and FIGARCH [10,11] models.

On the other hand, agent-based models allow reproducing and explaining some stylized facts of financial markets [12]. Interestingly, several works have recently been appeared in the literature analyzing a complete order book by real-time simulation [13–15]. Regarding the volatility clustering, it is worth mentioning that Lux et al. [16] highlighted that volatility is explained by market instability. Later, Raberto et al. [17] introduced an agent-based artificial market whose heterogeneous agents exchange only one asset, which exhibits some key stylized facts of financial markets. They found that the volatility clustering effect is sensitive to the model size, i.e., when the number of operators increases, the volatility clustering effect tends to disappear. That result is in accordance with the concept of market efficiency.

Krawiecki et al. [18] introduced a microscopic model consisting of many agents with random interactions. Then the volatility clustering phenomenon appears as a result of attractor bubbling. Szabolcs and Farmer [19] empirically developed a behavioral model for order placement to study endogenous dynamics of liquidity and price formation in the order book. They were able to describe volatility through the order flow parameters.

Alfarano et al. [20] contributed a simple model of an agent-based artificial market with volatility clustering generated due to interaction between traders. Similar conclusions were obtained by Cont [21], Chen [22], He et al. [23], and Schmitt and Westerhoff [24].

Other findings on the possible causes of volatility clusters are summarized below. Cont [21] showed that volatility is explained by agent behavior; Chen [22] stated that return volatility correlations arise from asymmetric trading and investors' herding behavior; He et al. [23] concluded that trade between fundamental noise and noise traders causes the volatility clustering; and Schmitt and Westerhoff [24] highlighted that volatility clustering arises due to the herding behavior of speculators.

Chen et al. [25] proposed an agent-based model with multi-level herding to reproduce the volatilities of New York and Hong Kong stocks. Shi et al. [26] explained volatility clustering through a model of security price dynamics with two kind of participants, namely speculators and fundamental investors. They considered that information arrives randomly to the market, which leads to changes in the viewpoint of the market participants according to a certain ratio. Verma et al. [27] used a factor model to analyze how market volatility could be explained by assets' volatility.

An interesting contribution was made by Barde in [28], where the author compared the performance of this kind of model with respect to the ARCH/GARCH models. In fact, the author remarked that the performance of three kinds of agent-based models for financial markets is better in key events. Population switching was found also as a crucial factor to explain volatility clustering and fat tails.

On the other hand, the concept of a volatility series was introduced in [2] to study the volatility clusters in the S&P500 series. Moreover, it was shown that the higher the self-similarity exponent of the volatility series of the S&P500, the more frequent the volatility changes and, therefore, the more likely that the volatility clusters appear. In the current article, we provide a novel methodology to calculate the probability of volatility clusters of a given size in a series with special emphasis on cryptocurrencies.

Since the introduction of Bitcoin in 2008, the cryptocurrency market has experienced a constant growth, just like the use of crypto assets as an investment or medium of exchange day to day. As of June 2020, there are 5624 cryptocurrencies, and their market capitalization exceeds 255 billion USD according to the website CoinMarketCap [29]. However, one of the main characteristics of cryptocurrencies is the high volatility of their exchange rates, and consequently, the high risk associated with their use.

Lately, Bitcoin has received more and more attention by researchers. Compared to the traditional financial markets, the cryptocurrency market is very young, and because of this, there are relatively few

research works on their characteristics, and all of them quite recent. Some of the authors analyzed the Bitcoin market efficiency by applying different approaches, including the Hurst exponent (cf. [30] for a detailed review), whereas others investigated its volatility using other methods. For instance, Letra [31] used a GARCH model for Bitcoin daily data; Bouoiyour and Selmi [32] carried out many extensions of GARCH models to estimate Bitcoin price dynamics; Bouri, Azzi, and Dyhberg [33] analyzed the relation between volatility changes and price returns of Bitcoin based on an asymmetric GARCH model; Balcilar et al. [34] analyzed the relation between the trading volume of Bitcoin and its returns and volatility by employing, in contrast, a non-parametric causality in quantiles test; and Baur et al. [35] studied the statistical properties of Bitcoin and its relations with traditional asset classes.

Meanwhile, in 2017, Bariviera et al. [36] used the Hurst exponent to compare Bitcoin dynamics with standard currencies' dynamics and detected evidence of persistent volatility and long memory, facts that justify the GARCH-type models' application to Bitcoin prices. Shortly after that, Phillip et al. [37] provided evidence of slight leverage effects, volatility clustering, and varied kurtosis. Furthermore, Zhang et al. [38] analyzed the first eight cryptocurrencies that represent almost 70% of cryptocurrency market capitalization and pointed out that the returns of cryptocurrencies exhibit leverage effects and strong volatility clustering.

Later, in 2019, Kancs et al. [39], based on the GARCH model, estimated factors that affect Bitcoin price. For it, they used hourly data for the period between 2013 and 2018. After plotting the data graphically, they suggested that periods of high volatility follow periods of high volatility, and periods of low volatility follow periods of low volatility, so in the series, large returns follow large returns and small returns small returns. All these facts indicate evidence of volatility clustering and, therefore, that the residue is conditionally heteroscedastic.

The structure of this article is as follows. Firstly, Section 2 contains some mathematical basic concepts on measure theory and probability (Section 2.1), the FD4 approach (Section 2.2), and the volatility series (Section 2.3). The core of the current paper is provided in Section 3, where we explain in detail how to calculate the probability of volatility clusters of a given size. A study of volatility clusters in several cryptocurrencies, as well as in traditional exchanges is carried out in Section 4. Finally, Section 5 summarizes the main conclusions of this work.

2. Methods

This section contains some mathematical tools of both measure and probability theories (cf. Section 2.1) that allow us to mathematically describe the FD4 algorithm applied in this article (cf. Section 2.2) to calculate the self-similarity index of time series. On the other hand, the concept of a volatility series is addressed in Section 2.3.

2.1. Random Functions, Their Increments, and Self-Affinity Properties

Let $t \geq 0$ denote time and (X, \mathcal{A}, P) be a probability space. We shall understand that $\mathbf{X} = \{X_t \equiv X(t, \omega) : t \geq 0\}$ is a random process (also a random function) from $[0, \infty) \times \Omega$ to \mathbb{R} , if X_t is a random variable for all $t \geq 0$ and $\omega \in \Omega$, where Ω denotes a sample space. As such, we may think of \mathbf{X} as defining a sample function $t \mapsto X_t$ for all $\omega \in \Omega$. Hence, the points in Ω do parameterize the functions $\mathbf{X} : [0, \infty) \rightarrow \mathbb{R}$ with P being a measure of probability in the class of such functions.

Let X_t and Y_t be two random functions. The notation $X_t \sim Y_t$ means that the finite joint distribution functions of such random functions are the same. A random process $\mathbf{X} = \{X_t : t \geq 0\}$ is said to be self-similar if there exists a parameter $H > 0$ such that the following power law holds:

$$X_{at} \sim a^H X_t \quad (1)$$

for each $a > 0$ and $t \geq 0$. If Equation (1) is fulfilled, then H is named the self-similarity exponent (also index) of the process \mathbf{X} . On the other hand, the increments of a random function X_t are said to be stationary as long as $X_{a+t} - X_a \sim X_t - X_0$ for all $t \geq 0$ and $a > 0$. We shall understand that the

increments of a random function are self-affine of the parameter $H \geq 0$ if the next power law stands for all $h > 0$ and $t_0 \geq 0$:

$$X_{t_0+\tau} - X_{t_0} \sim h^{-H} (X_{t_0+h\tau} - X_{t_0}).$$

Let X_t be a random function with self-affine increments of the parameter H . Then, the following T^H -law holds:

$$\mathcal{M}_T \sim T^H \mathcal{M}_1,$$

where its (T -period) cumulative range is defined as:

$$\mathcal{M}_{t,T} := \sup \{X(s, \omega) - X(t, \omega) : s \in [t, t + T]\} - \inf \{X(s, \omega) - X(t, \omega) : s \in [t, t + T]\},$$

and $\mathcal{M}_T := \mathcal{M}_{0,T}$ (cf. Corollary 3.6 in [40]).

2.2. The FD4 Approach

The FD4 approach was first contributed in [41] to deal with calculations concerning the self-similarity exponent of random processes. It was proven that the FD4 generalizes the GM2 procedure (cf. [42,43]), as well as the fractal dimension algorithms (cf. [44]) to calculate the Hurst exponent of any process with stationary and self-affine increments (cf. Theorem 3.1 in [41]). Moreover, the accuracy of such an algorithm was analyzed for samples of (fractional) Brownian motions and Lévy stable processes with lengths ranging from 2^5 to 2^{10} points (cf. Section 5 in [41]).

Next, we mathematically show how that parameter could be calculated by the FD4 procedure. First of all, let $\mathbf{X} = \{X_t : t \geq 0\}$ be a random process with stationary increments. Let $q > 0$, and assume that for each $X_t \in \mathbf{X}$, there exists $m_q(X_t) := E [|X_t|^q]$, its (absolute) q -order moment. Suppose, in addition, that there exists a parameter $H > 0$ for which the next relation, which involves (τ -period) cumulative ranges of \mathbf{X} , holds:

$$\mathcal{M}_\tau \sim \tau^H \mathcal{M}_1. \tag{2}$$

Recall that this power law stands for the class of (H)-self-similar processes with self-affine increments (of parameter H ; see Section 2.1), which, roughly speaking, is equivalent to the class of processes with stationary increments (cf. Lemma 1.7.2 in [45]). Let us discretize the period by $\tau_n = 2^{-n} : n \in \mathbb{N}$ and take q -powers on both sides of Equation (2). Thus, we have:

$$\mathcal{M}_{\tau_n}^q \sim \tau_n^{qH} \mathcal{M}_1^q \text{ for all } n \in \mathbb{N}. \tag{3}$$

Clearly, the expression in Equation (3) could be rewritten in the following terms:

$$X_n^q \sim \tau_n^{qH} X_0^q = 2^{-nqH} X_0^q \tag{4}$$

where, for short, the notation $X_n := \mathcal{M}_{\tau_n} = \mathcal{M}_{2^{-n}}$ is used for all $n \in \mathbb{N}$. Since the two random variables in Equation (4) are equally distributed, their means must be the same, i.e.,

$$m_q(X_n) = E [X_n^q] = 2^{-nqH} E [X_0^q] = 2^{-nqH} m_q(X_0). \tag{5}$$

Taking (2-base) logarithms on both sides of Equation (5), the parameter H could be obtained by carrying out a linear regression of:

$$H = \frac{1}{nq} \log_2 \frac{m_q(X_0)}{m_q(X_n)}. \tag{6}$$

vs. q . Alternatively, observe that the expression in Equation (4) also provides a relation between cumulative ranges of consecutive periods of \mathbf{X} , i.e.,

$$X_n^q \sim 2^{qH} X_{n+1}^q. \tag{7}$$

Since the random variables on each side of Equation (7) have the same (joint) distribution function, their means must be equal, namely,

$$m_q(X_n) = E[X_n^q] = 2^{qH} E[X_{n+1}^q] = 2^{qH} m_q(X_{n+1}) \text{ for all } n \in \mathbb{N}, \tag{8}$$

which provides a strong connection between consecutive moments of order q of \mathbf{X} . If (two-base) logarithms are taken on both sides of Equation (8), a linear regression of the expression appearing in Equation (9) vs. q allows calculating the self-similarity exponent of \mathbf{X} (whenever self-similar patterns do exist for such a process):

$$H = \frac{1}{q} \log_2 \frac{m_q(X_n)}{m_q(X_{n+1})}. \tag{9}$$

Hence, the FD algorithm is defined as the approach whose running is based on the expressions appearing in either Equation (5) or Equation (8). The main restriction underlying the FD algorithm consists of the assumption regarding the existence of the q -order moments of the random process \mathbf{X} . At first glance, any non-zero value could be assigned to q to calculate the self-similarity exponent (provided that the existence of that sample moment could be guaranteed). In the case of Lévy stable motions, for example, given q_0 , it may occur that $m_q(X_n)$ does not exist for any $q > q_0$. As such, we shall select $q = 0.01$ to calculate the self-similarity index of a time series by the FD algorithm, thus leading to the so-called FD4 algorithm. Equivalently, the FD4 approach denotes the FD algorithm for $q = 0.01$. In this paper, the self-similarity exponent of a series by the FD4 approach is calculated according to the expression in Equation (6). Indeed, since it is equivalent to:

$$\log_2 m_q(X_n) = \log_2 m_q(X_0) - nqH,$$

the Hurst exponent of the series is obtained as the slope of a linear regression, which compares $\log_2 m_q(X_n)$ with respect to n . In addition, notice that a regression coefficient close to one means that the expression in Equation (5) is fulfilled. As such, the calculation of $m_q(X_n)$ becomes necessary to deal with the procedure described above, and for each n , it depends on a given sample of the random variable $X_n \in \mathbf{X}$. For computational purposes, the length of any sample of X_n is chosen to be equal to 2^n . Accordingly, the greater n , the more accurate the value of $m_q(X_n)$ is. Next, we explain how to calculate $m_q(X_n)$. Let a log-price series be given, and divide it into 2^n non-overlapping blocks, $\mathcal{B}_i : i = 1, \dots, 2^n$. The length of each block is $k := 2^{-n} \cdot \text{length}(\text{series})$, so for each $i = 1, \dots, 2^n$, we can write $\mathcal{B}_i = \{B_1, \dots, B_k\}$. Then:

1. Determine the range of each block \mathcal{B}_i , i.e., calculate $R_i = \max\{B_j : j = 1, \dots, k\} - \min\{B_j : j = 1, \dots, k\}$ for each $i = 1, \dots, 2^n$.
2. The (q -order) sample moment is given by $m_q(X_n) = 2^{-n} \sum_{i=1}^{2^n} R_i^q$.

According to the step (1), both the minimum and the maximum values of each period are required to calculate each range R_i . In this way, notice that such values are usually known for each trading period in the context of financial series. It is also worth noting that when n takes the value $\log_2(\text{length}(\text{series}))$, then each block only consists of a single element. In this case, though, each range R_i can be still computed.

2.3. The Volatility Series

The concept of a volatility series was first contributed in Section 2.2 of [2] as an alternative to classical (G)ARCH models with the aim to detect volatility clusters in series of asset returns from the S&P500 index. It was found, interestingly, that whether clusters of high (resp., low) volatility appear in the series, then the self-similarity exponent of the associated volatility series increases (resp., decreases).

Let r_n denote the log-return series of a (index/stock) series. In financial series, the autocorrelation function of the r_n 's is almost null, though the $|r_n|$ series is not. The associated volatility series is defined

as $s_n = |r_n| + s_{n-1} - m$, where $|\cdot|$ refers to the absolute value function, m is a constant, and $s_0 = 0$. For practical purposes, we set $m = \text{mean } |r_n|$.

Next, we explain how the Hurst exponent of the volatility series, s_n , could provide a useful tool to detect volatility clusters in a series of asset returns. Firstly, assume that the volatility of the series is constant. Then, the values of the associated volatility series would be similar to those from a sample of a Brownian motion. Hence, the self-similarity exponent of that volatility series would become close to 0.5. On the contrary, suppose that there exist some clusters of high (resp., low) volatility in the series. Thus, the graph of its associated volatility series becomes smoother, as illustrated in Figure 1, which also depicts the concept of a volatility series. Hence, almost all the values of the volatility series are greater (resp., lower) than the mean of the series. Accordingly, the volatility series turns out to be increasing (resp., decreasing), so its self-similarity exponent also increases (resp., decreases).

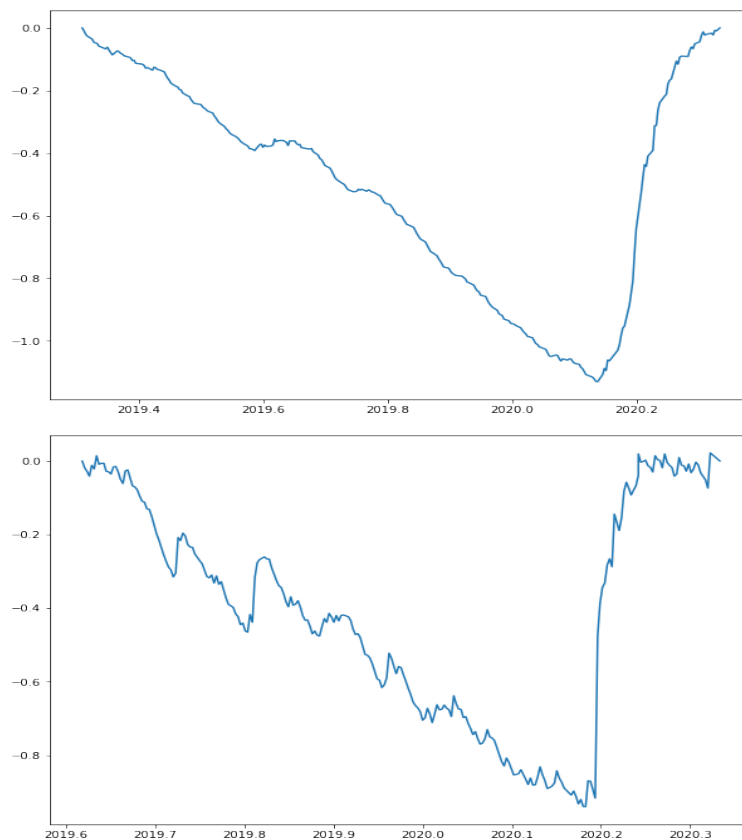


Figure 1. The picture at the top depicts the volatility series of the S&P500 index in the period ranging from March 2019 to March 2020. A self-similarity exponent equal to 0.94 is obtained by the FD4 approach. The graph below illustrates the volatility series of the Bitcoin/USD index in a similar period (both series contain 250 data (one year of trading), but recall that the Bitcoin/USD currency remains active also on weekends). In that case, a self-similarity exponent equal to 0.65 is obtained.

Following the above, the Hurst exponent of the volatility series of an index or asset provides a novel approach to explore the presence of volatility clusters in series of asset returns.

3. Calculating the Probability of Volatility Clusters of a Given Size

In this section, we explore how to estimate the probability of the existence of volatility clusters for blocks of a given size. Equivalently, we shall address the next question: What is the probability that a volatility cluster appears in a period of a given size? Next, we show that the Hurst exponent of a volatility series (see Sections 2.2 and 2.3) for blocks of that size plays a key role.

We know that the Hurst exponent of the volatility series is high when there are volatility clusters in the series [2]. However, how high should it be?

To deal with this, we shall assume that the series of (log-)returns follows a Gaussian distribution. However, it cannot be an i.i.d. process since the standard deviation of the Gaussian distribution is allowed to change. This hypothesis is more general than an ARCH or GARCH model, for example. Since we are interested in the real possibility that the volatility changes and, in fact, there exist volatility clusters, a static fixed distribution cannot be assumed. In this way, it is worth noting that the return distribution of these kinds of processes (generated from Gaussian distributions with different standard deviations) is not Gaussian, and it is flexible enough to allow very different kinds of distributions.

As such, let us assume that the series of the log-returns, r_n , follows a normal distribution, $N(0, \sigma(n))$, where its standard deviation varies over time via the function $\sigma(n)$. In fact, some classical models such as ARCH, GARCH, etc., stand as particular cases of that model. As such, we shall analyze the existence of volatility clusters in the following terms. We consider that there exist volatility clusters as long as there are, at least, both, a period of high volatility and a period of low volatility. Figure 2 illustrates that condition. Indeed, two broad periods could be observed concerning the volatility series of the S&P500 index. The first one has a low volatility (and hence, a decreasing volatility series) and the second one a high volatility (and hence, an increasing volatility series). In this case, the effect of the higher volatility (due to the COVID-19 crisis) is evident, thus being confirmed by a very high Hurst exponent of the corresponding volatility series (equal to 0.94).

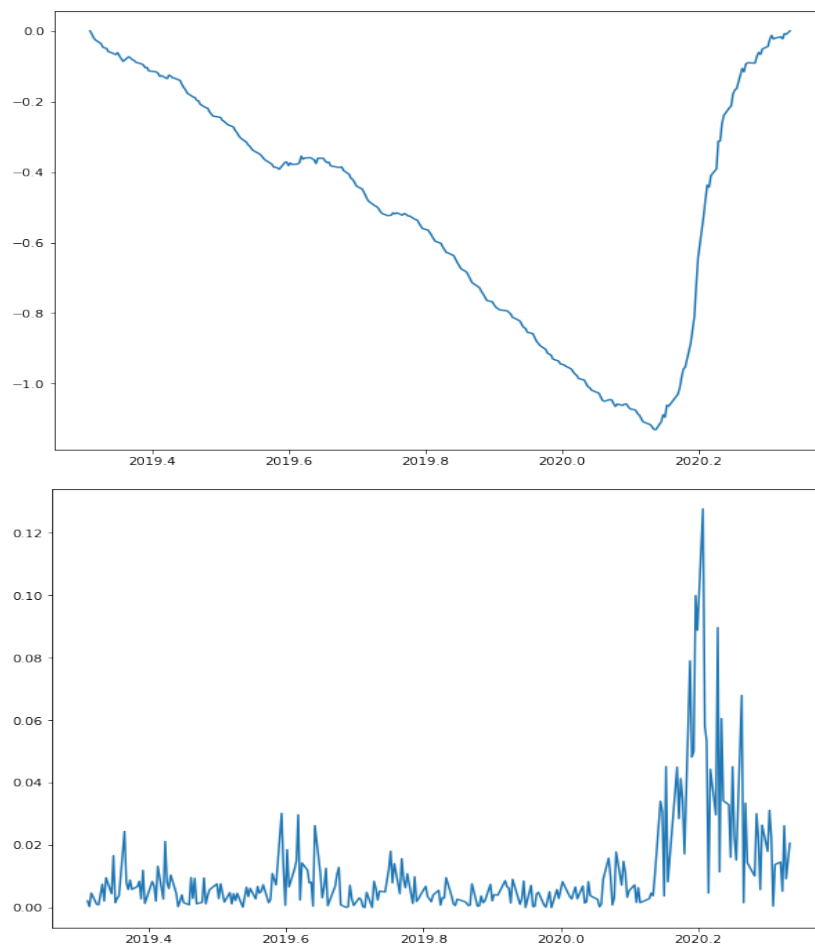


Figure 2. The graph at the top depicts the volatility series of the S&P500 index in the period ranging from March 2019 to March 2020. On the other hand, the chart at the bottom shows the series of absolute values of the log-returns of the S&P500 in the same period. That period, the self-similarity index of the volatility series of the S&P500, was found to be equal to 0.94 by the FD4 algorithm.

On the other hand, Figure 3 depicts the volatility series of the S&P500 index in the period ranging from January 2017 to January 2018. A self-similarity index equal to 0.55 was found by the FD4 algorithm. In this case, though, it is not so clear that there are volatility clusters, which is in accordance with the low Hurst exponent of that volatility series.

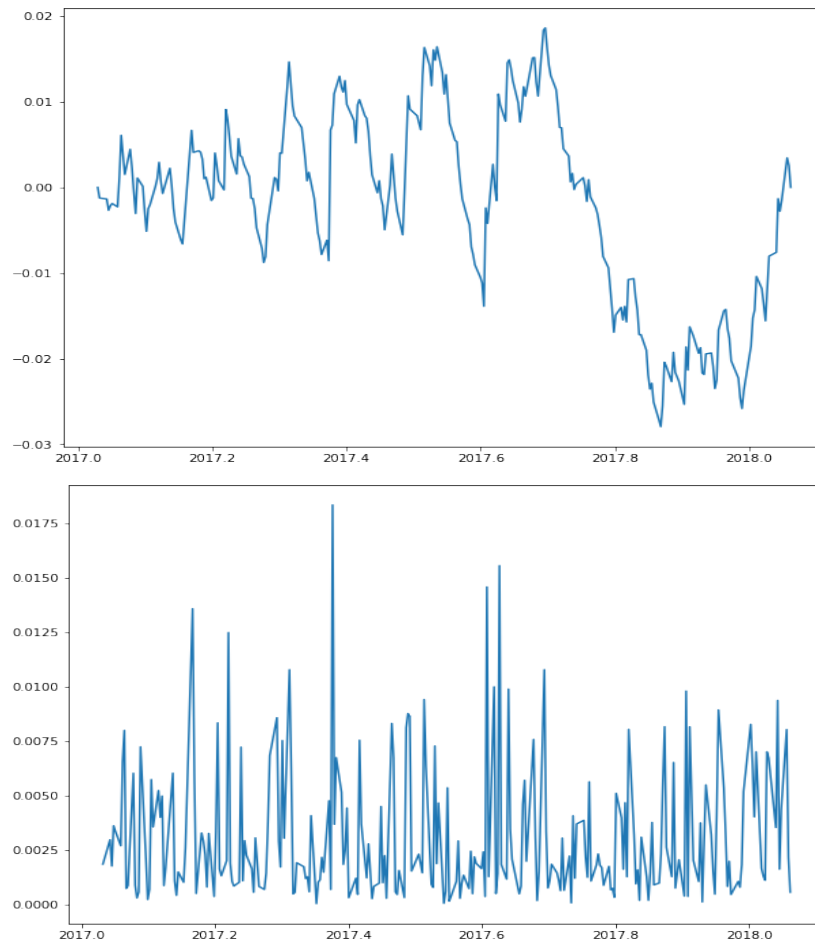


Figure 3. The plot at the top illustrates the volatility series of the S&P500 index in the period ranging from January 2017 to January 2018, whereas the graph at the bottom depicts the series of absolute values of the log-returns of the S&P500 index in the same period. In this case, the self-similarity exponent of the volatility series was found to be equal to 0.55 by the FD4 approach.

As such, the Hurst exponent of the volatility series of a Brownian motion will be considered as a benchmark in order to decide whether there are volatility clusters in the series. More precisely, first, by Monte Carlo simulation, a collection of Brownian motions was generated. For each Brownian motion, the Hurst exponents (by FD4 approach) of their corresponding volatility series were calculated. Hence, we denote by $H_{lim}(n)$ the value that becomes greater than 90% of those Hurst exponents. Observe that $H_{lim}(n)$ depends on n , the length of the Brownian motion sample. In fact, for a short series, the accuracy of the FD4 algorithm to calculate the Hurst exponent is lower. Accordingly, the value of $H_{lim}(n)$ will be higher for a lower value of n . Figure 4 illustrates (for the 90th percentile) how the benchmark given by $H_{lim}(n)$ becomes lower as the length of the Brownian motion series increases.

Therefore, we will use the following criteria. We say that there are volatility clusters in the series provided that the Hurst exponent of the corresponding volatility series is greater than H_{lim} . Then, we will measure the probability of volatility clusters for subseries of a given length as the ratio between the number of subseries with volatility clusters to the total amount of subseries of the given length.

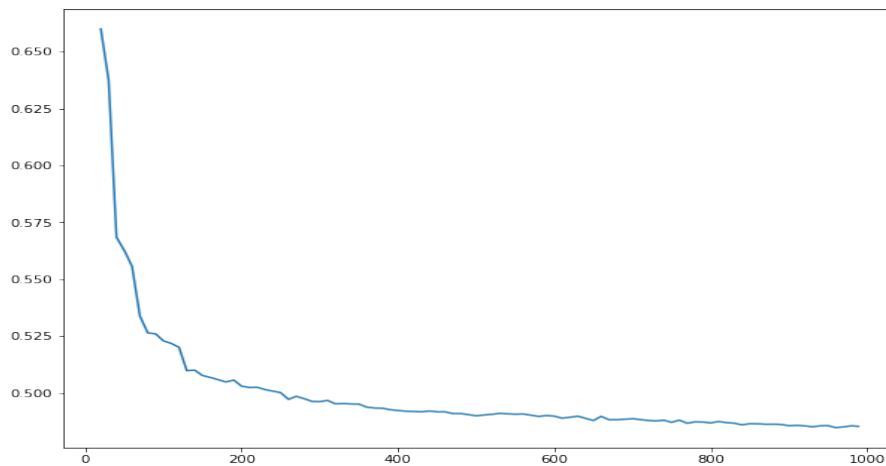


Figure 4. The 90th percentile, $H_{lim}(n)$, of the Hurst exponents of the volatility series of Brownian motions for several values of the length of the series, n .

In order to check that measure of the probability of volatility clusters, we will test it by artificial processes with volatility clusters of a fixed length (equal to 200 data). A sample from that process is generated as follows. For the first 200 data, generate a sample from a normal distribution $N(0, 0.01)$; for the next 200 data, generate a sample from a normal distribution $N(0, 0.03)$; for the next 200 data, generate a sample from a normal distribution $N(0, 0.01)$, and so on. It is worth pointing out that a mixture of (samples from) normal distributions with distinct standard deviations can lead to (a sample from) a heavy-tailed distribution. Following that example, Figure 5 depicts the distribution of that artificial process with volatility clusters compared to the one from a Gaussian distribution and also to the S&P500 return distribution (rescaled). It is clear that the process is far from Gaussian even in that easy example.

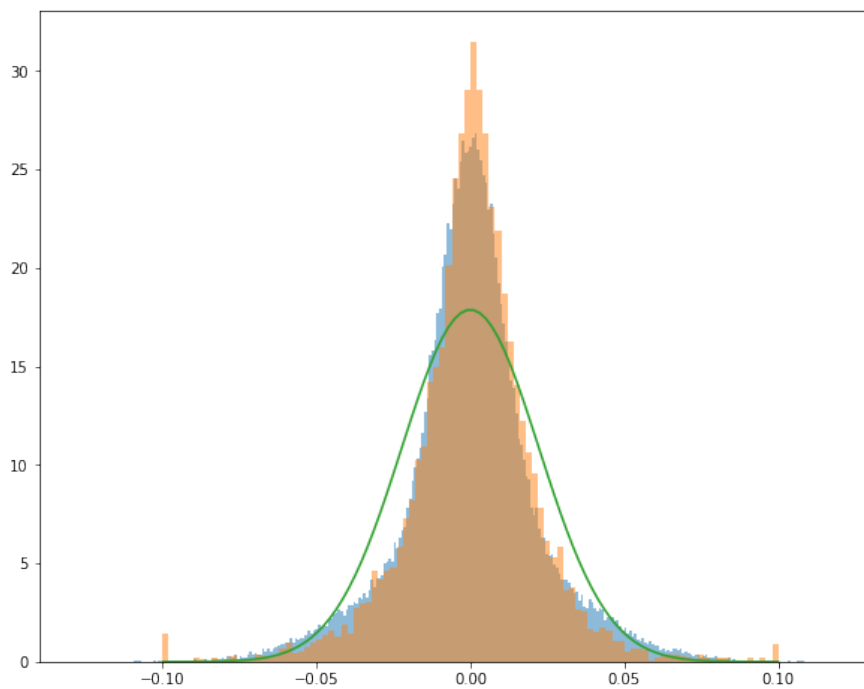


Figure 5. Histogram or density function of: (orange) return distribution of the S&P500 (rescaled and clipped to the interval $[-0.1, 0.1]$); (blue) process with volatility clusters of a fixed length (200 data); (green) normal distribution.

For that process, consider one random block of length 50. It may happen that such a block fully lies in a 200 block of fixed volatility. In this case, there will be no volatility clusters. However, if the first 20 data lie in a block of volatility equal to 0.01, with the remaining 30 data lying in a block of volatility equal to 0.03, then such a block will have volatility clusters. On the other hand, it is clear that if we have one block of length 50 with the first 49 data lying in a block of volatility equal to 0.01, whereas the remaining one datum lies in a block of 0.03 volatility, we cannot say that there are volatility clusters in such a block. Therefore, we shall consider that there are volatility clusters if there are at least 10 data in blocks with distinct volatilities. In other words, we shall assume that we cannot detect clusters with less than 10 data.

On the other hand, note that we are using a confidence level of 90%, and hence, if we get a probability of volatility clusters of, say, $x\%$, that means that there are no volatility clusters regarding the $(100 - x)\%$ of the blocks of the given size. However, for that confidence level of 90%, we are missing 10% of that $(100 - x)\%$, and hence, we will have the following theoretical estimates.

- Theoretical probability of volatility clusters considering clusters of at least 10 data: $(x - 20)/200$.
- Theoretical probability of volatility clusters considering clusters of at least 10 data detected at a confidence level of 90%: $(x - 20)/200 + (1 - (x - 20)/200) \cdot 0.1$.

Figure 6 graphically shows that the proposed model for estimating the probability of volatility clusters could provide a fair approximation to the actual probability of volatility clusters for such an artificial process.

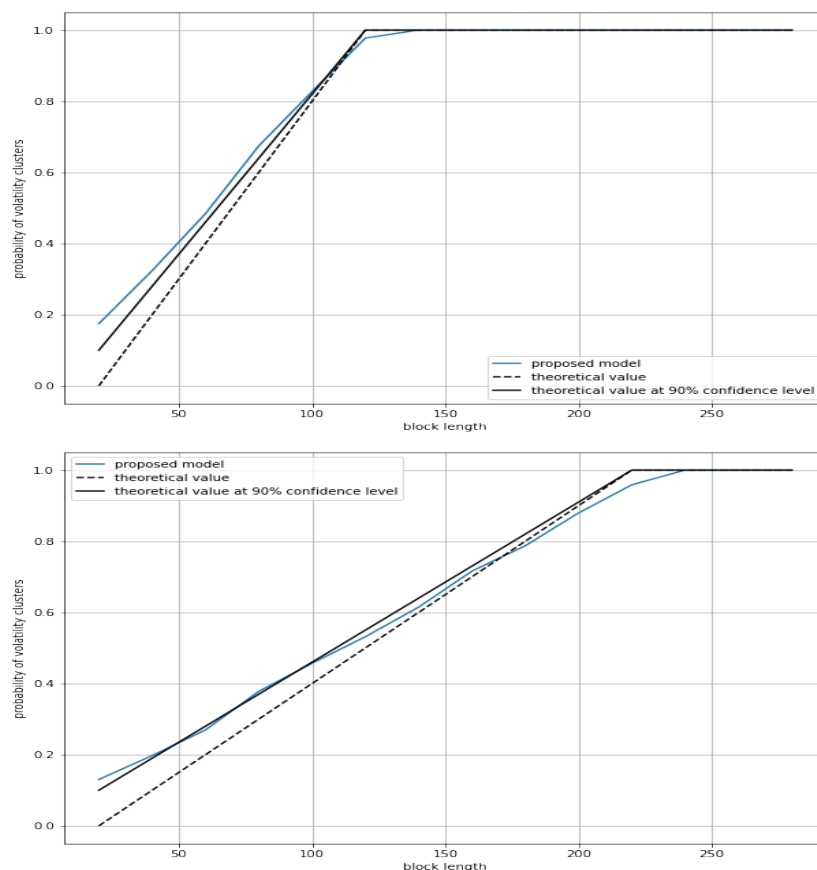


Figure 6. Cont.

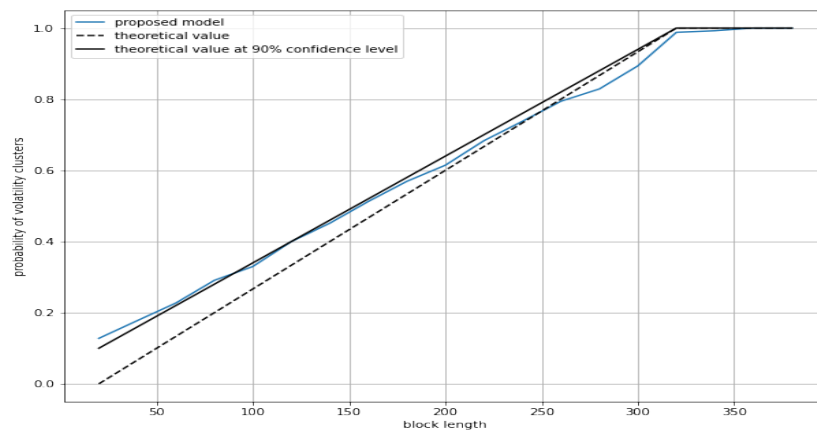


Figure 6. Probability of the existence of volatility clusters for an artificial process with volatility clusters of a fixed length (equal to 100, 200, and 300, from top to bottom).

4. Volatility Clusters in Cryptocurrencies

One of the main characteristics of cryptocurrencies is the high volatility of their exchange rates, and consequently, the high risk associated with their use.

In this section, the methodology provided in Section 3 to calculate the probability of volatility clusters is applied to different financial assets, with a special interest in cryptocurrency markets.

First, Figure 7 shows a similar profile in regard to the probabilities of volatility clusters of an index (S&P500) and a stock (Apple). On the other hand, the probability of volatility clusters of the Euro/USD exchange results in being quite lower.

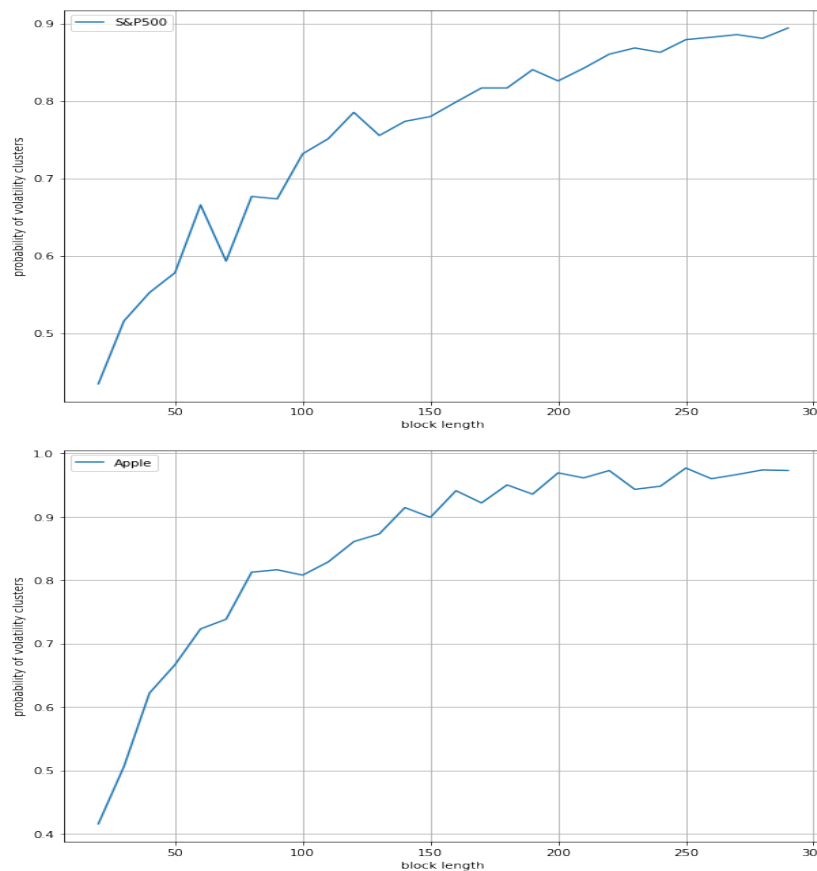


Figure 7. Cont.

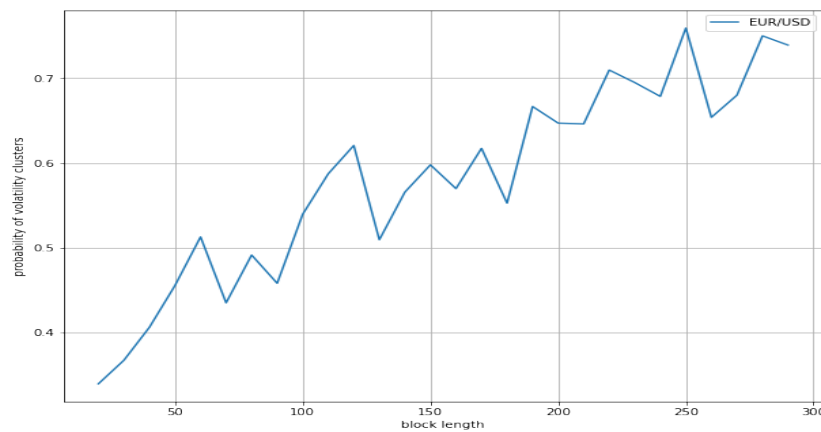


Figure 7. Probabilities of volatility clusters of the next assets. From top to bottom, an index (S&P500), a stock (Apple), and a forex pair (Euro/USD).

On the other hand, Figure 8 depicts the probability of volatility clusters of the three main cryptocurrencies, namely Bitcoin/USD, Ethereum/USD, and Ripple/USD. A similar profile appears for all such cryptocurrencies with the probabilities of their volatility clusters much greater than the ones for the three asset classes displayed in Figure 7.

These results suggest that the volatility in cryptocurrencies changes faster than in traditional assets, and much faster than in forex pairs.

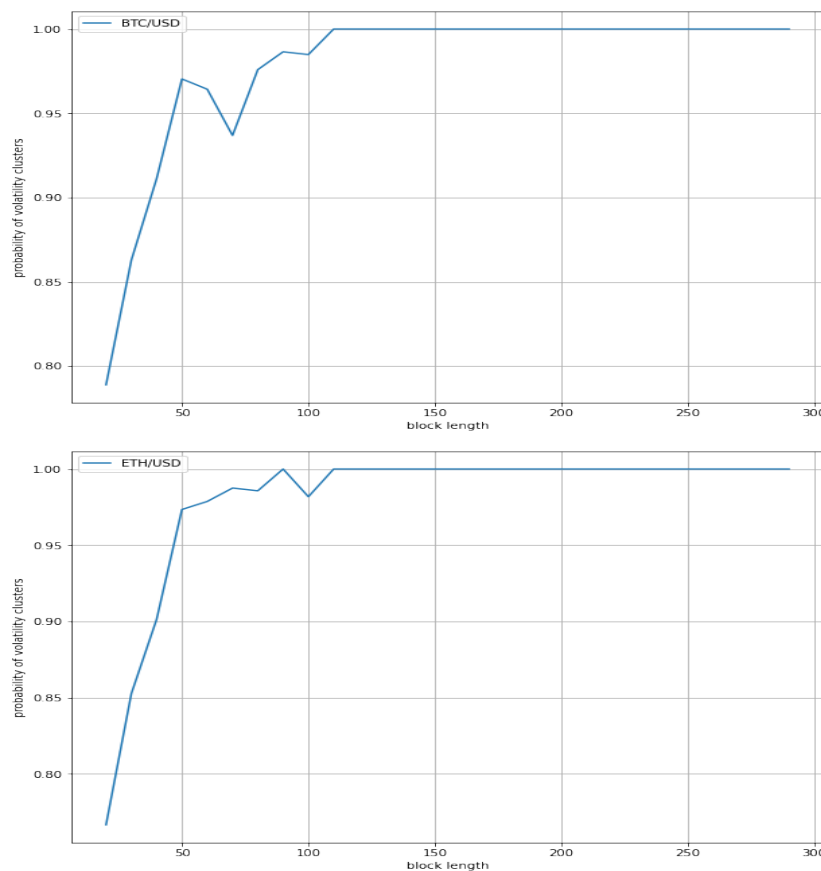


Figure 8. Cont.

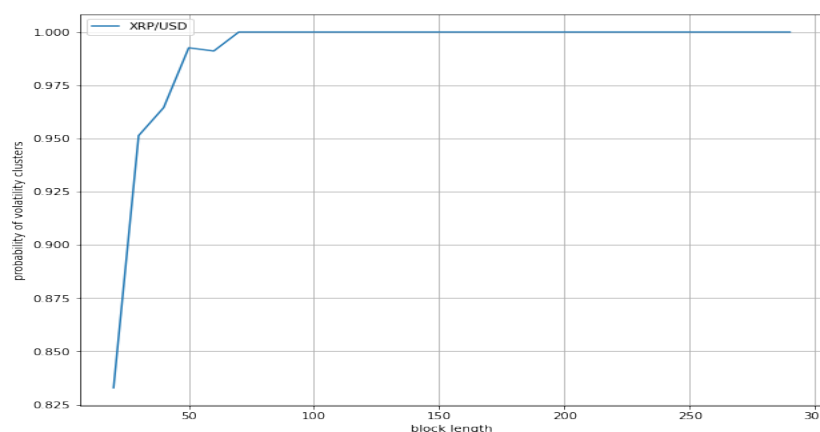


Figure 8. Probabilities of volatility clusters of the following cryptocurrencies. From top to bottom, Bitcoin/USD, Ethereum/USD, and Ripple/USD.

5. Conclusions

One of the main characteristics of cryptocurrencies is the high volatility of their exchange rates. In a previous work, the authors found that a process with volatility clusters displays a volatility series with a high Hurst exponent [2].

In this paper, we provide a novel methodology to calculate the probability of the volatility clusters of a series using the Hurst exponent of its associated volatility series. Our approach, which generalizes the (G)ARCH models, was tested for a class of processes artificially generated with volatility clusters of a given size. In addition, we provided an explicit criterion to computationally determine whether there exist volatility clusters of a fixed size. Interestingly, this criterion is in line with the behavior of the Hurst exponent (calculated by the FD4 approach) of the corresponding volatility series.

We found that the probabilities of volatility clusters of an index (S&P500) and a stock (Apple) show a similar profile, whereas the probability of volatility clusters of a forex pair (Euro/USD) results in being quite lower. On the other hand, a similar profile appears for Bitcoin/USD, Ethereum/USD, and Ripple/USD cryptocurrencies, with the probabilities of volatility clusters of all such cryptocurrencies being much greater than the ones of the three traditional assets. Accordingly, our results suggest that the volatility in cryptocurrencies changes faster than in traditional assets, and much faster than in forex pairs.

Author Contributions: Conceptualization, V.N., J.E.T.S., M.F.-M., and M.A.S.-G.; methodology, V.N., J.E.T.S., M.F.-M., and M.A.S.-G.; validation, V.N., J.E.T.S., M.F.-M., and M.A.S.-G.; formal analysis, V.N., J.E.T.S., M.F.-M., and M.A.S.-G.; writing—original draft preparation, V.N., J.E.T.S., M.F.-M., and M.A.S.-G.; writing—review and editing, V.N., J.E.T.S., M.F.-M., and M.A.S.-G.; These authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: Both J.E. Trinidad Segovia and M.A. Sánchez-Granero are partially supported by Ministerio de Ciencia, Innovación y Universidades, Spain, and FEDER, Spain, Grant PGC2018-101555-B-I00, and UAL/CECEU/FEDER, Spain, Grant UAL18-FQM-B038-A. Further, M.A. Sánchez-Granero acknowledges the support of CDTIME. M. Fernández-Martínez is partially supported by Ministerio de Ciencia, Innovación y Universidades, Spain, and FEDER, Spain, Grant PGC2018-097198-B-I00, and Fundación Séneca of Región de Murcia (Murcia, Spain), Grant 20783/PI/18.

Acknowledgments: The authors would also like to express their gratitude to the anonymous reviewers whose suggestions, comments, and remarks allowed them to enhance the quality of this paper.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; nor in the decision to publish the results.

References

1. Tseng, J.-J.; Li, S.-P. Asset returns and volatility clustering in financial time series. *Phys. A Stat. Mech. Appl.* **2011**, *390*, 1300–1314. [[CrossRef](#)]
2. Trinidad Segovia, J.E.; Fernández-Martínez, M.; Sánchez-Granero, M.A. A novel approach to detect volatility clusters in financial time series. *Phys. A Stat. Mech. Appl.* **2019**, *535*, 122452. [[CrossRef](#)]
3. Danielsson, J.; Valenzuela, M.; Zer, I. Learning from History: Volatility and Financial Crises. *Rev. Financ.* **2018**, *21*, 2774–2805. [[CrossRef](#)]
4. Valenti, D.; Fazio, G.; Spagnolo, B. The stabilizing effect of volatility in financial markets. *Phys. Rev. E* **2018**, *97*, 062307. [[CrossRef](#)]
5. Engle, R.F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* **1982**, *50*, 987–1007. [[CrossRef](#)]
6. Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *J. Econom.* **1986**, *31*, 307–327. [[CrossRef](#)]
7. Taylor, S.J. *Modelling Financial Time Series*; John Wiley & Sons, Ltd.: Chichester, UK, 1986.
8. Kim, Y.S.; Rachev, S.T.; Bianchi, M.L.; Fabozzi, F.J. Financial market models with Lévy processes and time-varying volatility. *J. Bank. Financ.* **2008**, *32*, 1363–1378. [[CrossRef](#)]
9. Engle, R.F.; Bollerslev, T. Modelling the persistence of conditional variances. *Econom. Rev.* **1986**, *5*, 1–50. [[CrossRef](#)]
10. Baillie, R.T.; Bollerslev, T.; Mikkelsen, H.O. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *J. Econom.* **1996**, *74*, 3–30. [[CrossRef](#)]
11. Bentes, S.R. Long memory volatility of gold price returns: How strong is the evidence from distinct economic cycles? *Phys. A Stat. Mech. Appl.* **2016**, *443*, 149–160. [[CrossRef](#)]
12. Patterson, G.; Sornette, D.; Parisi, D. Properties of balanced flows with bottlenecks: Common stylized facts in finance and vibration-driven vehicles. *Phys. Rev. E* **2020**, *101*, 042302. [[CrossRef](#)]
13. Biondo, A.E. Order book microstructure and policies for financial stability. *Stud. Econ. Financ.* **2018**, *35*, 196–218. [[CrossRef](#)]
14. Biondo, A.E. Order book modeling and financial stability. *J. Econ. Interact. Coord.* **2019**, *14*, 469–489. [[CrossRef](#)]
15. Sueshige, T.; Sornette, D.; Takayasu, H.; Takayasu, M. Classification of position management strategies at the order-book level and their influences on future market-price formation. *PLoS ONE* **2019**, *14*, e0220645. [[CrossRef](#)] [[PubMed](#)]
16. Lux, T.; Marchesi, M. Volatility clustering in financial markets: A microsimulation of interacting agents. *Int. J. Theor. Appl. Financ.* **2000**, *3*, 675–702. [[CrossRef](#)]
17. Raberto, M.; Cincotti, S.; Focardi, S.M.; Marches, M. Agent-based simulation of a financial market. *Phys. A Stat. Mech. Appl.* **2001**, *299*, 319–327. [[CrossRef](#)]
18. Krawiecki, A.; Holyst, J.A.; Helbing, D. Volatility Clustering and Scaling for Financial Time Series due to Attractor Bubbling. *Phys. Rev. Lett.* **2002**, *89*, 158701. [[CrossRef](#)]
19. Szabolcs, M.; Farmer, J.D. An empirical behavioral model of liquidity and volatility. *J. Econ. Dyn. Control* **2008**, *32*, 200–234
20. Alfarano, S.; Lux, T.; Wagner, F. Estimation of Agent-Based Models: The Case of an Asymmetric Herding Model. *Comput. Econ.* **2005**, *26*, 19–49 [[CrossRef](#)]
21. Cont, R. Volatility Clustering in Financial Markets: Empirical Facts and Agent-Based Models. In *Long Memory in Economics*; Teysnière, G., Kirman, A.P., Eds.; Springer: Berlin/Heidelberg, Germany, 2007.
22. Chen, J.J.; Zheng, B.; Tan, L. Agent-Based Model with Asymmetric Trading and Herding for Complex Financial Systems. *PLoS ONE* **2013**, *8*, e79531. [[CrossRef](#)]
23. He, X.Z.; Li, K.; Wang, C. Volatility clustering: A nonlinear theoretical approach. *J. Econ. Behav. Organ.* **2016**, *130*, 274–297. [[CrossRef](#)]
24. Schmitt, N.; Westerhoff, F. Herding behavior and volatility clustering in financial markets. *Quant. Financ.* **2017**, *17*, 1187–1203. [[CrossRef](#)]

25. Chen, J.-J.; Tan, L.; Zheng, B. Agent-based model with multi-level herding for complex financial systems. *Sci. Rep.* **2015**, *5*, 8399. [[CrossRef](#)] [[PubMed](#)]
26. Shi, Y.; Luo, Q.; Li, H. An Agent-Based Model of a Pricing Process with Power Law, Volatility Clustering, and Jumps. *Complexity* **2019**, *2019*, 3429412. [[CrossRef](#)]
27. Verma, A.; Buonocore, R.J.; di Matteo, T. A cluster driven log-volatility factor model: A deepening on the source of the volatility clustering. *Quant. Financ.* **2018**, 1–16. [[CrossRef](#)]
28. Barde, S. Direct comparison of agent-based models of herding in financial markets. *J. Econ. Dyn. Control* **2016**, *73*, 329–353 [[CrossRef](#)]
29. CoinMarketCap, 2020. Available online: <https://coinmarketcap.com/all/views/all/> (accessed on 24 March 2020).
30. Dimitrova, V.; Fernández-Martínez, M.; Sánchez-Granero, M.A.; Trinidad Segovia, J.E. Some comments on Bitcoin market (in)efficiency. *PLoS ONE* **2019**, *14*, e0219243. [[CrossRef](#)]
31. Letra, I. What Drives Cryptocurrency Value? A Volatility and Predictability Analysis. 2016. Available online: <https://www.repository.utl.pt/handle/10400.5/12556> (accessed on 24 March 2020).
32. Bouoiyour, J.; Selmi, R. Bitcoin: A beginning of a new phase? *Econ. Bull.* **2016**, *36*, 1430–1440.
33. Bouri, E.; Azzi, G.; Dyhrberg, A.H. On the return-volatility relationship in the Bitcoin market around the price crash of 2013. *Economics* **2017**, *11*, 1–16.
34. Balcihar, M.; Bouri, E.; Gupta, R.; Roubaud, D. Can volume predict Bitcoin returns and volatility? A quantiles-based approach. *Econ. Model.* **2017**, *64*, 74–81. [[CrossRef](#)]
35. Baur, D.G.; Hong, K.; Lee, A.D. Bitcoin: Medium of exchange or speculative assets? *J. Int. Financ. Mark. Inst. Money* **2018**. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2561183 (accessed on 1 May 2020)
36. Bariviera, A.F.; Basgall, M.J.; Hasperué, W.; Naiouf, M. Some stylized facts of the Bitcoin market. *Physics A* **2017**, *484*, 82–90. [[CrossRef](#)]
37. Phillip, A.; Chan, J.S.K.; Peiris, S. A new look at cryptocurrencies. *Econ. Lett.* **2018**, *163*, 6–9. [[CrossRef](#)]
38. Zhang, W.; Wang, P.; Li, X.; Shen, D. Some stylized facts of the cryptocurrency market. *Appl. Econ.* **2018**, *50*, 5950–5965. [[CrossRef](#)]
39. Kancs, D.; Rajcaniova, M.; Ciaian, P. *The Price of Bitcoin: GARCH Evidence from High Frequency Data*; 29598 EN; Publications Office of the European Union: Luxembourg, 2019; ISBN 978-92-7998570-6, JRC115098. [[CrossRef](#)]
40. Mandelbrot, B.B. (Ed.) *Gaussian Self-Affinity and Fractals*; Springer-Verlag: New York, NY, USA, 2002.
41. Fernández-Martínez, M.; Sánchez-Granero, M.A.; Trinidad Segovia, J.E.; Román-Sánchez, I.M. An accurate algorithm to calculate the Hurst exponent of self-similar processes. *Phys. Lett. A* **2014**, *378*, 2355–2362. [[CrossRef](#)]
42. Sánchez-Granero, M.A.; Trinidad Segovia, J.E.; García Pérez, J. Some comments on Hurst exponent and the long memory processes on capital markets. *Phys. A Stat. Mech. Appl.* **2008**, *387*, 5543–5551. [[CrossRef](#)]
43. Trinidad Segovia, M.; Fernández-Martínez, J.E.; Sánchez-Granero, M.A. A note on geometric method-based procedures to calculate the Hurst exponent, *Phys. A Stat. Mech. Appl.* **2012**, *391*, 2209–2214. [[CrossRef](#)]
44. Sánchez-Granero, M.A.; Fernández-Martínez, M.; Trinidad Segovia, J.E. Introducing fractal dimension algorithms to calculate the Hurst exponent of financial time series. *Eur. Phys. J. B* **2012**, 85–86. [[CrossRef](#)]
45. Fernández-Martínez, M.; Guirao, J.L.G.; Sánchez-Granero, M.A.; Trinidad Segovia, J.E. *Fractal Dimension for Fractal Structures: With Applications to Finance*, 1st ed.; Springer Nature: Cham, Switzerland, 2019; pp. 1–204.

