MNRAS 000, 1–10 (2024) Preprint 12 May 2024 Compiled using MNRAS LATEX style file v3.0

Time series analysis of Crypto-Currencies

Aman Bhalla[1]★
[1] University of Southampton, University Rd, Southampton SO17 1BJ
Student ID: 32929749

12 May 2024

ABSTRACT
Crypto-currencies are an evolving form of decentralised currency with complex market dynamics. This paper investigates the growth and relation of 3 popular cryptocurrencies through the use of time-series analysis. The paper will begin with an overview of the fundamentals of cryptocurrencies, followed by a methodology covering correlation, smoothing techniques, lag and volatility analysis. Lag analysis reveals strong correlations in the movements of Bitcoin and Ethereum, with some variability due to developments of Ethereum. Cross-correlation with Cardano uncovers potential lags in the data, suggesting that Cardano may lead Bitcoin and Ethereum. Finally, the volatility analysis has important implications for investors, suggesting that Bitcoin and Ethereum remain more stable relative to Cardano.

1 INTRODUCTION

Cryptocurrencies are digital forms of currency that allow peer-to-peer payments between people anywhere around the world. They have been adopted by over 295 million people since the inception of the first cryptocurrency, Bitcoin, in January of 2009.

Cryptocurrencies are stored in digital wallets where payments are recorded in a ledger known as the blockchain. The blockchain is a localised ledger that stores all transactions in a peer-peer network. This removes the need for a central authority which clears payments, such as a bank. The term "cryptocurrency" is a result of the encryption method used to verify transactions. Data is encrypted using an encryption key where the data can only be decrypted and read by a recipient with the correct key. For example, Bitcoin data (of any length) is sent into an algorithm which returns a 64-digit hexadecimal number, storing the encrypted data. This aims to provide security for users, ensuring the encryption of data between wallets and public ledgers (Kaspersky 2022; PwC 2023; Seth 2024).

The increasing adoption of cryptocurrencies is a result of the many benefits it provides over traditional currency. Decentralization of currency aims to combat large monopolies as governments can no longer set the value of currency. Additionally, the safety provided by cryptography methods and transactional speeds are more sophisticated than traditional currency (Tambe and Jain 2024). This has resulted in significant growths in cryptocurrencies. For example, the total value of Bitcoin as of April 10, 2024 stood at $1.39 trillion with over 19.68 million bitcoins in circulation. This is in significant contrast to its price in 2023 at $472 billion with 19.35 million Bitcoins in circulation (Reiff 2024). However, cryptocurrencies come with some drawbacks. Newer blockchains may be prone to attacks, with a constant risk of 51%. Attackers can then pause transactions, change the spending amount or prevent transactions from occurring. cryptocurrencies are also incredibly volatile, where prices often fluctuate due to media attention, government regulations and investor fears (Tambe and Jain 2024).

As such, there are a variety of methods which can be used to predict and understand the prices of different coins. These methods include machine learning, fundamental analysis and time series analysis. Machine learning models can be trained on historical data and then used to predict future prices. Models can include linear regression, neural networks and random forests to name a few. Fundamental analysis involves a deep-analysis of all information about a certain cryptocurrency, with a mixture of quantitative and qualitative approaches. Ultimately, this method aims to determine the intrinsic price of a cryptocurrency using information such as the market cap, total value, road map, community size and rate of adoption (Coin Tree 2021).

Time Series Analysis is a technique which aims to understand the time series data produced by cryptocurrencies as they fluctuate relative to another currency. The method is able to incorporate trends, seasonal components (eg. variation during holiday seasons) and major events (eg. global pandemics) into its forecasting of a time series. It is also useful in investigating the relations between cryptocurrencies. In specific, how a currency may lag with another, and how similar the trends of a currency are.

Time series analysis is not specific to understanding financial data, but has many applications. For example, climate forecasting, retail forecasting and astrophysics. In astrophysics time series data is collected in the form of the light we receive from astronomical objects, such as stars, planets and galaxies. We are able to determine the distribution of this data in terms of the observing time. Ultimately, this data contains an element of randomness which may be a derivative of a greater phenomena. What may appear as "noise" in the data may have be a deterministic variability such as a rotation period or luminosity decay of a stellar body . By comparing multiple time series, astrophysicists can observe lags between time series. This occurs when one time series may lag behind the other, indicating that whatever produced the original time series may be having an effect on the second time series. As such, astronomers can extract physical properties from this data, such as the mass of stars, the period of a binary system etc (Vaughan 2015).

In this paper, I will be using time series analysis to understand fluctuations in cryptocurrencies relative to each other. I will look at the correlations, lag and volatility of 3 currencies: Bitcoin, Ethereum and Cardano. Bitcoin and Ethereum were introduced in 2009 and 2015 respectively, and have become well established cryptocurrencies since their inception. Bitcoin and Ethereum are the 2 biggest cryptocurrencies and have a combined market capitalization of more than 60% of the crypto market (Hooson 2022). Cardano was established in 2017 and aims to evolve into a decentralised platform, with a goal to be an alternative to Ethereum. The paper will begin with an outline of the methods used, including the calculation of correlation coefficients, different smoothing methods, a cross-correlation lag analysis and finally a volatility analysis. I will then present the results of the analysis, followed by a discussion on the interpretation and causes of my findings.

2 METHOD

The following section outlines the methodology used to apply time series analysis to cryptocurrencies. I will begin by outlining how the data is sourced, the method behind determining Pearson coefficients and an outline of cross-correlation analysis. Finally, I will explain the steps involved in a volatility analysis of each coin.

2.1 Importing Data

To begin the analysis, I will use hourly data for Bitcoin, Etherium and Cardano relative to Tether. Tether is a stable coin cryptocurrency. A stable coin is tied to the value of another currency. In this case, Tether is tied to the US dollar. The data I am using can be accessed from the site "cryptodatadownload" : https://www.cryptodatadownload.com/data/binance/. I have imported this data into the pandas data frame in python which gives me access to its respective analytical tools.

The hourly data contains the high, low, open and closing price of the currency. The high and low prices indicate the maximum and minimum prices the data reached within the hour. The open and closing price represents the price at the start of the hour and at the end of the hour respectively. I will use the open price of the data in my analysis. This is because the closing price is simply the opening price at the next hour, therefore this choice is identical for large datasets. I did not use high and low prices as they may be a result of random volatility - a factor I would like to minimise.

2.2 Calculating Correlation

A primary aim of this project is to understand how similar the cryptocurrencies are. To do this I will calculate the correlation between the time series, using the Pearson coefficient. The Pearson coefficient represents the relationship between two data sets.

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \tag{1}$$

The numerator defines the covariance of the two data sets x and y, where the covariance measures the degree at which the two variables spread with each other. The denominator is the product of the standard deviation of x and y, which measures the spread of each dataset about the mean. x_i and y_i define each value in the x and y dataset respectively. $\bar{x}$ and $\bar{y}$ are the mean of the x and y dataset respectively . A positive Pearson's Coefficient means that both datasets progress in the same direction, however a negative coefficient indicates an inverse correlation between both data sets. If the value is close to 1, the correlation is strong, however a value close to 0 represents a weak correlation. The significance of this correlation can be interpreted through the p value. The p value indicates the probability of observing a correlation coefficient assuming there is no correlation between the two variables. A low p value suggests the observation is unlikely to have occurred by chance, indicating a significant correlation. In this paper I will take an acceptable p value to be < 0.05 (Fernando 2024).

2.3 Smoothing the Time Series

An important technique in revealing the underlying progression in a time series dataset is smoothing. This is because smoothing is able to reduce the random variation in the time series data, making it easier to understand larger developments in the data. In this section I will provide a brief overview of a selection of smoothing methods. In the results section I will include a direct comparison between these methods.

2.3.1 Simple Moving Average

A naive approach would be to simply take the mean of all the data, however, if the data possess a trend we would like to forecast, this would eliminate any fluctuations in the data. A rolling average is able to take into account these fluctuations. The rolling average takes the mean of small successive windows of the data. The basic formula for the simple moving average (SMA) in a window is given below:

$$SMA = \frac{x_1 + x_2 + \ldots + x_N}{N} \tag{2}$$

Here x_n defines the price at point x. N is the number of points in a window (Hayes 2023b).

2.3.2 Savitzky - Golay Filter

A Savitzky-Golay filter is a smoothing technique which uses polynomials to smooth windows of data. For each point in the data, a polynomial can be obtained from 2n+1 surrounding points. The polynomial is evaluated at the centre of this neighbourhood and can take on an order defined by the user. In this paper, I will use the Savitzky-Golay filter to smooth the data before calculating lags between the crypto-coins. Savitzky-Golay filters reduce noise in the data, providing a cleaner signal to the cross-correlation lag function (Wronski 2021).

2.4 Calculating Lags

Another aim of this project is to understand if there are any lags between the cryptocoins. More explicitly, I will investigate if developments in the price of one cryptocoin are a result of similar developments in another. This can be achieved by completing a cross-correlation analysis which tracks the movements of two time series relative to each other. As such, cross-correlations can reveal lags between cryptocurrencies (Hayes 2023a).

2.4.1 Normalising Data

Before calculating the lags, I normalised the data. Normalisation aims to eliminate the effects of data scaling. This will set the prices of each currency to be between 0 and 1, allowing for appropriate relative comparisons. Data that has not been normalised can also lead to abnormally large correlation values due to the difference in scale. The formula is given by:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{3}$$

In equation 3, x defines the time series data, x_{min} and x_{max} are the minimum and maximum of the data respectively (Google For Developers 2022).

2.4.2 *Correlation Function*

To get a plot of correlation vs lags, I used "matplotlib.pyplot.xcorr". This function calculates the correlation between two time series as a function of lag. This works by shifting two time series in time, and calculating the correlation at each shift. Peaks in the correlation can reveal information about the magnitude and duration of a lag.

The formula used for the correlation with lag k used in "matplotlib.pyplot.xcorr" is given by:

$$\sum_n x[n+k] \cdot y^*[n] \tag{4}$$

Two time series are represented by x and y, where y^* is the complex conjugate of time series y. For the crypto-coin data, the lag k will be in steps of 1 hour (Matplotlib 2024) When interpreting lags, a peak in the cross-correlation may indicate a lag between two series, at the lag value the graph peaks at. A negative lagging peak in a correlation of x and y indicates that y is leading x and vice versa.

2.5 Volatility Analysis

Volatility represents the dispersion of returns for a cryptocurrency. Returns define the percentage change in a cryptocurrency over a time period. Understanding the volatility of a cryptocurrency can provide important information to investors who would like to make a profit when trading cryptocurrencies. This is because volatility represents the amount of uncertainty/risk in the value of a cryptocurrency. Higher values of volatility may indicate dramatic changes in the currency over a short period of time, whereas a lower volatility indicates a stable and consistent progression of a cryptocurrency. I will be calculating volatility by taking the daily percentage return of a cryptocurrency, and then working out the standard deviation of the daily returns (Hayes 2024b).

2.5.1 *Percentage Returns*

To calculate percentage returns, I used the module "pandas.DataFrame.pct_change" along with the daily open price of each coin. This module calculates the fractional change in the price, compared to the previous day:

$$Return = \frac{x_d}{x_{d-1}} \tag{5}$$

In equation 5, x_d represents the daily price, where x_{d-1} is the price on the previous day.

	BTC/ETH	BTC/ADA	ETH/ADA
2018/04/17-2024/04/22	0.929	0.777	0.833
2021/01/01-2023/01/01	0.803	0.610	0.729

Table 1. Comparison of Pearson's coefficient between Bitcoin, Etherium and Cardano, from 2018-2024 and 2021-2024.

2.5.2 *Standard Deviation*

The standard deviation represents the dispersion of a dataset about its mean. It is calculated by taking the square root of the variance. The variance represents the spread of the dataset. As such, calculating the standard deviation of the returns can capture any variation or dispersion in the data (Hargrave 2023).

$$\sigma = \sqrt{\frac{(\sum_i R_i - \mu)^2}{N}} \tag{6}$$

In equation 6, R represents the values of percentage return, μ is the mean of the dataset and N is the total number of data points. I calculated the standard deviation of the percentage return to acquire a measure of the volatility for each cryptocurrency.

2.5.3 *Plotting Standard Deviation vs Time*

A useful result for investors would be to understand how this deviation evolves over time, and whether it is worth investing in a currency in the future when the volatility has reduced. To understand this, I also plotted the standard deviation for Bitcoin, sampled at multiple regions in time. I did this by choosing progressively increasing regions in time, and working out the standard deviation of the percentage returns in this region.

3 RESULTS

3.1 Qualitative Comparison

I will begin the analysis by taking a qualitative view on the predominant characteristics of the cryptocurrencies. A comparison of each currency over the entire data range is shown in figure 1. From the figure, it is evident that all series have a large amount of noise, however this can be resolved through data smoothing. All of the series have a generally increasing trend, although this effect is reduced in Cardano. There are significant peaks in the data during 2021 and a steep increase during 2024 in Bitcoin and Ethereum.

3.2 Correlation

To understand how each cryptocurrency correlates with each other, I calculated the Pearson's Coefficient between each time series. I did this over two different time periods, to understand how this correlation changes over time. (Note, all expressions of date/time are given in the format (yyyy/mm/dd))

I chose the date range 2018/04/17-2024/04/22 to get an idea of the general correlation across the entire available date range. I chose 2021/01/01-2023/01/01 to see how the correlation has developed recently. I will also take any correlation above 0.7 to be strong, and 0.5-0.7 to be moderate as these are widely accepted ranges. From Table

Figure 1. The open price (USD) of Bitcoin, Ethereum and Cardano from 2018 to 2024.

1 we can see that over the entire hourly data range from 2018/04/17-2024/04/22, Bitcoin/Ethereum and Ethereum/Cardano has a strong correlation, above 0.8. Although Bitcoin/Cardano has a slightly lower correlation at 0.777, this can still be classified as a strong correlation. However, when the range of dates shifts to 2021/01/01-2023/01/01 the correlations reduce. In this region, all currencies have similar correlations, averaging to 0.770. Note, the p value for all of these correlations came out to 0, implying that the presented correlations are significant.

3.3 Smoothing Comparison

In section 2.3 I introduced 2 methods of smoothing data. The goal of data smoothing is to eliminate noise whilst maintaining the general trends of the time series. This allows for a better visualisation of the data and is often a requirement of predictive time-series models. In this section I will compare simple moving averages (SMA) to the Savitzky-Golay Filter. In figure 2 I have plotted hourly Bitcoin data from 2020/01/01 to 2024/04/22 and applied both smoothing methods. I used a window size of 500 for both plots, with a polynomial order of 3 for the Savitzky-Golay filter. Increasing the polynomial order seemed to decrease the level of smoothing, until an order of 7. At an order of 7 and above, the smoothing becomes unstable and the filter returns an error. As such I chose an order of 3, which provides a moderate amount of smoothing relative to orders in the stable range.

Comparing the Savitzky-Golay filter to the moving average, I find that moving averages result in higher levels of smoothing for the same window size as the Savitzky-Golay filter. It is important to mention that the moving average smoothing results in undefined values at the end of the time series, in the plot I have replaced them with 0s for visualisation. As such, when computing lags I will use the Savitzky-Golay filter to avoid the loss of data at the edges,

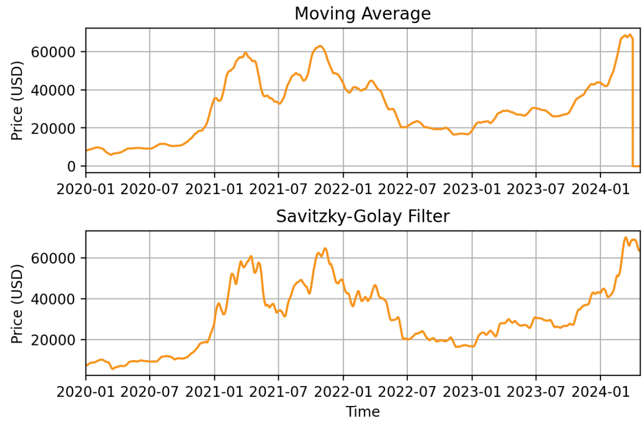

Figure 2. A comparison of moving average and Savitzky-Golay smoothing methods for Bitcoin time series data. Smoothing windows = 500 with a Savitzky-Golay order of 3.

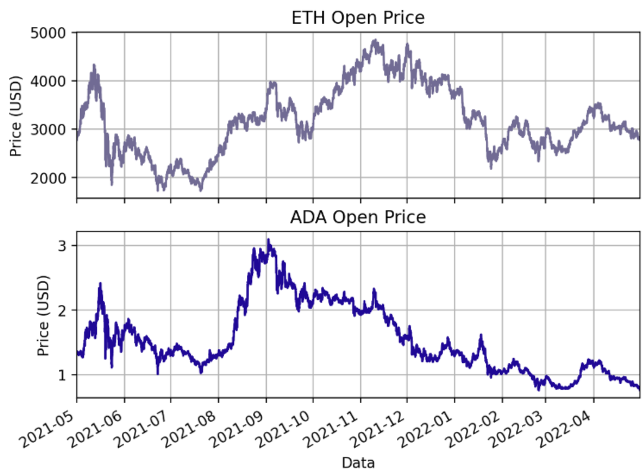

Figure 3. Open price of Ethereium and Cardano, from 2021-05-01 to 2022-05-01.

however both methods provide a significant reduction to the noise of the data without losing their characteristic features.

3.4 Lag Analysis

To calculate Lags, I used a cross correlation function to understand the relationship between each cryptocurrency. All data has been normalised, and smoothed using Savitzky-Golay filtering, with a window of 500 and polynomial order = 3. I started by looking at the cross correlation in the date range 2021-05-01 to 2022-05-01. I chose this range as Cardano has an earlier peak relative to the other currencies during this time. This can be seen in 3. Figure 4 shows the cross correlations for all 3 currencies in this date range. For BTC/ETH I find a peak of 0.981 at a lag of 53 hours. This implies that Bitcoin is leading Ethereum by approximately 2 days in this region. However, looking at the cross correlation of BTC with Cardano, I find 2 peaks. The first peak is centred near 0 lag, implying a strong relation between the movement of the currencies. However, the second and stronger peak is located at a lag of -1300 hours, equal to approximately -8 weeks. This implies that Bitcoin is following Cardano by 8 weeks.

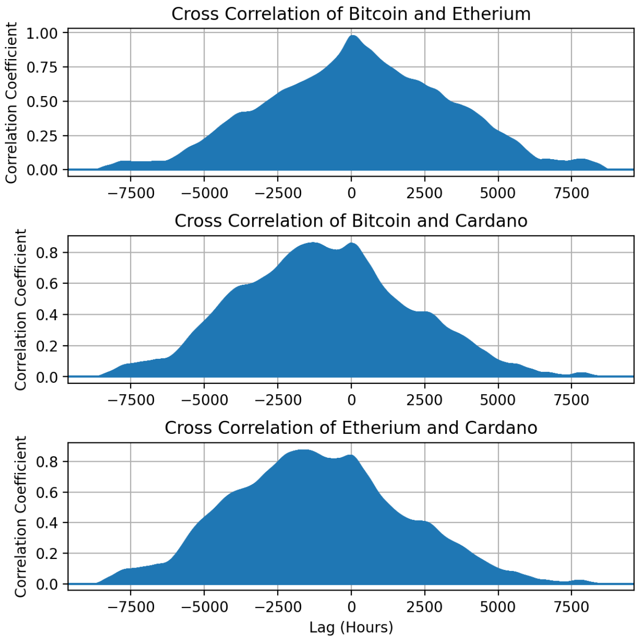

Figure 4. Cross-Correlation functions between Bitcoin, Etherium and Cardano, from 2021/05/01 to 2022/05/01.

A similar peak in ETH/ADA implies that Ethereum is also following Cardano, with a lag of approximately -8 weeks.
Taking the cross-correlation of the entire available date range (2018-04-17 to 2024-04-22) I found no lag between Bitcoin and Ethereum, with the peak in correlation occuring at 0 hours. When correlating BTC/ETH with ADA, I found 2 peaks again. 1 peak at 0 hours indicating no lag, however with a lower secondary peak at a correlation of ≈ 0.5. Taking the cross correlation at an earlier range (2018-04-17 to 2021-01-01), results in peaks at 0 lags, with similar distributions for all currencies. This implies that at this time, there are negligible lags between each currency.

3.5 Volatility Analysis

To analyse the volatility, I calculated the daily percentage returns on each of the cryptocurrencies, and then computed the standard deviation of the returns. For the entire range of dates (2018-04-17 to 2024-04-22) the standard deviations are given by: $\sigma_{BTC} = 3.69$, $\sigma_{ETH} = 4.87$, $\sigma_{ADA} = 5.45$ (all values quoted to 3 s.f.). I find that Bitcoin has the least deviation in returns whereas Cardano has the most deviation. To get a better idea of how this deviation changes over time, I repeatedly measured the variation for increasing date ranges. This is shown in figure 5. I used a linear model to fit this data, outlined in equation 7.

$$S = mx + c \tag{7}$$

In this equation, S represents the fit of the standard deviation, the gradient is defined by m, and c represents the y-intercept. The parameters are defined as -0.0308 ± 0.0109 and c = 4.072 ± 0.304. The negative gradient implies a reduction in standard deviation over time, however as m is small, this reduction is not large. The R^2 value between the fit and data = 0.145, implying a weak correlation between the data and the fit.

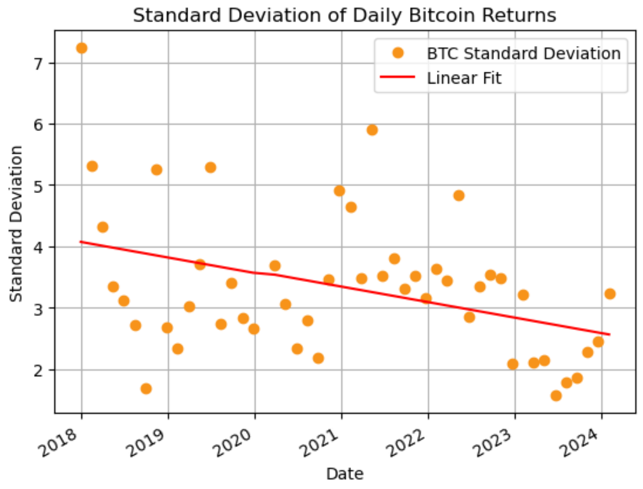

Figure 5. The standard deviation of daily percentage return on Bitcoin, from 2018 to 2024, linear fit. Standard Deviation in units %

4 DISCUSSION

In this section, I will outline interpretations of the results, and a comparison with real-world events.

An important technique in time-series analysis is data smoothing. In this paper I carried out a comparison between moving averages and Savitzky-Golay filtering. I found that moving averages seemed to have a more dramatic smoothing at similiar window sizes compared to Savitzky-Golay filtering. This is likely a result of the differences in window averaging, where the moving average simply takes a mean of all the data in the window. The Savitzky-Golay filter takes a polynomial fit of this data, which would ultimately capture more variation in the data resulting in a less drastic smoothing. I also noticed data loss, at the edges of the moving average filter. This is because undefined data cannot be incorporated into a simple mean. However, the Savitzky-Golay filter carries out a weighted regression, which takes into consideration the distance of points away from the centre of the window. Coupled with the polynomial fitting, the increased complexity in the Savitzky-Golay smoothing allows for 0 data loss when smoothing.
However, a development to this paper would be to forecast the growth of a time series. In this regard, an exponential filtering method would be more suitable than the ones I have used. Exponential filtering gives recent data points in the time series more weighting. This is crucial in making accurate predictions that do not rely on outdated developments of the time series (Brownlee 2020).

The price of a cryptocurrency can be governed by real-world developments, but also the developments in similar cryptocoins. As such, understanding the correlation between Bitcoin, Ethereum and Cardano was a primary aim of this project. Upon calculating the correlation between each currency, I found that over large ranges of time, there are significant correlations between the prices of Bitcoin and Ethereum, (Pearson's coefficient = 0.929). This implies that both currencies evolve similarly and that one may be the cause of another. However, when narrowing down the range of dates from 2021-2024, I found a decrease in this correlation by -0.126. A possible cause for this could have been the introduction of an upgrade to Etherium's blockchain known as the 'Shanghai hard fork'. The Shanghai hard

fork allows users who have staked their Ethereum to withdraw staked assets. The introduction of this new technology may have caused the rift in the correlation between Bitcoin and Ethereum. These findings have implications for investors who partake in hedging. Hedging is a strategy used to offset losses by investing in a related currency. As such, investors should alter their hedging strategies, as strategies may rely on cross hedging one asset for the other. The discrepancy between Bitcoin and Ethereum may affect this process. (Torpey 2023).
Looking at the correlation between BTC/ETH and Cardano, I found positive correlations with Cardano, but of lower magnitude. In the full date range this correlation = 0.777, between 2021 and 2024. A reason for lower correlations, are likely due to the inherent differences in the underlying mechanism of each currency. Cardano is an example of a proof-of-stake cryptocurrency. In this mechanism, "validators" are chosen to validate transactions on the blockchain, whilst staking a certain number of coins as collateral. Validators are chosen in proportion to the amount of coins they hold. However, Bitcoin is an example of a proof-of-work cryptocurrency which where specific miners are chosen to verify data, who are rewarded with Bitcoin upon accurately validating data. This fundamental difference could account for the lower correlations between BTC/ADA. Ethereum is also an example of a proof-of-stake currency, however this change was only introduced in 2022, which could explain why it has higher correlations with Cardano compared to BTC/ADA (Binance; Napoletano 2024).

In section 3.4 I calculated the lags between each time series, over different date ranges. From 2021/05/01 to 2022/05/01 I found Bitcoin and Ethereum to be strongly correlated with Bitcoin leading Ethereum by 2 days. However, across the entire date range I found the lags to be at 0 hours between Bitcoin and Ethereum. This shows that in general, there is no discernible relation between either currency. Both must evolve due to external factors, such as market demand, mining costs and general economic trends. The fact that Bitcoin led Ethereum from 2021/05/01 to 2022/05/01 could be a result of a delayed reaction to external factors by Ethereum, or some low-level relationship between each coin.
I found that Cardano seemed to lead both Bitcoin and Ethereum in the entire date range, especially from 2021/05/01 to 2022/05/01. This can be visualised in figure 3. I will use the comparison of ETH/ADA as BTC/ADA follows a similar evolution. The cross-correlation indicated a 10 week lag between Etherium and Cardano, this would indicate that Etherium is following/reacting to the movement of Cardano, or delayed in its reaction to external factors. However, looking at figure 3 it is evident that Cardano is not actually leading Ethereum. An earlier peak in Cardano at 2021-09 is simply a stronger version of the same peak in Ethereum. This implies the cross-correlation has erroneously taken the peak in Ethereum at 2021-11 to be a lagged version of the Cardano peak at 2021-09. This has the implication that Cardano has a more unstable evolution relatively to Bitcoin and Ethereum, with the price peaking/dropping more aggressively than BTC and ETH. I further investigated this phenomena by taking a volatility analysis of all cryptocurrencies.

In the volatility analysis, this instability in Cardano was confirmed when I found its standard deviation on returns to be higher than both Bitcoin and Ethereum. Although the deviations did not differ significantly, this may impact investors who would like to invest in Cardano, as it may pose an increased risk relative to Bitcoin/Ethereum. Cardano being a relatively new cryptocurrency means that it is still under development. This means that it likely has a smaller market which could be a reason for these deviations. I also investigated how the deviations changed over time, specifically for Bitcoin. I found that there is a weak, decreasing trend in the volatility of Bitcoin over time. As such, investors who may have avoided Bitcoin in the past due to this volatility may find that Bitcoin is becoming more stable and should consider investing. Ultimately, this analysis has shown that Bitcoin, Ethereum and Cardano all possess significant fluctuations on returns, that should be considered when making investments in these cryptocurrencies.

To build upon the findings in this paper, a predictive model can be used to predict how the value of each cryptocurrency will develop in the future. This will provide useful data to investors who may be anticipating a decline/increase in the price of a certain currency they have invested in. For example, a model such as Autoregressive Integrated Moving Average (ARIMA) could be used. ARIMA uses lagged moving averages to smooth time series data and can be used in technical analysis to forecast prices of a time series (although this model is only used for short-term forecasting) (Hayes 2024a). The volatility analysis I completed took on a relatively naive approach in calculating the standard deviation of the returns, as such a model such as GARCH could be used to improve this. The Generalised Autoregressive Conditional Heteroskedasticity (GARCH) model can estimate the volatility of financial markets whilst taking into account volatility clustering (Kenton 2020).

5 CONCLUSION

In this paper, I completed a time series analysis to understand the complex dynamics and relationships between Bitcoin, Ethereum and Cardano. By employing time series analysis I was able to understand the correlations between each cryptocurrency. I found that Bitcoin and Ethereum had strong correlations from 2018-2024. During 2021-2024 this dropped, likely a result of technological advancements in Ethereum. Cardano had low correlations with Bitcoin and Ethereum, likely a result of the inherent differences in each coin. I found that Savitzky-Golay filtering provided an optimal method of smoothing the data, to reveal underlying trends. Moving averages resulted in data loss. Savitzky-Golay effectively reduces noise, providing a clean time series for the cross-correlation analysis. I found no lags between Bitcoin and Ethereum from 2018-2024, however from 2021 to 2022, Bitcoin appeared to lead Ethereum. This variation is likely due to external factors or some low-level relationship between each coin. I found that Cardano seemed to lead Bitcoin and Ethereum, however this was a false positive due to the volatile nature of Cardano. Finally, I found that the volatilty in Bitcoin has decreased from 2018, although not significantly. Ultimately, these findings may provide valuable insights to investors who would not want to invest in a volatile coin. I also find that for investors who would like to use Bitcoin as a predictor of coins, it may be important to focus on external factors that influence the progression of cryptocurrencies.

6 ADDENDUM

The addendum outlines 3 techniques: Integration, The Runge-Kutta Method and Fourier Transforms. I will include 3 applications in academia, and 3 outside academia for each technique.

6.1 Integration

Integration is a fundamental component of calculus. It is a technique used to calculate the sum of a continuous line, area or volume by summing over infinitely small components (Berggen 2018).

6.1.1 Integration in Finance, Medicine and Fluid Mechanics

In finance, sophisticated mathematical techniques have provided significant benefits in understanding financial assets. In finance, an asset is a contract specifying a claim to a payment. These assets are often traded, however, this trade can be agreed to occur at a later time T. This is known as a forward contract. A forward contract is a derivative asset, as before time T, the price of the underlying asset may become positive or negative. A variant of the forward contract is the European call option, which states the holder has the option to buy the asset at time T, but is not obliged. The Black-Scholes formula defines the value at a time t of a European asset. In this context, integration can be used to improve this model, by deriving the partial differential equation behind this formula (Muldowney and Wojdowski).

In medicine, integration is used in Pharmocokinetics. This field aims to understand the dynamics of drug metabolism, distribution and excretion in the body. A key problem in this field is understanding how drugs are eliminated in the body. This process can be measured in the form of clearance which defines the ratio of the administered dose to the "area under the curve" (AUC). AUC represents the total exposure to a drug and is calculated via integration. AUC is important as it allows scientists to understand the effects of changes in drug concentrations on the clearance, such as altering a drug schedule (Pollock and Weichselbaum 2003).

In Industrial Instrumentation, integration is used to understand fluid flow rates. Flow-meters are used to measure the flow rate of fluid outputs, where the integration of the output signal is used to get the volume of the total fluid. An example of such a device is the "Foxboro Pneumatic Flow integrator", which displays the total mass of fluid passing through a pipe using an odometer. It utilises a jet of compressed air to spin a turbine wheel. The speed of the wheel is proportional to the flow rate, which is able to turn the odometer and register the total fluid passing through (Control Automation).

6.1.2 Integration in Computer Science, Quantum Mechanics and Chemistry

In Quantum Mechanics, integration is used to predict fundamental properties of different particles. Quantum mechanics aims to predict uncertain observables of particles, such as momentum and position. Due to the wave/particle like nature of all atoms, estimating such quantities involves interpreting the probability distributions that govern the behaviour of these particles. The probability distribution is defined as the square of the particles wave function, where the wave function is a solution to the Schrödinger equation. This distribution can be integrated over all momentum or position space to find the expectation value of these variables (Griffiths 2018).

In Chemistry, integration is used in chemical kinetics. This is the study of the rates of changes in concentrations of chemicals during reactions. The Rate Law defines the instantaneous rate of a reaction as a function of concentration. A 1st order reaction follows a rate law with an exponential change in concentration, resulting in a linear natural log relation. The integrated rate law performs integration on this 1st order law which can provide further information about the relationship between concentration and time (Florida State University).

In computing, there are a variety of ways to calculate integrals through the use of numerical integration. For example, the trapezium rule aims to calculate the area under a curve by dividing it into N trapeziums. The trapeziums act as discrete elements which are summed over to compute the total area. Some other examples include Monte Carlo integration, Simpson's rule and the midpoint rule (Krishna 2021).

6.2 Runge-Kutta

The Runge-Kutta Method is used to solve differential equations. Runge-Kutta constructs higher order solutions to a problem without the explicit calculation of the functions derivative. This is useful as high order derivatives which may require significant amounts of computation (Zheng and Zhang 2017).

6.2.1 Runge-Kutta in Engineering, Finance and Medical Simulations

In Engineering, the Runge-Kutta method is used in understanding vibratory systems. Phase lag is an important problem when designing systems with some form of signal transmission, to ensure responsiveness in the system. In such systems Runge-Kutta pairs are used to derive ordinary differential equations, with solutions that govern the free oscillations of high and low frequencies in the system. Traditional techniques may introduce a phase lag when deriving these solutions, as such the Runge-Kutta method is used to minimise these lags (Tsitouras 2002).

In Finance, Runge-Kutta methods are used for pricing "American Contracts". An American Contract is a type of derivative security. Derivative securities are traded on stock markets, and can include call options, bonds and dividend stocks. In an American Contract, a holder can exit the contract at any time before its expiration. Valuating these contracts can prove difficult, and in this regard numerical methods must be used to price them. Runge-Kutta methods can be used to numerically integrate a system of ordinary differential equations that define the pricing of contracts (de Frutos 2006).

Telemedicine is a field of medicine which aims to simulate deformations of tissues and organs in real time, as realistically as possible. For example, lung surgery simulation platforms can be built using OpenGL that verify the realism and accuracy of such models. Model accuracy is important, however this comes at the cost of real time simulation. To balance these components, the Runge-Kutta method is used to solve for the relationship between applied force on the lungs and its subsequent deformation. Results using this method have shown real-time reproduction of lung models in remote surgery . These models also have the ability of describing lungs suffering from disease (Zhang et al. 2022).

6.2.2 Runge-Kutta in Physics, Chemistry and Ecology

In Physics, the Runge-Kutta method is used to simulate planetary motions in the Solar System. Newton's Theory of gravity provides an explanation for the dynamics and evolution of the motion of bodies in our Solar System. A set of differential equations describing this motion can be derived and then solved numerically with initial values. This problem can be solved using Runge-Kutta methods to analyse the solutions. This method can provide increased accuracy compared to traditional techniques such as Euler's method, and provide stable solutions that are consistent over a wide range of conditions (Seyr and Puschnig 2020).

In Chemistry, Runge-Kutta Methods are used to understand molecular dynamics. In simulating molecular dynamics, it is important to find an efficient solution to evaluating the interactions between particles. This is because it involves a large number of time steps, and evaluations of force. Traditionally the "Verlet Velocity Method" is used, and works with small time steps, but cannot compute longer time steps. As such, Runge-Kutta Methods have been used which prove to be more accurate and efficient, particularly in models where individual behaviour is important, such as protein folding (Dahlgren et al.).
In Ecology, models are created to understand reaction-diffusion systems that describe the dynamics in different ecological interactions between species. These systems can be defined by partial differential equations that feature low-order nonlinear and high-order linear terms. Runge-Kutta methods can be used to understand stability in these models. For example, how perturbations in the system or environmental changes effect the model. Runge-Kutta can also provide dynamical consistency and reliability in deriving these computational results (Owolabi and Patidar 2016).

6.3 Fourier Transforms

Fourier transforms are used to represent the strength of a time-dependant sinusoidal in frequency space. These components can then be integrated with a combination of sines and cosines to represent the complete sinusoid, in what is called an inverse Fourier transform (Rouse 2020).

6.3.1 Fourier Transforms in Astronomy, Acoustics and Finance

In Astronomy, Fourier Transforms are used when processing flux distributions from stellar body. For example, light curves can be observed from stars, in the form of time series data. A power spectrum of the data can be taken, where the power spectrum represents the modulus squared of the Fourier components. This a result of the complex-nature of a Fourier transform. The power spectrum may reveal details about the system, for example, peaks could be a result of differential rotation or star spot activity, coupled with magnetic interactions from a companion star (Longair 2012).
In acoustics, fourier transforms can be used to reduce the noise from domestic appliances. For example, dishwashers, refrigerators and washing machines can all produce unwanted vibrations during their operation. Manufacturers aim to mitigate these effects by using Fast Fourier Transforms (FFT). By recording the sound produced by these appliances and applying the FFT, manufacturers can convert this time series into frequency space and analyse the fundamental components. This reveals natural frequencies of the vibrating system, allowing for more targeted measures to reduce vibrations and minimise noise in the appliance (Harčarik et al. 2012).
In Financial Economics, Fourier transforms are a widely used tool that can model real time pricing of assets, whilst taking into account volatility and randomness in the price of assets. The Fast Fourier Transform can be used in binomial option pricing where option prices can be efficiently computed by decomposing a payoff function into its frequency components. This allows for increased accuracy when estimating option pricing models. For investors, this may prove useful in analysing the outcomes of option strategies and completing risk assessments (Černý 2004).

6.3.2 Fourier Transforms in Quantum Mechanics, Wave Physics

In Quantum mechanics, free particles can be described as plane waves. In this description, particles have a wavelength described by the de Broglie equation which relates the wavelength of the plane wave to its momentum. Fourier transforms can be used to understand the position distribution of the particle. Performing a Fourier transformation on this expression results in a Dirac function centred on the location of the particle (Rioux 2022).
In Wave Physics, Fourier Transforms can be used to understand Fraunhofer diffraction. Fraunhofer diffraction is observed when waves encounter an obstacle/slit and diffract as they spread around it. The plane waves form diffraction patterns as they interfere with each other. Fourier transforms can be used to understand how the complex wave-forms contribute to the diffraction patterns. This can be achieved by decomposing wave-fronts into respective plane waves. This can extract features of the incoming waves such as wave intensity (Freegarde 2013).
In Chemistry Fourier Transforms are used in crystallography. Instead of using a lens to collect scattered diffraction from a crystal, a Fourier transform can be used to mathematically construct the image. It also has applications in spectroscopy, where the Fourier transform can decompose complex spectra. Fourier transforms can also enhance a signal due to technical limitations of detectors, or limitations in diffracting slits. FTs can extract the true spectrum of the physical process (Glasser 1987).

REFERENCES

Berggren, J.L. (2018). calculus | Definition Facts. In: Encyclopædia Britannica. [online] Available at: https://www.britannica.com/science/calculus-mathematics

Binance (n.d.). Bitcoin vs Cardano | Compare BTC and ADA | Binance.com. [online] Binance. Available at: https://www.binance.com/en-GB/event/bitcoin-vs-cardano [Accessed 12 May 2024].

Brownlee, J. (2020). A Gentle Introduction to Exponential Smoothing for Time Series Forecasting in Python. [online] Machine Learning Mastery. Available at: https://machinelearningmastery.com/exponential-smoothing-for-time-series-forecasting-in-python/.

Cerny, A. (2004). Introduction to Fast Fourier Transform in Finance. SSRN Electronic Journal. doi:https://doi.org/10.2139/ssrn.559416.

Coin Tree (2021). Cointree. [online] Cointree.com. Available at: https://www.cointree.com/learn/the-three-types-of-cryptocurrency-analysis/ [Accessed 12 May 2024].

Control Automation (). Calculus in Industrial Process Measurement and Control. [online] Available at: https://control.com/textbook/calculus/numerical-integration.

Dahlgren, S., Kling, H. and Lindblad, E. (n.d.). New Runge-Kutta algorithms for Molecular Dynamics. Available at: https://www.it.uu.se/edu/course/homepage/projektTDB/vt04/projekt4/mdrapport.pdf de Frutos, J. (2006). Implicit–explicit Runge–Kutta methods for financial derivatives pricing models. European Journal of Operational Research, 171(3), pp.991–1004. doi:https://doi.org/10.1016/j.ejor.2005.01.013.

Fernando (2024). The Correlation Coefficient: What It Is, What It Tells Investors. [online] Investopedia. Available at: https://www.investopedia.com/terms/c/correlationcoefficient.asp#:~:text=To%20calculate%20the%20Pearson%20correlation.

Florida State University (n.d.). Kinetics. [online] www.chem.fsu.edu. Available at: https://www.chem.fsu.edu/chemlab/chm1046course/integrated.html#:~:text=Integrated%20Rate%20Law%3A&text=This%20equation%20is%20a%20differential[Accessed12May2024].

Freegarde, T. (2013). Introduction to the Physics of Waves. [online] Cambridge University Press. Cambridge: Cambridge University Press. Available at: https://www.cambridge.org/core/books/introduction-to-the-physics-of-waves/A3A5CEFF7AD4BE9B7859B7C17CC700DE [Accessed 12 May 2024].

Glasser, L. (1987). Fourier transforms for chemists. Part 1. Introduction to the Fourier transform. Journal of Chemical Education, 64(10), p.A228. doi:https://doi.org/10.1021/ed064pa228.

Google For Developers (2022). Normalization. [online] Google Developers. Available at: https://developers.google.com/machine-learning/data-prep/transform/normalization.

Griffiths, D.J. and Schroeter, D.F. (2018). Introduction to Quantum Mechanics. [online] Higher Education from Cambridge University Press. Available at: https://www.cambridge.org/highereducation/books/introduction-to-quantum-mechanics/990799CA07A83FC5312402AF6860311E#overview.

Harčarik, T., Bocko, J. and Masláková, K. (2012). Frequency Analysis of Acoustic Signal using the Fast Fourier Transformation in MATLAB. Procedia Engineering, 48, pp.199–204. doi:https://doi.org/10.1016/j.proeng.2012.09.505.

Hargrave, M. (2023). Standard deviation formula and uses vs. variance. [online] Investopedia. Available at: https://www.investopedia.com/terms/s/standarddeviation.asp.

Hayes (2023a). What Is Cross-Correlation? Definition, How It's Used, and Example. [online] Investopedia. Available at: https://www.investopedia.com/terms/c/crosscorrelation.asp#:~:text=In%20its%20simplest%20version%2C%20the [Accessed 12 May 2024].

Hayes, A. (2023b). Simple Moving Average - SMA. [online] Investopedia. Available at: https://www.investopedia.com/terms/s/sma.asp.

Hayes, A. (2024a). Autoregressive Integrated Moving Average (ARIMA). [online] Investopedia. Available at: https://www.investopedia.com/terms/a/autoregressive-integrated-moving-average-arima.asp.

Hayes, A. (2024b). Volatility Definition. [online] Investopedia. Available at: https://www.investopedia.com/terms/v/volatility.asp.

Hooson, M. (2022). Bitcoin vs. Ethereum. [online] Forbes Advisor. Available at: https://www.forbes.com/advisor/investing/cryptocurrency/bitcoin-vs-ethereum/.

Kaspersky (2022). What Is Cryptocurrency and How Does It work? [online] Kaspersky. Available at: https://www.kaspersky.com/resource-center/definitions/what-is-cryptocurrency.

Kenton, W. (2020). GARCHP rocess. [online] Investopedia. Available at: https://www.investopedia.com/terms/g/generalalizedautogregressiveconditionalheteroskedasticity.asp#:~:text=The%20generalized%20autoregressive%20conditional%20heteroskedasticity%20(GARCH)%20process%20is%20an%20approach.

Krishna (2021). What is numerical integration? [online] Goseeko blog. Available at: https://www.goseeko.com/blog/what-is-numerical-integration/.

Longair, M.S. (2012). High Energy Astrophysics. 3rd ed. [online] Cambridge University Press. Cambridge: Cambridge University Press. Available at: https://www.cambridge.org/core/books/high-energy-astrophysics/CF25E2E5FC0EDFC51FCD7846A262C0AE [Accessed 12 May 2024].

Matplotlib (2024). matplotlib.pyplot.xcorr — Matplotlib 3.8.4 documentation. [online] matplotlib.org. Available at: https://matplotlib.org/stable/api/_as_gen/matplotlib.pyplot.xcorr.html [Accessed 12 May 2024].

Muldowney, P. and Wojdowski, W. (n.d.). Integration Theory in Financial Mathematics. [online] Available at: https://www.emis.de/proceedings/Toronto2000/papers/wojdowski.pdf.

Napoletano (2024). What Is Proof Of Work? – Forbes Advisor. [online] www.forbes.com. Available at: https://www.forbes.com/advisor/investing/cryptocurrency/proof-of-work/#:~:text=Proof%20of%20work%20is%20a%20consensus%20mechanism%20to%20choose%20which.

Owolabi, K.M. and Patidar, K.C. (2016). Numerical simulations of multicomponent ecological models with adaptive methods. Theoretical biology and medical modelling, 13(1). doi:https://doi.org/10.1186/s12976-016-0027-4.

Pollock and Weichselbaum (2003). Principles of Pharmacokinetics. [online] www.ncbi.nlm.nih.gov. Available at: https://www.ncbi.nlm.nih.gov/books/NBK12815/#:~:text=AUC%2C%20the%20area%20under%20the.

PwC (2023). Making sense of bitcoin and blockchain. [online] PwC. Available at: https://www.pwc.com/us/en/industries/financial-services/fintech/bitcoin-blockchain-cryptocurrency.html.

Reiff, N. (2024). How Much of All Money Is in Bitcoin? [online] Investopedia. Available at: https://www.investopedia.com/tech/how-much-worlds-money-bitcoin/.

Rioux (2022). 1.13: Quantum Mechanics and the Fourier Transform. [online] Chemistry LibreTexts. Available at: https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Quantum_Tutorials_(Rioux)/01%3A_Quantum_Fundamentals/1.13%3A_Quantum_Mechanics_and_the_Fourier_Transform [Accessed 12 May 2024].

Rouse, M. (2020). Fourier Transform. [online] Techopedia. Available at: https://www.techopedia.com/definition/7292/fourier-transform.

Seth, S. (2024). Explaining the Crypto in Cryptocurrency. [online] Investopedia. Available at: https://www.investopedia.com/tech/explaining-crypto-cryptocurrency/.

Seyr, A. and Puschnig, P. (2020). Numerical simulation of the planetary motions in the solar system with Runge Kutta Methods. [online] Available at: https://static.uni-graz.at/fileadmin/_Persoenliche_Webseite/puschnig_peter/unigrazform/Theses/BachelorThesis_Seyr_2020.pdf.

Tambe, N. (2024). Advantages and Disadvantages of Cryptocurrency in 2023. [online] Forbes Advisor INDIA. Available at: https://www.forbes.com/advisor/in/investing/cryptocurrency/advantages-of-cryptocurrency/#:~:text=A%20cryptocurrency%20is%20a%20digital.

Torpey (2023). What Does Lower Bitcoin-Ether Correlation Mean for Investments? [online] Investopedia. Available at: https://www.investopedia.com/lower-bitcoin-ether-correlation-7484225.

Tsitouras, C. (2002). Explicit Runge–Kutta pairs appropriate for engineering applications. Applied Mathematical Modelling, 26(1). doi:https://doi.org/10.1016/S0307-904X(01)00042-7.

Vaughan, S. (2015). Random time series in astronomy. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 371(1984), p.20110549. doi:https://doi.org/10.1098/rsta.2011.0549.

Wronski (2023). Study of smoothing filters – Savitzky-Golay filters. [online] Bart Wronski. Available at: https://bartwronski.com/2021/11/03/study-of-smoothing-filters-savitzky-golay-filters/.

Zhang, X., Zhang, W., Sun, W. and Song, A. (2022). A new soft tissue deformation model based on Runge-Kutta: Application in lung. Computers in Biology and Medicine, [online] 148, p.105811. doi:https://doi.org/10.1016/j.compbiomed.2022.105811.

Zheng and Zhang (2017). Runge-Kutta Method - an overview | ScienceDirect Topics. [online] www.sciencedirect.com. Available at: https://www.sciencedirect.com/topics/mathematics/runge-kutta-method.

This paper has been typeset from a TEX/LATEX file prepared by the author.