Forecasting Bitcoin Market Price and Measuring It's Risk

A report submitted to the Department of Mathematics as a partial fulfillment of the Bachelor of Science Honors Degree in Financial Mathematics and Industrial Statistics

By

B.L.A.W. Lakshika

Supervisor :

Dr. N. Yapage

Department of Mathematics

University of Ruhuna

Matara

31.01.2025

Declaration

I, B.L.A.W.Lakshika, declare that the presented project report titled, “Forecasting Bitcoin Market Price and Measuring It’s Risk” is uniquely prepared by me based on the group project carried out under the supervision of Dr. N. Yapage, Department of Mathematics, Faculty of Science, University of Ruhuna, as a partial fulfillment of the requirements of the level III , Case Study course unit, MFM 3151 of the Bachelor of Science Honours Degree in Financial Mathematics and Industrial Statistics in Department of Mathematics, Faculty of Science, University of Ruhuna, Sri Lanka.

It has not been submitted by me to any other institution or academic program for any other purpose.

Signature : ....
Date : ....

Supervisor’s Recommendation

I certify that this study was carried out by B.L.A.W. Lakshika under my supervision.

Signature : ....
Date : ....

Department of Mathematics
Faculty of Science
University of Ruhuna

Acknowledgement

First of all, my sincere gratitude goes to our supervisor, Dr. N. Yapage, for his unconditional support for us to succeed in this research. As well, I would like to thank our course coordinator Ms. I.C. Liyanage, who gave us more valuable instructions to fulfill our case study successfully. And also, I extend my thanks to all the lecturers in the department for their great support towards us. Moreover, I extend my thanks to all the instructors who helped me to make this research a success. Also, I express my sincere appreciation to all our batch mates for sharing their knowledge with me. Finally, I am thankful to everyone who contributed to making this research a success.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACF	Auto Correlation Function
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
JB	Jarque-Bera
MSE	Mean-square Error
PACF	Partial Auto Correlation Function
RMSE	Root Mean-square Error
DES	Double Exponential Smoothing
VaR	Value at Risk
ARIMA	Autoregressive Integrated Moving Average
NA	Not Available

Abstract

Bitcoin is one of the currencies in the crypto market characterized by extreme volatility and uncertainty, making quantifying traditional financial risk measures insufficient for capturing its complexities and unpredictable price behavior. This study aims to forecast Bitcoin prices and measure market risk using double exponential smoothing for time series forecasting and Shannon entropy, volatility, and value at risk (VaR) for risk assessment. Historical Bitcoin price data from 2018 to 2024 was analyzed using statistical techniques in R programming. The forecasting model achieved a Mean Absolute Percentage Error (MAPE) of 2.34%, indicating strong predictive accuracy. Risk analysis revealed that Bitcoin's volatility and VaR were highly correlated, with significant market risks during economic disruptions. Shannon entropy ranged between 0.4 and 1.0, with higher values indicating increased uncertainty. Results indicate that Shannon entropy offers valuable insights into the dynamic behavior of Bitcoin prices and invented a reliable tool for risk assessment in this volatile market. This study elaborates on the importance of alternative approaches to risk evaluation and forecasting to address the unique challenges posed by Bitcoin, one of the most volatile assets in the financial landscape.

Key words: Bitcoin, Time Series Forecasting, Holt's Linear Trend Method, Risk metrics.

Chapter 1

Introduction

1.1 Background of the study

Cryptocurrencies are digital and virtual currencies that rely on the blockchain network that uses encryption for transactions. This is regulated by the government, and like most fiat currencies, there is one issuing authority. Cryptocurrencies have a decentralized parallel system that makes them safe and convenient.

After being introduced in 2008, the rise in the price of bitcoin and the popularity of other cryptocurrencies triggered a growing discussion about how much energy was consumed during the production of this currency. Making cryptocurrency the most expensive and most popular, both the business world and the research community have begun to study the development of bitcoin. In this study, bitcoin price predictions are performed using the double exponential smoothing method.

But it is clear that cryptocurrencies, especially their markets prices fluctuating based on the factors of speculation, market nature, regulations and technology use. This also applies to Bitcoin, a type of cryptocurrency. Not all risks and price movements used in traditional finance are suitable for Bitcoin. To solve this problem, this study introduces Shannon entropy, Professor Shannon, a measure of uncertainty in the domain of information theory, to provide a proper assessment of risk in virtual asset ecosystems.

1.2 Literature review

Nowadays, cryptocurrency markets have attracted significant attention from global investors due to their innovative nature, widespread online accessibility, growing market capitalization, and potential for high returns. So many researchers pay their attention to the new invention in that field, Such as forecasting the prices and quantifying the market risk. The study done by Baharaeen and Masud (1986) convey that exponential smoothing is suitable to forecast the bitcoin data, which are often used to solve time series data problems. According to the study, forecasting in exponential smoothing is done by repeating calculations continuously using the latest data, where newer values are given a relatively greater weight than older observations. MAPA has been used to measure the suitability of the model. Moreover, they are highlighted that the lowest MAPA value gave the most suitable model rather than others.

The crypto market is a highly volatile, dynamic market. Because of that, some investors are afraid to invest in the market. So researchers have made an attempt to quantify the risk in the market. The study by Ormos and Zibriczky (2014) gave us valuable insight into using Shannon entropy in the Crypto market. According to the study done by Huang et al. (2019) , it has been shown that volatility has been one of the most widely studied and utilized measures of financial risk in modern finance.

1.3 Problem statement

The rapid fluctuations and high volatility of Bitcoin prices present significant challenges for traders, investors, and financial analysts. Accurate forecasting of Bitcoin prices and effective risk assessment are useful when making decisions in this market. From this study, we aim to find how to accurately forecast Bitcoin price by using double exponential smoothing and capturing the inherent seasonality and trends of the cryptocurrency market. Additionally, the study will be engaged to find how to model Shannon entropy as a quantitative risk measure to evaluate the uncertainty within the market. By comparing entropy-based risk measures and traditional risk measures, this research seeks to provide a comprehensive framework for predicting Bitcoin prices and quantifying risks in the Bitcoin market.

1.4 Objectives

Our key goals are as follows,

- To accurately forecast Bitcoin prices using Double Exponential Smoothing.
- To quantify market risk and uncertainty using Shannon Entropy.
- To compare Shannon entropy with traditional financial risk metrics for evaluating market volatility.

1.5 Significance of the study

The cryptocurrency industry is now quite well-known among youthful investors. Among these, Bitcoin is one of the most powerful digital currencies. Finding an appropriate model for predicting the price of bitcoin and comparing risk metrics like Shannon entropy, volatility, and VaR to build the optimum model for identifying risk are the most important aspects of this study.

From this study, Bitcoin holders will understand the crypto market, and newcomers can invest their money without doubt. Plus, the study reveals effective risk measures for investors.

Chapter 2

Methodology

2.1 Research approach

This study used the time series and entropy-based quantitative research approach.to forecast and quantify uncertainty and volatility in cryptocurrencies. The study investigated hypotheses related to the models' performance and the importance of their forecasts. These quantitative techniques are suitable for identifying instances, situations and logical relationships between results within datasets.

2.2 Research design

This study employed a longitudinal study design to forecast the data set by using double exponential smoothing analysis. Moreover, it applied Shannon entropy, volatility, and VaR as tools to measure the uncertainty and unpredictability in the price movements of Bitcoin. So the stepwise approach was adopted to forecast the data set and calculate the entropy, volatility, and VaR to make a conclusion. The design applies the wide view to reveal significant information that may help cryptocurrency holders.

2.3 Methods used for data analysis

In this study, we used time series analysis, Shannon entropy, volatility, and VaR. And also R software was used to do the analysis and to interpret results.

2.3.1 Time series analysis

Time series analysis is a specific way of analyzing a sequence of data points collected over an interval of time. In time series analysis, analysts record data points at consistent intervals over a set period of time rather than just recording the data points intermittently or randomly.

Time series analysis achieves the best results when researchers maintain datasets comprising plenty of data points to reach consistency and reliability. Extensive data collection allows for selecting an adequate representative sample that enables robust analysis across the noisy data. The analysis verifies that the detected trends do not originate from exceptional observations while considering regular seasonal variations. Furthermore, time series data provides a basis for forecasting, which involves predicting future data points from historical data points.

Mathematically, a time series is defined by the values $x_1, x_2, \cdots, x_n$ of a variable x ob served at times $t_1, t_2, ..., t_n$.

Thus, x is a function of t, symbolized by:

$$\boldsymbol{x = f(t)},$$

where

$$x = \text{The observed variable or value.}$$
$$t = \text{The corresponding time point.}$$

A time series consists of 4 components. They are,

1. **Trend** $(\mathbf{T}_t)$

   It shows long-term movement or direction in the data set. It can be upward, downward, or flat. These movements represent whether the data is increasing, decreasing, or stable over time.

2. **Seasonal variation** $(\mathbf{S}_t)$

   It is short-term fluctuations in a time series that occur periodically in a year. This continues to repeat year after year.

3. **Cyclical variation** $(\mathbf{C}_t)$

   It shows recurrent upward or downward movements in a time series, but the period of the cycle is greater than a year. Also, these variations are not regular as seasonal variation.

4. **Irregular variation** $(\mathbf{I}_t)$

   Irregular variations are fluctuations in time series that are short in duration, have an irregular nature, and follow no regularity in the occurrence pattern.

By using the above components, we can decompose a time series like below. It can be either additive or multiplicative, depending on how the components interact.

1. Additive Model

$$\boldsymbol{X_t = T_t + S_t + C_t + I_t},$$

2. Multiplicative Model

$$\boldsymbol{X_t = T_t \times S_t \times C_t \times I_t},$$

where

X_t is the observed time series value at time t.
T_t is the secular Trend.
S_t is the seasonal Variation.
C_t is the cyclical Variation.
I_t is the irregular or random variation.

- **Autocorrelation Function (ACF)**

  This function measures the correlation between a time series and its lagged values over different time periods. It helps to identify patterns, trends, and dependencies within the dataset.

  The autocorrelation function at lag k ,

$$\boldsymbol{\rho_k} = \mathrm{ACF}(\boldsymbol{k}) = \frac{\mathrm{Cov}(\boldsymbol{X_t, X_{t-k}})}{\sqrt{\mathrm{Var}(\boldsymbol{X_t}) \cdot \mathrm{Var}(\boldsymbol{X_{t-k}})}} = \frac{\boldsymbol{\gamma_k}}{\boldsymbol{\gamma_0}},$$

  where

  ρ_k: Autocorrelation at lag k
  X_t: The value of the time series at time t
  X_{t-k}: The value of the time series at lag k
  $\mathrm{Cov}(X_t, X_{t-k})$: The covariance between X_t and X_{t-k},
  $\mathrm{Var}(X_t)$: The variance of X_t
  $\mathrm{Var}(X_{t-k})$: The variance of X_{t-k}

The autocorrelation coefficient ACF(k) lies in the range,

$$-1 \leq \text{ACF}(k) \leq 1$$

- **Partial Autocorrelation Function (PACF)**

  This function is also used in time series to get the idea about the relationship between a time series variable at different lags, after removing the effects of previous lags. It is particularly useful in determining the order of AR models.

  The PACF at lag k is the correlation between X_t and X_{t-k} after controlling for the effects of the intermediate time points $X_{t-1}, X_{t-2}, \ldots, X_{t-(k-1)}$. Essentially, it gives the direct relationship between X_t and X_{t-k}, excluding the influence of the series values between them.

  Mathematically, the PACF at lag k, denoted ϕ_{kk}, can be expressed as:

$$\boldsymbol{\phi_{kk}} = \text{corr}(\boldsymbol{X_t, X_{t-k} \mid X_{t-1}, X_{t-2}, \ldots, X_{t-(k-1)}}),$$

  Where

$$\text{corr}(\boldsymbol{X_t, X_{t-k} \mid X_{t-1}, X_{t-2}, \ldots, X_{t-(k-1)}}),$$

  is the partial correlation of X_t and X_{t-k} after accounting for the influence of the values in between.

The relationship between AR models and MA models based on the ACF and PACF plots can be shown as follows.

Model Type	**ACF Behavior**	**PACF Behavior**	**Key Points**
AR(p) Model	Exponentially decaying, tails off slowly.	Significant at lags 1 through p, cuts off after lag p.	The ACF decays, and the PACF cuts off after lag p.
MA(q) Model	Cuts off after lag q.	Exponentially decaying, tails off slowly.	The ACF cuts off after lag q, and the PACF decays.
ARMA(p, q) Model	A mixture of exponential decay (from AR part) and cutoff (from MA part).	Significant at lags 1 through p, cuts off after lag p.	A combination of AR and MA characteristics, with ACF decaying and PACF cutting off.
White Noise	ACF is zero at all lags.	PACF is zero at all lags.	Both the ACF and PACF are zero for white noise because there's no correlation at any lag.

Table 2.1: Relationship between AR and MA models using ACF and PACF plots.

- **Stationarity and weakly stationarity**

  A time series is said to be stationary if there is no systematic change in mean, variance, covariance, and standard deviation that does not change over time. In other words, a stationary time series does not have trend or seasonal components.

  Weak stationarity, also simply known as stationarity or second-order stationarity, is a property of a time series where its mean and variance are constant over time, and the covariance between two time periods depends only on the time distance between them, not on the actual time at which the covariance is computed. This concept plays an important role in the field of time series analysis and econometrics because stationary processes can be easy to model and forecast.

Aspect	Strict Stationarity	Weak Stationarity
Definition	The entire distribution (all moments) of the series does not change over time.	Only the first two moments (mean and variance) and autocovariance are time-invariant.
Applicability	Requires invariance of higher-order moments, not practical for real-world data.	Commonly used and sufficient for most statistical and econometric models.
Moment Dependency	Considers all moments (mean, variance, skewness, kurtosis, etc.).	Considers only first and second moments (mean and variance).
Testing	Difficult to test directly in practice.	Easier to test using tools like ADF (Augmented Dickey-Fuller) or KPSS tests.

Table 2.2: Key differences between stationarity and weak stationarity.

Non-linear Exponential Curve Fitting

Exponential curve fitting is a mathematical method for modeling time series data that increases or decreases rapidly over time using an exponential curve, also known as a geometric curve. To fit an exponential curve to time series data, we typically model the expected value $y(t)$ of the data at time t using an exponential function.

General formula for an exponential smoothing,

$$\boldsymbol{y = a \cdot e^{b \cdot x} + c},$$

where

- y : Dependent variable.

- x : Independent variable.
- a, b, c : Parameters to be estimated.
  - a : Scaling factor.
  - b : Growth (or decay) rate.
  - c : Vertical shift.

Exponential smoothing

Exponential Smoothing measures give large weights for more reason observations, and the weights decrease exponentially as the observations become older. These methods are most effective when the parameters describing the time series are changing slowly over time.

There are three type of exponential smoothing,

1. Simple exponential smoothing.
   Simple exponential smoothing is used when the time series data does **not exhibit a trend or seasonality**. The forecast is based on a weighted average of past observations, with more recent observations given higher weights.

2. Holt's linear trend method (Double exponential smoothing)

   This method is also called as Holt's trend corrected or second-order exponential smoothing. This is a technique that works with data **having a trend but no seasonality**. In order to make predictions on the data, the Holt's Method uses two smoothing parameters, alpha and beta, which correspond to the level components and trend components and giving more weight to recent observations. Therefore, this method is often used in short-term forecasting.

   The forecast for the next time period,

$$\boldsymbol{\hat{y}_{t+h} = \ell_t + h \cdot b_t},$$

   where

   $\hat{y}_{t+h}$ = forecast for h-steps ahead,
   ℓ_t = smoothed level at time t,
   b_t = smoothed trend (slope) at time t,

h = forecast horizon.

The level and trend components are as follows,

$$\boldsymbol{\ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1})},$$

$$\boldsymbol{b_t = \beta(\ell_t - \ell_{t-1}) + (1 - \beta)b_{t-1}},$$

where

y_t is the actual value at time t,
α is the smoothing parameter for the level ($0 \leq \alpha \leq 1$),
β is the smoothing parameter for the trend ($0 \leq \beta \leq 1$).

The initial values are typically chosen as;

$$\boldsymbol{\ell_0 = y_1},$$

$$\boldsymbol{b_0 = y_2 - y_1},$$

3. Holt-Winters method (triple exponential smoothing).

   The Holt-Winters method is an extension of Holt's method that also includes a seasonality component, making it ideal for data **with both trend and seasonality**. It measures the level, trend, and seasonal fluctuations in the data and is suitable for periodic data.

2.3.2 Risk metrics

Shannon entropy

Shannon entropy is a concept from information theory, introduced by Claude Shannon in 1948. It is used to quantify the uncertainty or information content of a random variable or a source of information. In simple terms, it measures the amount of unpredictability or surprise in a set of outcomes.

Shannon entropy can also be applied to estimate the risk and the uncertainty of the financial market, which has a high level of volatility and a complex market structure, using entropy-based analysis. Shannon entropy considers price movements or returns as random , thus providing a clearer picture stability or non stability and predictability of the cryptocurrency market.

Formula for the Shannon entropy,

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_b p(x_i),$$

H(X) = Entropy of the random variable X.
= The probability of the i^{th} outcome of the random variable.
n = Number of possible outcomes in the system.

- **Higher entropy** : If all outcomes of X are equally likely, the entropy is maximized. This means there's more uncertainty or surprise about the outcome; that means there's a high risk in there.

- **Lower entropy** : If one outcome is far more likely than the others, the entropy is lower. This indicates less uncertainty because the outcome is predictable.It means that there is a low risk in there.

Stepwise calculation for Shannon entropy is shown below.

1. Daily log returns calculation.

   Log returns are calculated to measure the relative change in Bitcoin prices over consecutive days.

   $$R_t = \ln\left(\frac{\text{Price at time } t}{\text{Price at time } (t-1)}\right),$$

   $$R_t = \text{log return.}$$

2. Arrange log returns into bins.

   To calculate probabilities, the range of log returns is divided into 10 equal-width bins. This allows the creation of a frequency distribution of returns over a rolling window of 30 days.

3. Entropy calculation.

   By using the above Shannon entropy formula and probability value. We can calculate the entropy for each 30 days window.

4. Entropy calculation. To make entropy values comparable, the calculated entropy is normalized by dividing it by the maximum possible entropy for 10 bins, which is log(10).

   Normalized entropy H_o is,

$$H_o = \frac{H}{H_{\max}} = \frac{H}{\ln n}, \quad \text{where} \quad 0 \leq H_o \leq 1$$

   The maximum possible entropy is,

$$H_{\max} = \ln n$$

H_o	:	Normalized entropy value,
H	:	Shannon entropy,
$H_{\max}$	:	Maximum possible entropy,
n	:	Number of possible outcomes in the system.

Properties of Shannon entropy

- **Continuity:** H must be a continuous function of $p_X(x_i)$, where $i = 1, \ldots, n$.

- **Monotonicity:** If all probabilities are equal ($p_X(x_i) = \frac{1}{n}$), the uncertainty, H must increase as n increases.

- **Additivity:** If an option is divided into two successive options, the original H must equal the weighted sum of the uncertainties of the individual options.

Shannon demonstrated that any measure satisfying these properties can be scaled by a positive constant to define the unit of measurement. The key properties that make this uncertainty measure particularly useful are,

Key Properties

- $H(X) = 0$: This happens if and only if all probabilities $p_X(x_i)$ are zero except one.

- $H(X)$ is maximized when all probabilities are equal ($p_X(x_i) = 1/n$). In this case, $H(X) = \log(n)$, indicating maximum uncertainty.

- The joint entropy $H(X, Y)$ satisfies,

$$H(X, Y) \leq H(X) + H(Y),$$

  with equality if and only if X and Y are statistically independent, i.e.,

$$p(x_i, y_j) = p(x_i)p(y_j).$$

- Any shift toward equalizing the probabilities $p_X(x_i)$ will result in an increase in H (greater uncertainty).

- For joint events, the uncertainty satisfies:

$$H(X, Y) = H(X) + H(Y|X),$$

  meaning the total uncertainty of the joint event (X, Y) is the sum of the uncertainty of X and the conditional uncertainty of Y given X.

- $H(Y) \geq H(Y|X)$: This implies that knowledge of X never increases the uncertainty of Y. Instead, it reduces $H(Y)$, unless X and Y are independent, in which case the uncertainty remains unchanged.

Volatility

Volatility is a statistical measure of the dispersion of returns for a given financial asset, portfolio, or market index to measure the risk.

We can calculate standard deviation as below,

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(r_i - \bar{r})^2},$$

where

- σ: Volatility (standard deviation of returns),
- r_i: Individual returns,
- $\bar{r}$: Mean (average) of the returns, given by $\bar{r} = \frac{1}{n} \sum_{i=1}^{n} r_i$,
- n: Total number of observations.

After calculating the values, we can get an understanding of the risk as follows:

- **High volatility:** Indicates large fluctuations in price or returns, implying higher risk.
- **Low volatility:** Indicates smaller fluctuations, implying lower risk and more stability.

Value at Risk

statistical measure used in finance and risk management to assess the potential loss in value of an investment, portfolio, or financial asset over a specified time period, given normal market conditions and a certain level of confidence.

VaR is calculated using the following formula:

$$\text{VaR} = \mu - Z_\alpha \cdot \sigma,$$

where

μ = the mean (expected return) of the asset or portfolio.
Z_α = the Z-score corresponding to the desired confidence level (α).
σ = the standard deviation (volatility) of the asset or portfolio returns.

By using the VaR value, we can get a brief idea as follows:

Low VaR:

- Lower potential loss, typically associated with conservative, low-risk investments.
- Indicates a safer, more stable position with less volatility in the asset or portfolio.

High VaR:

- Higher potential loss, typically associated with more volatile and risky investments.
- Indicates a higher risk profile, reflecting assets with more price fluctuations and larger potential losses.

2.4 Bitcoin data

We collected data on Bitcoin prices from 15th September 2018 to 14th September 2024. And used daily Bitcoin closed prices for this study.
The R code used to read this data is in the appendix.

2.4.1 Metadata

The source of these data is from investing.com.
(`https://www.investing.com/`)

The data collection took place on 16th September 2024.

2.4.2 Data dictionary

Variable Name	Variable Description	Variable Type	Measurement Unit
Date	Day observed	Numerical	Calendar day
Price	Closed price of the day	Numerical	U.S. dollars
High	Highest price of the day	Numerical	U.S. dollars
Low	Lowest Price of the day	Numerical	U.S. dollars
Open	Opening price of the day	Numerical	U.S. dollars

Table 2.3: Data dictionary table

2.4.3 Preparation for analysis

Our data set is fully clean, but in other contexts, we will have to deal with numerical issues that are important in time series forecasting and entropy calculation depending on the programming language and libraries we are using; even a single NA value or null value can prevent models from running. So we made sure that there were no NA values and null values in the data set.

Chapter 3

Results

In this chapter we did both time series analysis and risk matrix calculation. Result as follows.

3.1 Time series analysis

To show the features of the particular time series in terms of the stability of the time series on average and variance, we draw the data against time, using the time series plot.

Figure 3.1shows the closing price of Bitcoin from 14th September 2018 to 15th September 2024. Fluctuations in the closing prices are clearly shown here. The price, which gradually increased from 2018 to mid-2020, again gradually decreased from mid-2021 to mid-2023, and then it rapidly increased again by the middle of the year 2023. The highest value went nearly 80,000. From 2020 to late 2021, the closing price fluctuated significantly due to the impact of the COVID-19 pandemic, so it still showed an overall upward trend.

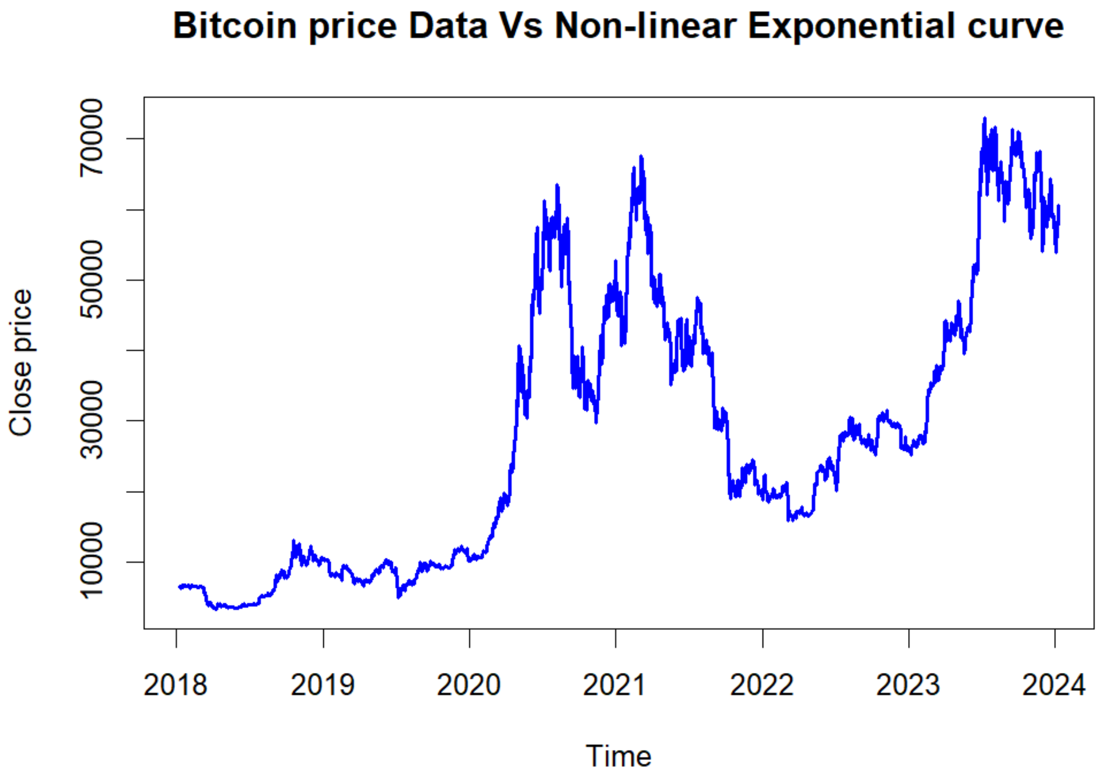

Figure 3.1: Time series plot of Bitcoin closing prices

3.1.1 Non-linear exponential curve

We identified from the time series plot that the Bitcoin data shows an overall upward trend. An exponential curve is also an upward curve. So, I used the exponential curve fitting technique here, taking the time period as the independent variable and the exponential values of the daily stock prices as the dependent variable.

$$\text{Bitcoin\_price\_ts} \sim a \cdot \exp(b \cdot \text{time\_index})$$

Parameter	**Estimate**	**Std. Error**	**t value**	**Pr(>|t|)**
a	1.100×10^4	359.9	30.56	$< 2 \times 10^{-16}$
b	7.516×10^{-4}	0.00001939	38.76	$< 2 \times 10^{-16}$

Table 3.1: Curve fitting results

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 Residual standard error: 14040 on 2190 degrees of freedom.
Number of iterations to convergence: 65.
Achieved convergence tolerance: 6.495×10^{-6}.

By using above result,

$$y = a \cdot e^{b \cdot x}$$

$$y = 11000 \cdot e^{0.0008 \cdot x}$$

where

y : Bitcoin price at time t.
x : time.

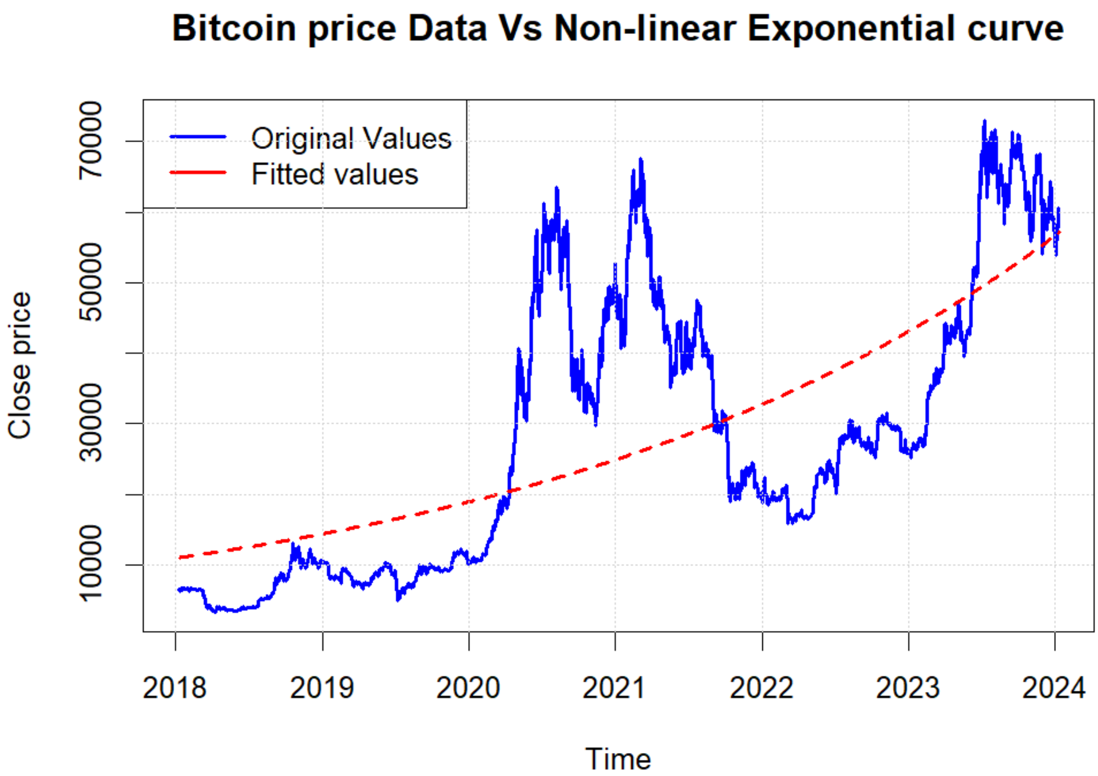

Figure 3.2: Exponential Curve Vs Bitcoin prices

The above figure 3.2 shows the non-linear curve fitting and Bitcoin price data on the same time axis.

3.1.2 Double exponential smoothing

This is a technique that works with data that have a trend but no seasonality. I used the decompose command in the R software to identify the components of the Bitcoin price.

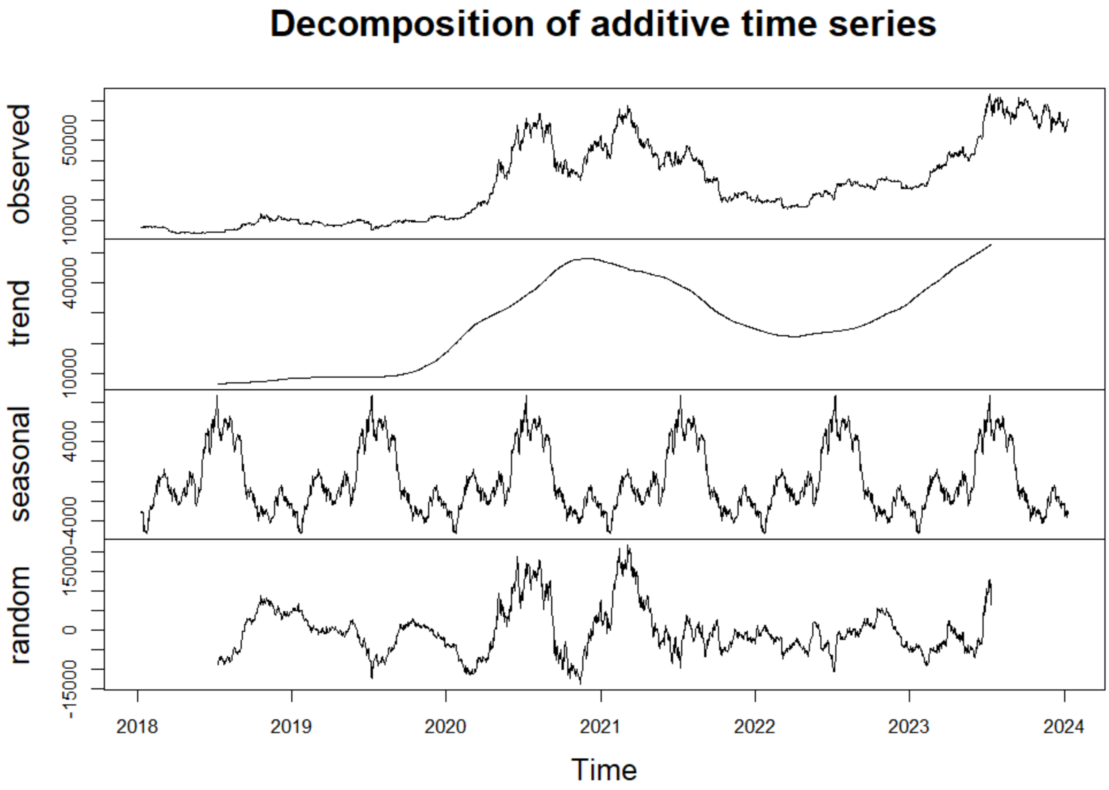

Figure 3.3: Decomposed Bitcoin price data

As observed in Figure 3.3, the dataset exhibits a trend. To formally verify this, I conducted the Mann-Kendall test, which confirmed the presence of a trend in the data. However, this test does not provide insights into the seasonal component. Therefore, to assess seasonality, I applied the isSeasonal() function from the seastest package. The results indicate that the original dataset does not exhibit a significant seasonal pattern.

To perform double-exponential smoothing, we used the holt() function. Here is the summary of this model fitting:

Forecast method: Holt's method

Model information:
Holt's method

Call:
`holt(y = Bitcoin_price_ts, h = 30)`

Smoothing parameters:

- $\alpha = 0.9434$
- $\beta = 1 \times 10^{-4}$

Initial states:

- $l = 6448.8917$
- $b = 35.0029$

sigma: 1123.738

AIC, AICc, BIC:

AIC:	47663.15
AICc:	47663.18
BIC:	47691.61

Error measures:

ME:	-9.43458	RMSE:	1122.712	MAE:	651.0441	MPE:	-0.2124811
MAPE:	2.337721	MASE:	0.03320922			ACF1:	-0.001494951

By using above result,
Smoothing Equation:

$$\ell_t = \alpha y_t + (1 - \alpha)(\ell_{t-1} + b_{t-1})$$

$$\ell_t = 0.9434 y_t + 0.0566(\ell_{t-1} + b_{t-1})$$

Trend Equation :

$$b_t = \beta(\ell_t - \ell_{t-1}) + (1 - \beta)b_{t-1}$$

$$b_t = 0.0001(\ell_t - \ell_{t-1}) + 0.9999b_{t-1}$$

Forecasted Equation :

$$\hat{y}_{t+h} = \ell_t + h \cdot b_t$$

$$\hat{y}_{t+h} = 6448.8917 + 35.0029h$$

$h = 1, 2, 3, \ldots$

In here,

- **Alpha (level smoothing):** $\alpha = 0.9434$
  Indicates that the model places a high weight on recent observations when smoothing the level.
- **Beta (trend smoothing):** $\beta = 0.0001$
  Suggests minimal smoothing of the trend, meaning the trend component is very stable over time.

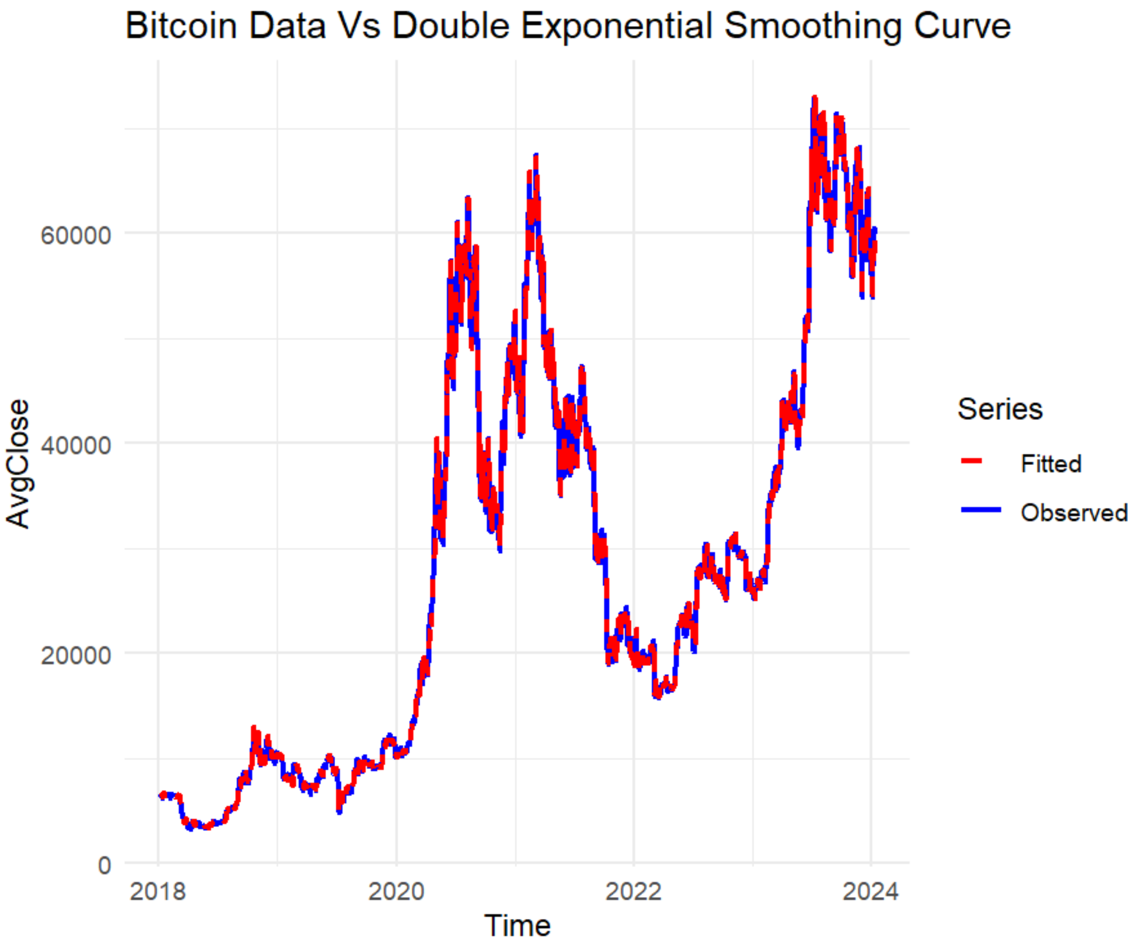

Figure 3.4: Bitcoin data vs Double Exponential Curve

The above figure 3.4 shows the fitting of the double exponential curve and the Bitcoin data on the same time axis. In this graph we can see that fitted values are somewhat exactly matched with the real-time bitcoin data set. Further, we checked AIC, BIC, MSE, and MAPE.

Metric	**Value**
AIC	4.766315×10^4
BIC	4.769161×10^4
MSE	1.260483×10^6
MAPE	2.337721

Table 3.2: AIC, BIC, MSE and MAPE values

According to the above table 3.2, we can see that the Akaike Information Criterion (AIC) is $47,663.15$, and the Bayesian Information Criterion (BIC) is $47,691.61$. These values are relatively close, indicating that the model achieves a reasonable balance between fit and complexity. However, the absolute values suggest the model may not be highly parsimonious, as both metrics penalize overly complex models.

Mean Squared Error (MSE) of $1,260,483$ reflects the average squared difference between actual and predicted values. While this is a large number, its acceptability depends on the scale of the data.

Mean Absolute Percentage Error (MAPE) of 2.34% indicates that the model's predictions are, on average, only 2.34% off from the actual values. This demonstrates strong predictive performance, especially for financial forecasting.

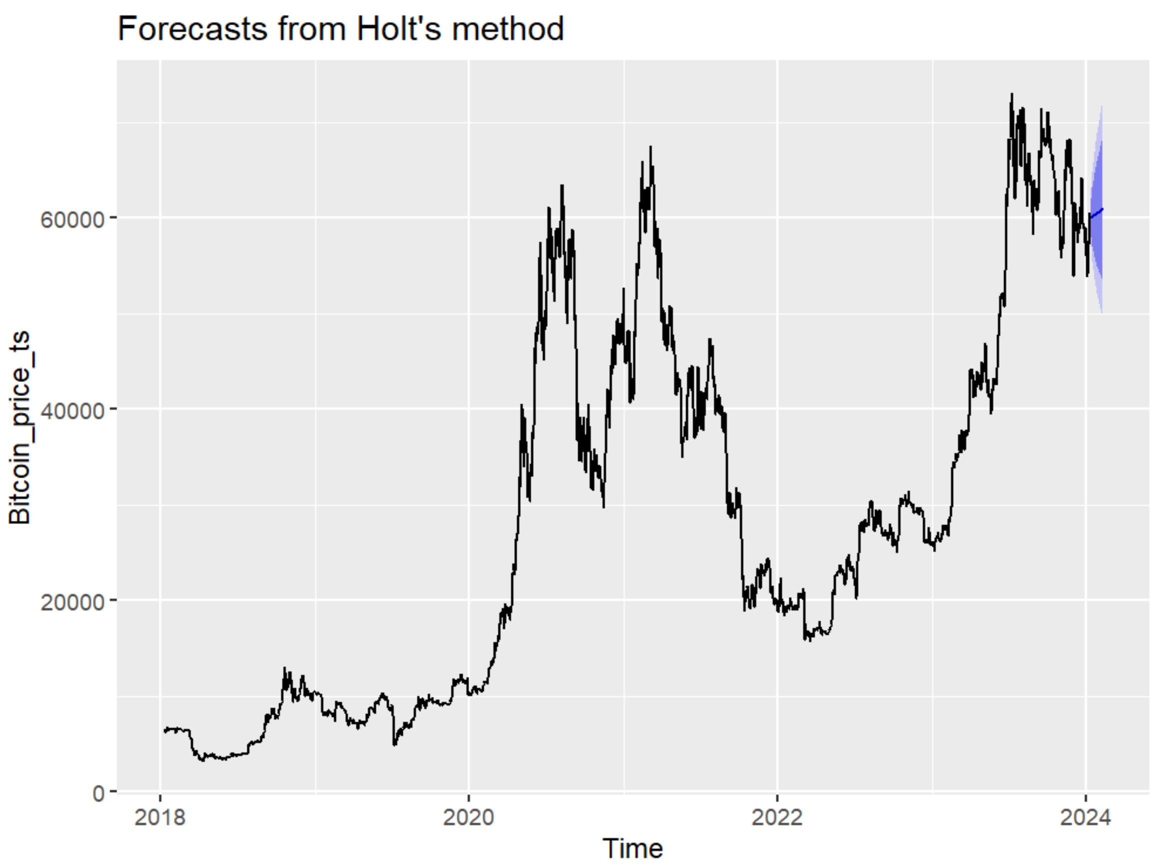

Figure 3.5: Forecasted graph

Above figure 3.5 The blue line represents the point forecasts (predicted Bitcoin prices) for the near future.The shaded blue area represents the confidence intervals, which indicate the uncertainty around the predictions. The wider the interval, the greater the uncertainty. So the forecasted values are consistent with the historical trend and suggest a continuation of the current price level, with moderate volatility.

3.2 Risk metrics

Risk metrics are essential in finance, economics, and decision-making for several reasons. So in here we used Shannon entropy, volatility, and VaR to quantify the risk in the bitcoin market.

3.2.1 Shannon entropy

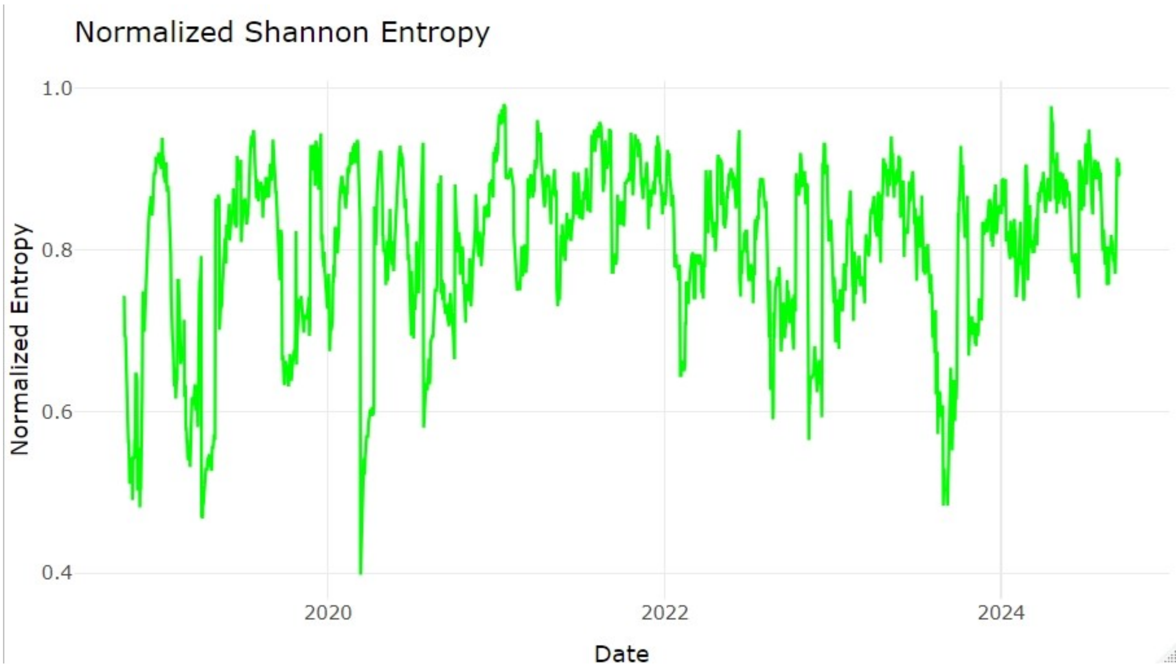

Figure 3.6: Normalized Shannon Entropy Graph

This figure **??** shows the normalized Shannon entropy of bitcoin in 2018-2024. This analysis was done by using Appendix B. This shows variations in market uncertainty of Bitcoin over time, and values vary from 0.4 to 1.0. Normalized entropy is the measure of unpredictability of the daily log returns. Thus, the greater the value of entropy indicate the more uncertain the behavior of the market.

On August 24, 2022, the normalized entropy was around 0.5904, indicating a moderate level of uncertainty in the market. When the entropy approaches 1.0, it signifies highly random price movements, reflecting elevated uncertainty and risk. Conversely, when the entropy decreases toward 0.4, the market becomes more predictable, indicating reduced uncertainty and greater stability.

The plot shows repeated increases and decreases in entropy, which align with major market events or changes in investor behavior. High entropy levels often occur during volatile market periods, serving as warnings for investors. Notable peaks in entropy are visible at the end of 2020 and in 2021, likely linked to significant events in the cryptocurrency market, such as large price swings or important regulatory announcements.

3.2.2 Volatility

In this study, volatility is measured as the rolling standard deviation of daily log returns over a 30-day period. This approach tracks how price fluctuations change over time, showing periods of higher or lower market activity. High volatility means bigger price swings and higher risk, while low volatility indicates more stable price behavior. (The code used for analysis is in Appendix B.).

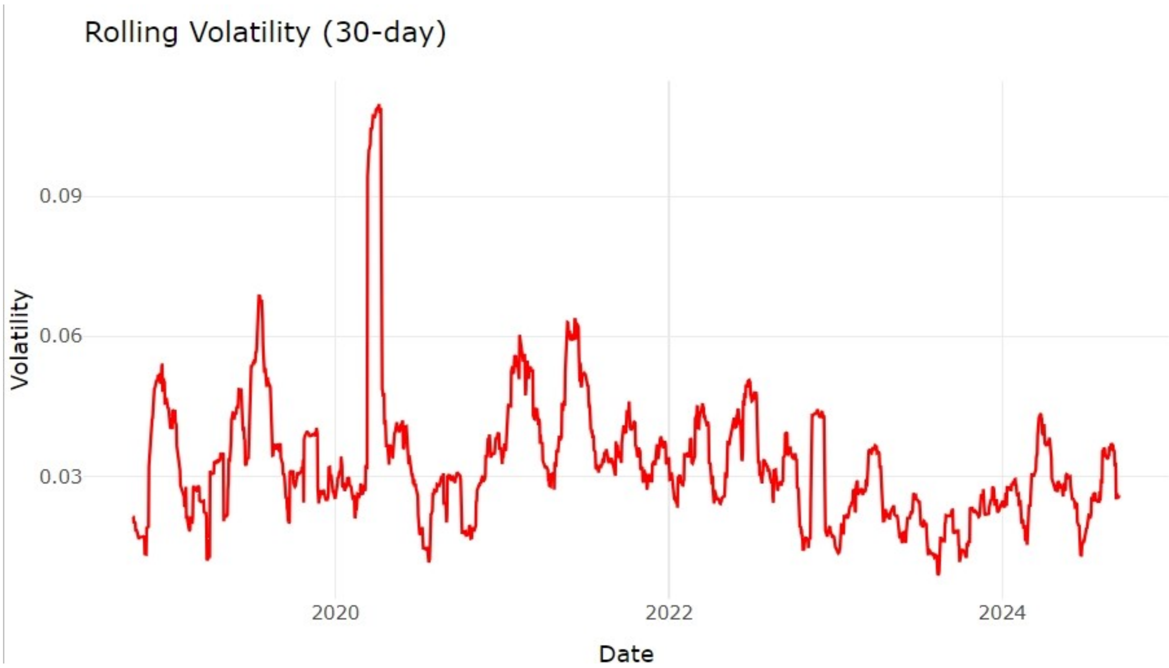

Figure 3.7: Rolling volatility graph

The 30-day rolling volatility captures Bitcoin's changing risk levels based on its daily log return. A noticeable spike occurred in early 2020, reaching levels above 0.09, at the time of the global COVID-19 pandemic. This period was marked by high market uncertainty and significant price fluctuations across all asset classes, including bitcoin prices. Similarly, the volatility of Bitcoin reached its maximum during late 2021. At the time, major market events happened, highlighting the strong impact of external factors on the Bitcoin market.

During 2022 and 2023, volatility remained relatively low, ranging between 0.02 and 0.05, indicating reduced market risk compared to earlier periods. The rolling volatility also highlights periods of clustering, where phases of high volatility are followed by more stable periods. This underscores the importance of volatility as a key market risk indicator, helping traders and investors easily understand uncertainties in the bitcoin market more effectively.

3.2.3 Value at Risk (VaR)

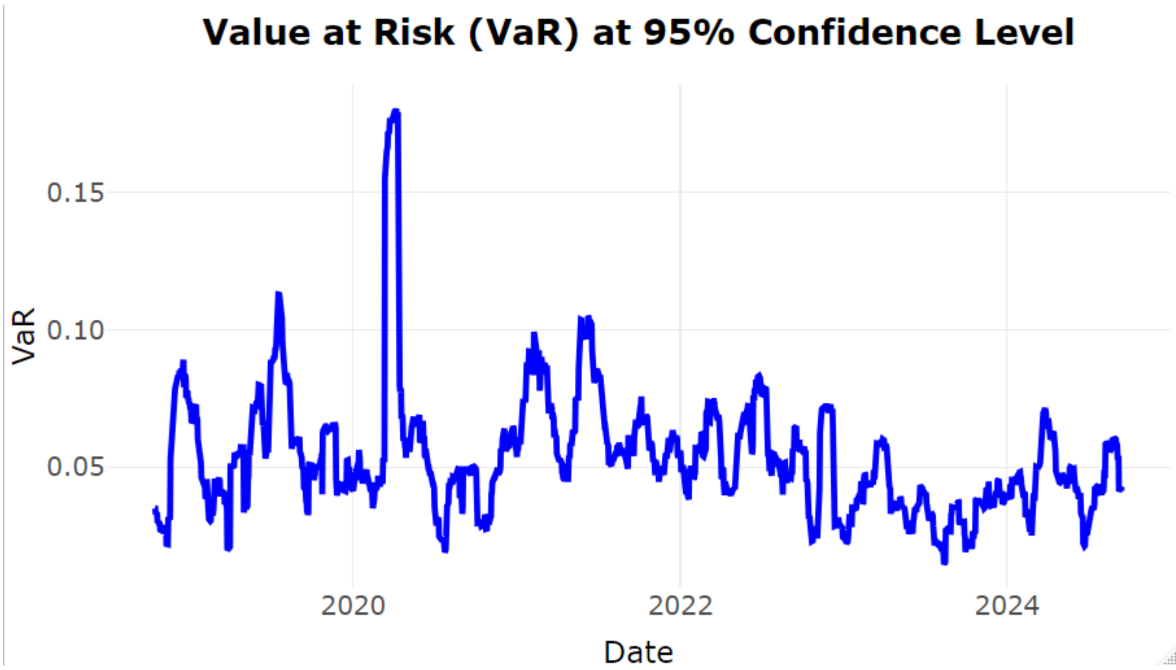

Figure 3.8: VaR graph

The graph 3.8 shows the quantified maximum possible loss over a certain period, as obtained from historical data. It is observed that there is a great deal of variation in the VaR values over time. For example, in 2020, VaR peaked at about 0.15, meaning an investor should expect a loss of as high as 15% of the value of the portfolio during this period. These spikes usually indicate increased market volatility and risk, possibly as a consequence of events in the global economy or other factors related to the bitcoin market.

During stable periods like late 2023 and early 2024, VaR values remained consistently low, ranging from 0.03 to 0.05, indicating reduced market uncertainty. These changes in VaR highlight the volatile nature of the bitcoin market, where conditions can shift rapidly. The plot shows the importance of tracking VaR during high-risk periods to make well-informed decisions about risk management and portfolio adjustments.

3.2.4 Volatility vs VaR

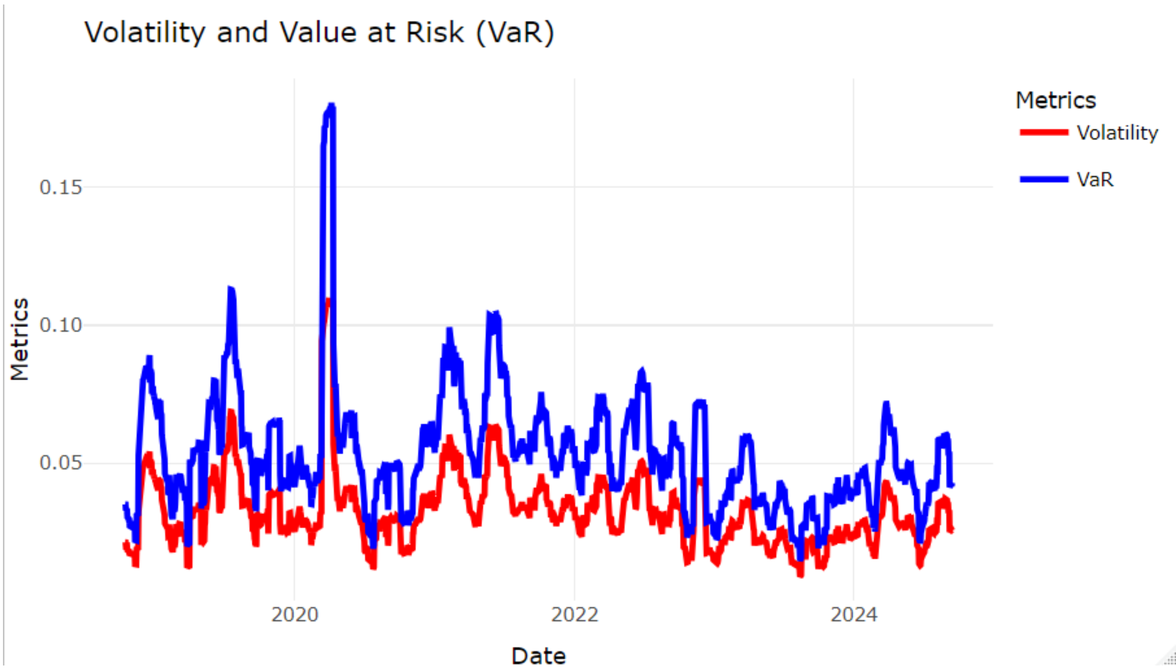

Figure 3.9: Volatility vs VaR

The relationship between Bitcoin's volatility and Value at Risk (VaR) develops over time according to the data presented in Figure 3.9. As volatility hits high levels during early parts of 2020. Bitcoin's volatility goes above 0.1 while VaR exceeds 0.15 during the time when market risk increases due to pandemic panic and market uncertainty caused by COVID-19. Market volatility correlates directly with higher potential loss amounts. Both volatility and VaR experienced declines in the stable market conditions that occurred between late 2022 and early 2023. Market uncertainty decreased along with financial risk while volatility maintained levels between 0.02 and 0.05 within this period.

In here we can see how both of the measures help in understanding Bitcoin's risk. Volatility highlights how much the market is fluctuating, while VaR turns that fluctuation into a clear measure to quantify the risk. Together, they give investors a clearer picture of market risk, helping them make smarter decisions and allowing traders to adjust their strategies based on their risk.

3.2.5 Volatility vs VaR vs Shannon entropy

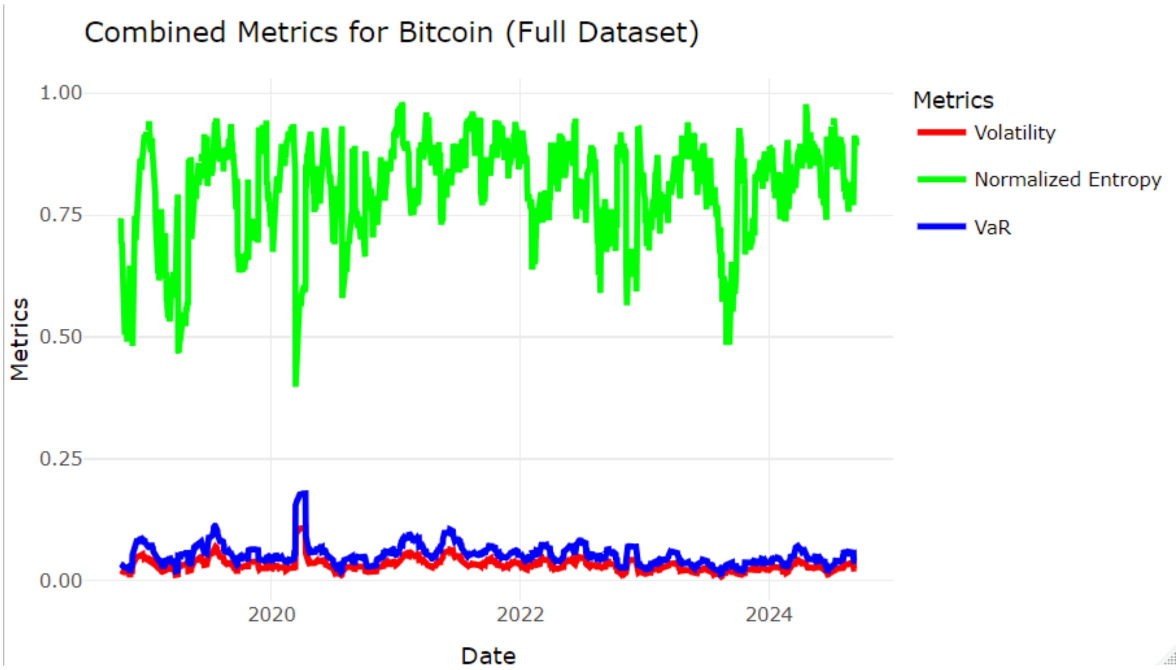

Figure 3.10: Volatility vs VaR vs Shannon entropy

The graph 3.4 shows how entropy, volatility, and VaR changed over time. Entropy stayed high, between 0.6 and 1.0, showing that Bitcoin's price movements have been uncertain. Volatility (the red line) and VaR (the blue line) are closely connected, with both spiking during unstable times, like the start of 2020. For example, during the COVID-19 crisis, volatility went over 0.1, VaR rose above 0.15, and entropy reached close to 1.0, indicating maximum market uncertainty.

During the periods from late 2022 to early 2023, entropy stayed between 0.75 and 0.8, while both volatility and VaR were low (below 0.05). This shows lower financial risks and less market instability. However, the relatively high entropy values still suggest that there's some level of unpredictability in Bitcoin's market.

Chapter 4

Discussion and Conclusion

4.1 Discussion

This study analyzed Bitcoin's price trends and market risk from 2018 to 2024 using time series forecasting and risk assessment metrics. The research focused on the relationship between volatility, Value at Risk (VaR), and Shannon entropy to quantify market uncertainty.

The findings indicate that double exponential smoothing is a suitable time series analysis for forecasting the data, and it is fitted with a MAPE value of 2.34% with the best alpha at 0.94. Bitcoin exhibits significant volatility, with pronounced fluctuations during periods of global financial instability. For instance, during early 2020, the COVID-19 pandemic captures extreme market movements, leading to high volatility (above 0.1) and elevated VaR (exceeding 0.15). This suggests that external macroeconomic factors heavily influence Bitcoin's price behavior. Conversely, in more stable periods such as late 2022 to early 2023, both volatility and VaR declined to below 0.05, reflecting reduced market uncertainty. However, entropy values remained relatively high (0.75–0.8), indicating that some degree of unpredictability in the Bitcoin market prices.

The integration of Shannon entropy as a risk measure provided deeper insights into market unpredictability. While volatility and VaR capture price fluctuations and potential financial losses, entropy quantifies the randomness of price movements. The results confirm that high entropy values correspond to volatile periods, highlighting the need for alternative risk measures beyond traditional financial models.

4.2 Conclusion

As a conclusion, we can say that double exponential smoothing is a suitable time series method to analyze the trend pattern and forecast accurately within the 95% confidence interval. Moreover. This study indicates that Bitcoin's market is characterized by high volatility and unpredictability, making traditional risk assessment methods insufficient. The use of Shannon entropy, in combination with volatility and VaR, provides a more comprehensive risk evaluation framework. So we can conclude that the combination of the above 3 risk measures gave valuable insight into risk assessment and was suitable for bitcoin analysis. Moreover, according to the result, we can conclude that,

- Bitcoin's price movements are highly volatile during global economic disruptions, leading to significant financial risks.

- VaR and volatility are strongly correlated that higher price fluctuations lead to greater potential losses as well as lower price fluctuation lead to smaller potential losses.

- Shannon entropy remains elevated, even in stable periods, indicating persistent uncertainty in Bitcoin's market.
- Alternative risk assessment techniques, such as entropy, enhance market risk evaluation, offering investors better tools for decision-making.

Analysis of bitcoin prices requires the implementation of advanced statistical measures, according to this research. Research should focus on developing complex forecasting tools along with new risk indicators to strengthen digital asset management practices in the evolving cryptocurrency environment.

Appendix A

The R codes that are used for time series double exponential smoothing are as follows:

```
library(TSA)
library(tseries)
library(lmtest)
library(forecast)
library(ggplot2)
library(FinTS)

# attach data file
bitcoin<- attach(Bitcoin_price)



#Checking for obvious errors
which(is.na( bitcoin))

bitcoin$Price<-
  as.numeric(as.character(gsub(",","",bitcoin$Price)))

#Converting the data set into time series object
Bitcoin_price_ts<- ts(as.vector(bitcoin$Price),
                      start=c(2018,9,15),frequency = 365)
plot.ts(Bitcoin_price_ts,
        main="Time series plot of Bitcoin Prices")



# Decompose the data
decomposed_data <- decompose(Bitcoin_price_ts)
plot(decomposed_data)

#check seasonality
isSeasonal(Bitcoin_price_ts,freq = 365)

# Perform Mann-Kendall test(check trend)
library(trend)
mk_test <- mk.test(Bitcoin_price_ts)
```

```
# Display the result from mk test
print(mk_test)

# Apply Holt's method (double exponential smoothing)
DE_model <- holt(Bitcoin_price_ts,h=30)
DE_model
summary(DE_model)

#fit  the model
fitteddes <- fitted(DE_model)

observed_data <- data.frame(Time = as.numeric(time(tsObj))
                            , Values = as.numeric(tsObj))
fitted_data <- data.frame(Time = as.numeric(time(tsObj)),
                          Values = as.numeric(fitteddes))

#plot the orginal  data with fitted  value
ggplot() +
  geom_line(data = observed_data,
            aes(x = Time, y = Values,
                color = "Observed"),
            linewidth = 1) +
  geom_line(data = fitted_data,
            aes(x = Time, y = Values,
                color = "Fitted"), linewidth = 1,
            linetype = "dashed") +
  scale_color_manual(name = "Series",
                     values = c("Observed" = "blue",
                                "Fitted" = "red")) +
  ggtitle("Bitcoin Data Vs Double Exponential Smoothing Curve") +
  xlab("Time") +
  ylab("AvgClose") +
  theme_minimal()

#check AIC,BIC,MSE and MAPE
aic_des <- AIC(DE_model$model)
bic_des <- BIC(DE_model$model)
mse_des <- mean((DE_model$residuals)^2,
                na.rm = TRUE)
```

```
mape_des <- mean(abs((tsObj - fitteddes) / tsObj),
                 na.rm = TRUE) * 100

des_comparoson <- data.frame(
  Metric = c("AIC", "BIC", "MSE", "MAPE"),
  Value = c(aic_des, bic_des, mse_des, mape_des)
)

print(des_comparoson)

#forecast  the data
forecast_data <-forecast(DE_model,h=40)
autoplot(DE_model)



#plot foretasted value
library(seastests)
par(mfrow=c(1,1))
isSeasonal(tsObj,freq =365)
plot(DE_model,
     main = "Time Series with Double Exponential Smoothing",
     xlab = "Time", ylab = "Values",
     col = "blue", lwd=2)

# Add Fitted Values from Holt-Winters Model
lines(DE_model$fitted, col = "red",
      lwd = 2, lty=2)

# Add Legend
legend("topleft",
       legend = c("Original Data", "Fitted (Holt-Winters)"),
       col = c("blue", "red"),
       lty = c(1,2),lwd=2)
```

Appendix B

```
# Load necessary libraries
```

```
library(dplyr)
library(tidyr)
library(entropy)  # For entropy calculations
library(zoo)      # For rolling calculations
library(ggplot2)
library(plotly)   # For interactive plots

# Replace 'Bitcoin_full_data' with your
actual data frame or file path.
data <- Bitcoin_full_data  # Ensure this is loaded in your R environment

# Convert Date column to Date type and arrange by date
data <- data %>%
  mutate(Date = as.Date(Date)) %>%
  arrange(Date)

# Calculate daily log returns
data <- data %>%
  mutate(Log_Returns = log(Price / lag(Price)))

# Calculate rolling volatility (e.g., 30-day rolling standard deviation)
data <- data %>%
  mutate(Volatility = zoo::rollapply(Log_Returns, width = 30,
FUN = sd, fill = NA, align = "right"))

# Calculate Shannon entropy for daily log returns
data <- data %>%
  mutate(Entropy = sapply(seq_along(Log_Returns),
function(i) {
    if (i < 30) return(NA)
    subset_returns <- Log_Returns[(i - 29):i]
    subset_returns <- subset_returns[!is.na(subset_returns)]
    freq <- table(cut(subset_returns, breaks = 10))
    entropy::entropy(freq)
  }))

# Normalize entropy (relative to max entropy)
max_entropy <- log(10)  # Max entropy for 10 bins
data <- data %>%
```

```
  mutate(Normalized_Entropy = Entropy / max_entropy)

# Calculate Value at Risk (VaR) at 95% confidence level
confidence_level <- 0.95
z_score <- qnorm(1 - confidence_level)
data <- data %>%
  mutate(VaR = -z_score * Volatility)

# Remove rows with NA values
data <- data %>% drop_na()

# Save the results to a CSV file
write.csv(data, "bitcoin_full_analysis.csv", row.names = FALSE)

# ----- Separate Plots -----
# Plot 1: Daily Log Returns
plot1 <- ggplot(data, aes(x = Date, y = Log_Returns)) +
  geom_line(color = "black") +
  labs(title = "Daily Log Returns", x = "Date",
 y = "Log Returns") +
  theme_minimal()

# Plot 2: Rolling Volatility
plot2 <- ggplot(data, aes(x = Date, y = Volatility)) +
  geom_line(color = "red") +
  labs(title = "Rolling Volatility (30-day)",
 x = "Date", y = "Volatility") +
  theme_minimal()

# Plot 3: Normalized Shannon Entropy
plot3 <- ggplot(data, aes(x = Date, y = Normalized_Entropy)) +
  geom_line(color = "green") +
  labs(title = "Normalized Shannon Entropy",
 x = "Date", y = "Normalized Entropy") +
  theme_minimal()

# Plot: Value at Risk (VaR)
var_plot <- ggplot(data, aes(x = Date, y = VaR)) +
  geom_line(color = "blue", size = 1) +
```

```
  labs(
    title = "Value at Risk (VaR) at 95% Confidence Level",
    x = "Date",
    y = "VaR"
  ) +
  theme_minimal() +
  theme(
    plot.title = element_text(hjust = 0.5,
size = 16, face = "bold"),
    axis.title = element_text(size = 14),
    axis.text = element_text(size = 12),
    legend.position = "none"
  )

# ----- Interactive Plots with Plotly -----
# Interactive Log Returns Plot
plotly::ggplotly(plot1)

# Interactive Rolling Volatility Plot
plotly::ggplotly(plot2)

# Interactive Normalized Entropy Plot
plotly::ggplotly(plot3)

# Render the plot interactively using Plotly
interactive_var_plot <- plotly::ggplotly(var_plot)

# Print the interactive plot
interactive_var_plot




# Combined Plot: Volatility and VaR
volatility_var_plot <- ggplot(data, aes(x = Date)) +
  geom_line(aes(y = Volatility,
 color = "Volatility"), size = 1) +
  geom_line(aes(y = VaR, color = "VaR"), size = 1) +
  labs(
```

```
    title = "Volatility and Value at Risk (VaR)",
    x = "Date",
    y = "Metrics"
  ) +
  scale_color_manual(
    values = c("blue", "red"),
    name = "Metrics",
    labels = c("Volatility", "VaR")
  ) +
  theme_minimal() +
  theme(
    legend.position = "bottom",
    legend.title = element_blank()
  )

# Render the combined plot interactively using Plotly
plotly::ggplotly(volatility_var_plot)


# Combined plot with entropy,volatility,VaR
combined_plot <- ggplot(data, aes(x = Date)) +
  # Volatility
  geom_line(aes(y = Volatility,
 color = "Volatility"), size = 1) +
  # Normalized Entropy
  geom_line(aes(y = Normalized_Entropy,
color = "Normalized Entropy"), size = 1) +
  # VaR
  geom_line(aes(y = VaR,
 color = "VaR"), size = 1) +
  labs(
    title = "Combined Metrics for Bitcoin (Full Dataset)",
    x = "Date",
    y = "Metrics (Raw Values)"
  ) +
  scale_color_manual(
    values = c("green", "blue", "red"),
    name = "Metrics",
    labels = c("Volatility", "Normalized Entropy", "VaR")
```

```
  ) +
  theme_minimal() +
  theme(
    legend.position = "bottom",
    legend.title = element_blank()
  ) +
  # Set y-axis breaks for the primary y-axis
  scale_y_continuous(
    name = "Metrics",
    breaks = seq(0, 1, by = 0.25),  # Custom breaks at 0, 0.25, 0.5, 0.75, 1
    sec.axis = sec_axis(~ .,
name = "Metrics (Alternative Scale)")
  )

# Render the combined plot interactively
interactive_combined_plot <- plotly::ggplotly(combined_plot)

# Print the interactive combined plot
interactive_combined_plot
```

Table of Contents

Bibliography

Baharaeen, S. and Masud, A. S. (1986). A computer program for time series forecasting using single and double exponential smoothing techniques. *Computers & Industrial Engineering*, 11(1-4):151–155.

Huang, D., Schlag, C., Shaliastovich, I., and Thimme, J. (2019). Volatility-of-volatility risk. *Journal of Financial and Quantitative Analysis*, 54(6):2423–2452.

Ormos, M. and Zibriczky, D. (2014). Entropy-based financial asset pricing. *PloS one*, 9(12):e115742.