2022 Off-Cycle Internship Report

Value-at-Risk on Decentralised Finance: Methodologies and Use Cases

Submitted by

Amy Oumayma KHALDOUN

Master 2 Probability & Finance
Sorbonne University & Ecole Polytechnique

Under the guidance of

Anne-Claire Maurice & Théo Lafitte

Abstract

Value-At-Risk plays a central and prominent role in risk management and there exist several methods to estimate it. From historical simulation to GARCH models and variance-covariance methods and so on. In fact, the report provides overall results of each approach and for different mono-asset and multi-asset portfolios using a range of crypto-currencies and exchanges. To check the goodness of fit of the approaches, backtesting methods are used. On the other hand, we conclude the report with different use cases of the VaR to hedge one's portfolios and minimise the losses.

Contents

Acknowledgments

It is a genuine pleasure to express my warmest appreciation and acknowledgment towards my team members: Anne-Claire, Emmanuel Gobet, Mnacho Echenim and Théo Lafitte who made this internship possible. Their guidance, advice, persistent help and encouragement guided me through all the stages of my work. I learned a lot from them and I will never thank them enough for the time they dedicated to me.

In addition, a big thank you to my colleagues: Iliass Bouchriha, Axel Battut, Desi Nedelcheva, Anastasia Melachrinos, Mathieu Jalvé, Bediss Cherif, Nicolas Bazire and all the Kaikoes for making every moment at Kaiko delightful and extremely pleasant, every conversation engaging and captivating and every lunch enjoyable.

Moreover, it is my privilege to thank the CEO of Kaiko Ambre Soubiran, for dedicating time to me to get to know her and get to learn all about the Kaiko's history. She is truly and utterly an inspirational woman who's been building along with the Kaikoers a massive and successful crypto empire.

Last but not least, I'd like to thank my mother for her love, her sacrifices and relentless support. Without her I wouldn't be able to achieve my goals and my dreams.

Chapter 1

Notations

- V_t is the value of the whole portfolio whether it's a mono-asset or a multi-asset one.
- P_t^i is the value of the asset i in the portfolio at time t.
- δ_t^i is the quantity of the asset i at time t.
- w_t^i is the weight of the asset i at time t.
- $P\&L_t = V_t - V_{t-1}$.
- $X_{t+1} = -P\&L_{t+1} = V_t - V_{t+1}$ represents the loss of a portfolio between time t and $t+1$.
- $\mathcal{F}_t$ is a σ-algebra on V_t, $\mathcal{F}_t = \sigma(P_0^i, ..., P_t^i, 1 \leq i \leq n)$ where n is the number of the assets in the porfolio. It holds the information available at time t.
- Y_t is a variable that designates a type of returns between $t-1$ and t, see (2.8).
- $q_t^\beta(Y_{t+1})$ is a β-quantile of the return Y_{t+1} conditional to $\mathcal{F}_t$:

  $$q_t^\beta(Y_{t+1}) = -q_t^{1-\beta}(-Y_{t+1})$$

  where q_t^β is defined in (2.4).
- VaR_t^α is the Value-at-Risk of the portfolio loss at risk level $\alpha \in (0,1)$, conditionally to the information available at time t: this is defined by

  $$VaR_t^\alpha(X_{t+1}) := q_t^\alpha(X_{t+1}) = q_t^\alpha(V_t - V_{t+1}).$$

  We will typically consider $\alpha = 95\%$ so that we expect the quantity to be positive.
- Z_t is the innovation process.
- $\mu_{t+1|t}$ is the conditional mean of the returns (specified beforehand) at time t, i.e. $\mu_{t+1|t} = \mathbb{E}(Y_{t+1}|\mathcal{F}_t)$
- $\sigma_{t+1|t}$ is the conditional volatility of the returns (specified beforehand) at time t, i.e. $\sigma_{t+1|t} = \mathbb{V}(Y_{t+1}|\mathcal{F}_t)$.

Chapter 2

Introduction

The question why to choose to do an internship at the crypto company named Kaiko instead of a traditional one has been asked several times especially when we think about the whole controversy and uncertainty around the Crypto world. I'll begin answering that with an introduction of Kaiko and eventually a little story of money.

2.1 Kaiko

Kaiko is the leading source of cryptocurrency market data, providing businesses with industrial grade and regulatory compliant data. It empowers market participants, institutional investors, leading academic institutions, as well as enterprises in the digital finance industry, like Chainlink, Bank of Canada, Bloomberg, Uniswap, Deutsche Borse Group with global connectivity to real-time and historical data feeds across the world's leading centralized and decentralized cryptocurrency exchanges. Kaiko's mission is to bridge traditional and blockchain ecosystems by providing reliable and actionable financial data and services. its proprietary products range from portfolio valuation to strategy backtesting, performance reporting, charting, analysis, indices, pre- and post-trade.

Kaiko covers over 100 crypto centralised and decentralised exchanges and have over 10 years of historical market data. Their data ranges between trade data, to order book data and quantitative data.

As a quantitative analyst, I've worked for six month within the analytics department with a team of three people: Anne-Claire Maurice, Emmanuel Gobet Mnacho Echenim and Theo Lafitte.

2.2 Why The Crypto World ?

This section is purely the opinion of the author and is independant from the internship topic and does not reflect the views of Kaiko.

Money has always been the center of the human story with its several functions in different communities and has been represented in different material forms in the past 10000 years, going from grains to cattle and metal coins to seashell and paper or just numbers in your bank account. The concept of money has gone through several big events too: we can recall the period when it was deriving its value from a commodity to the Nixon Shock in 1971. One characteristic that seemed quite dominant from the late 17th century with the monopolisation in ancient Rome and the birth of the Bank of England up to now is the centralisation of money. In other words, the existence of a central authority that governs money and decides how everything related to it is managed.

There are obviously many advantages to this monetary system, among those, it's worth mentioning the effective control and supervision. Also, the fact that every job in the organization that requires specialists can be trained thanks to centralized office organization, which also helps to standardize work. Naturally, when operations are centralized, they will either be in the hands of one person or a group of people, but they will still be directly under his or her supervision. As a result, activities will be more consistent, resulting in consistent decision-making and processes which also results in no duplication of work.

However, there's a very controversial political philosophy called Libertanism which seeks to maximize autonomy and political freedom, and minimize the state's encroachment on and violations of individual liberties. In fact, libertarians have a deep distrust of centralised state power and try to limit the government interventions since they believe that centralised financial institutions are the major cause of the greatest economic crisis along with the inflation that is considered to be the result of the government mismanagement of the supply of money and a big torture to humanity. As a consequence, we have witnessed a significant amount of revolutionary technologies such as the Internet whose infrastructure is completely different from anything invented before. This latter exists in both the real world and a separate realm within cables and electricity which removed all communication barriers, creating an open world with easy, instant and endless communication.

According to the Libertanians, there are some drawbacks that are noteworthy and that constitute the base for a new technology to emerge. As a matter of fact, the money centralisation destroys individual initiative since it revolves around one small entity. In other words, one man takes all the decisions and decides the modes of implementing them despite the glaring fallacies. Adding the fact that this type of decision making triggers some kind of distance from the customers or the people in general which limits creativity and communication. Furthermore, the fact that it gives all responsibilities to few people in the organization causes them to remain over-burdened with routine work and slow down the operations which is not the most time-optimised system. Besides, we can mention the lack of secrecy. In fact, in a centralized system, secrecy cannot be upheld because the government can have access to anyone's bank account information.

The internet has opened the possibility of new type of money that removed the centralised entities interventions and enabled people to have a secure and private transaction that requires no bank permission, no ID and no delays. This is a crucial life changing money system, especially when we observe that more than 1.7 billion people in the world have no access to a bank account whereas you can have a digital wallet in a minute with only two things: a laptop and WiFi connection. Nonetheless, it is true that the internet is nothing more but connected cables that can be shut down in some particular or extreme incident (like the wars). Thus, find oneself unable to access their money which is something unlikely to happen in a centralized system.

The internet money has been baptised: crypto-currency and it operates within the so-called blockchain technology. The blockchain can simply be illustrated as a list of transactions in chronological order, held in blocks tied together in a chain. All the transactions are public but anonymous thanks to the power of cryptography. The first blockchain ever created was Bitcoin (also the name of the crypto-currency) back in 2009 by Satoshi Nakamoto. Nowadays, we count more than $10,000$ crypto-currencies and several blockchains.

Additionally, the global crypto-currency market size was valued at \$1.49 trillion in 2020 and soared to an all-time high, reaching a market cap of \$3 trillion. Obviously, this is a drop in the bucket compared with the U.S. stock market's \$48 trillion value. Nonetheless, for an asset class that was little more than a decade old, it was significant. In fact, the crypto-currency market is expected to witness promising growth in the coming years owing to improved data transparency and independency across payments in banks, financial services, insurance and so on. Moreover, the market is getting more and more regulated and thus more trusted by investors since regulation protects them from fraud and other risks.

Figure 2.1: Crypto Market Cap from April 2013 to September 2022

Finally, it's impossible to ignore the consequences of the emergence of bitcoin: the creation of Decentralised Finance (DeFi) which is a financial system that is open to everyone and minimizes the need to trust and rely on a central authority, the appearance of the Initial coin offerings (ICOs) which are a popular way to raise funds for products and services usually related to cryptocurrency, Web3 (a new iteration of the World Wide Web which embodies concepts such as decentralization, blockchain technologies, and token-based economics), Decentralized autonomous organization (DAO) and so on.

All in all, as far as I'm concerned, it makes sense to view DeFi and conventional, centralized banking infrastructure as rival or conflicting forces. But in reality, collaborating and working together in the future is their greatest option. By reducing the barriers and silos connected to centralized finance, DeFi may aid currency institutions in becoming more resilient. Besides, I see the crypto and blockchain world as a golden opportunity for research, experimentation and innovations and discover how different the two worlds are mathematically speaking.

2.3 Value-at-risk: Definition

Investors and traders are getting more and more interested in the crypto world, there is now

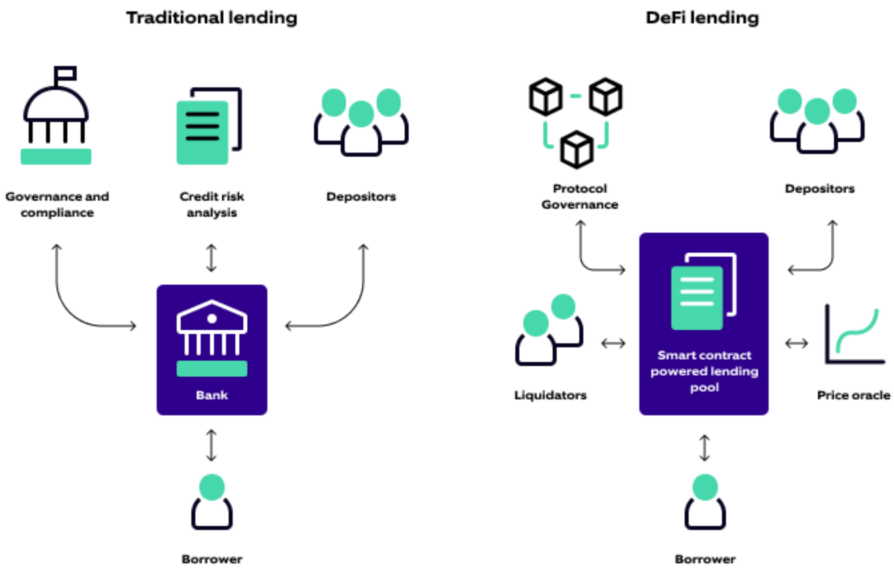

Figure 2.2: Traditional vs. DeFi finance in the lending space scenario

more than ever, a need for instruments for risk management to prevent major losses since the the most concerning aspect of this new technology is the high volatility of crypto-currencies (whose volatility can vary from 70% to 200%) compared to traditional finance (15 per cent on average at least for equities and indexes).

Modern financial markets have identified several major kinds of risk: credit risk, operational risk, liquidity risk, and market risk. In recent years, researchers and market practitioners have paid more attention to Value-at-Risk (VaR) for the analysis of market risk. The VaR was widely used in trading portfolios to measure market risk in the 1990s. Its origins can be traced back to 1992, when the New York Stock Exchange imposed capital requirements on member firms. It has its origins in portfolio theory, as well as a crude VaR measure published in 1995.

This statistical technique is used to estimate the maximal amount (in dollars) that can be lost by a portfolio, over a period of time or future time horizon, and for a given risk level. In other words, there are three key elements to describe the Value at Risk (VaR): the time period over which the risk is assessed, the risk level and the dollar amount of VaR. Typically, assume the 95% daily VaR of my portfolio is \$100k. This can be interpreted in two equivalent ways: my portfolio has 5% chance of losing as least \$100k over the next one-day period or on average, my portfolio will lose at least \$100k, once every 20 days.

Mathematically speaking, the value-at-risk is a statistic that is used to forecast a great possible losses over a certain time period under a certain probability. It is computed across various confidence levels on either simulated or historical data.

We consider losses P&L of a portfolio over a certain time horizon and we define:

$$X_t := -P\&L_t, \tag{2.1}$$

$$P\&L_t = V_t - V_{t-1} \tag{2.2}$$

where $t > 0$ and the $P\&L_t$ is computed between times $t-1$ and t. A loss means a positive X_t. Intuitively, the VaR represents the threshold above which losses (with changed sign) occur with probability at most α. It should be computed conditionally to $\mathcal{F}_t$ where $(F_t : t \geq 0)$ is the filtration modelling all the information available at different times t.

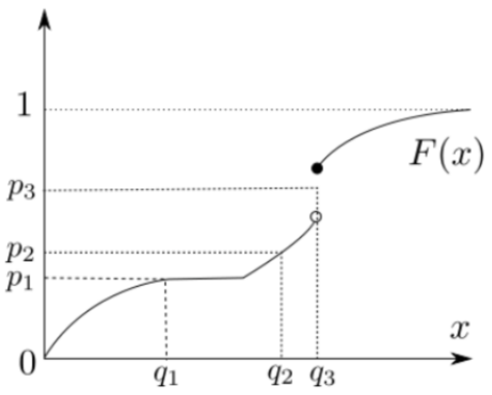

Figure 2.3: Example of c.d.f. as a function of x, and related values of $\mathbf{q}_\alpha(q_i) = p_i$ for different probability levels p_i

Definition: (For a general reference on VaR, see [Sto11] and [AJME15])
Denote by F_t the cumulative distribution function (c.d.f.) of X_{t+1}, conditionally to the available information represented by $\mathcal{F}_t \quad t \geq 0$:

$$F_t(x) := \mathbb{P}(X_{t+1} \leq x|\mathcal{F}_t), \;\; \mathcal{F}_t := \sigma(P_0^i, ..., P_t^i, 1 \leq i \leq n) \tag{2.3}$$

where n is the number of the assets in the portfolio and P_t^i the price of the asset i at time t.

The VaR at the confidence-level α is the lower α-quantile of the conditional distribution of X_{t+1} is defined by:

$$VaR_t^\alpha(X_{t+1}) := q_t^\alpha(X_{t+1}) = \inf\{q : F_t(q) \geq \alpha\}, \tag{2.4}$$
$$q_t^\alpha(X_{t+1}) = -q_t^{1-\alpha}(-X_{t+1}) \tag{2.5}$$

where $0 < \alpha < 1$.

Important remark: We typically consider the risk level $\alpha = 95\%$ or 99% instead of 5% and 1% so that we expect the quantile to be positive.

On the other hand, it is obvious that $VaR_t^\alpha(X_{t+1})$ is an increasing function of α.

VaR Decomposition:

Let us introduce the first two moments of X_{t+1} conditional on $\mathcal{F}_t$, the information available at time t:

$$\mu_{t+1|t} = \mathbb{E}_t(X_{t+1}|\mathcal{F}_t),$$
$$\sigma^2_{t+1|t} = \mathbb{V}_t(X_{t+1}|\mathcal{F}_t).$$

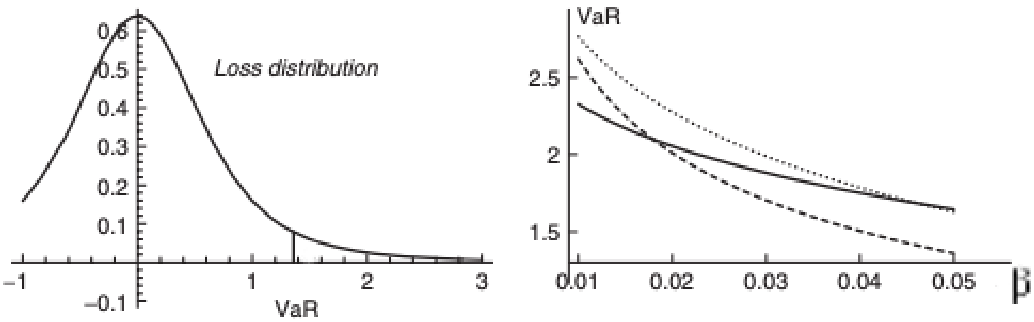

Figure 2.4: VaR is the β-quantile of the conditional loss distribution (left). The righthand graph displays the VaR as a function of $\beta \in [1\%, 5\%]$ for Gaussian distribution (solid line), a Student t distribution with 3 degrees of freedom S (dashed line) and a double exponential distribution E (thin dotted line). The three laws are standardized so as to have unit variances.

Suppose that:

$$X_{t+1} = \mu_{t+1|t} + \sigma_{t+1|t} \times X^*_{t+1} \tag{2.6}$$

where X^*_{t+1} is a standardized (conditionally centered, unit variance) random variable with cumulative distribution function G.

Generally speaking, G is time dependant. However, in this hypothesis, since X^* is centered with a unit variance, we consider that the dependence is all contained in the conditional mean and the conditional variance.

Denote by $\tilde{G}$ the quantile function of the variable X^*_{t+1}. If $\tilde{G}$ is continuous and strictly increasing, we simply have: $\tilde{G} = G^{-1}$, where G^{-1} is the ordinary inverse of G. In view of (2.4) and (2.6) it follows that:

$$\mathbb{P}\left(VaR^\alpha_t(X_{t+1}) \geq \mu_{t+1|t} + \sigma_{t+1|t}X^*_{t+1}\right) = \alpha = G\left(\frac{VaR^\alpha_t(X_{t+1}) - \mu_{t+1|t}}{\sigma_{t+1|t}}\right)$$

Consequently,

$$VaR^\alpha_t(X_{t+1}) = \mu_{t+1|t} + \sigma_{t+1|t} \times \tilde{G}(\alpha). \tag{2.7}$$

VaR can thus be decomposed into an 'expected loss' $\mu_{t+1|1}$, the conditional mean of the loss, and an 'unexpected loss' $\sigma_{t+1|1}\tilde{G}(\alpha)$, also called economic capital.

On the other hand, it is sometimes more interesting to compute the quantile on variations other than losses. Traditionally, besides the $P\&L$, there are two other ways of computing portfolio variations: arithmetic returns (AR) and log returns (LR). In this case, we finally compute the VaR after finding the relationship between $VaR^\alpha_t(X_{t+1})$ and $q^\alpha_t(Y_{t+1})$ where $Y_{t+1} \in \{\mathbf{AR}_{t+1}, \mathbf{LR}_{t+1}\}$.

$$\mathbf{AR}_t = \frac{P\&L_t}{V_{t-1}}, \quad \mathbf{LR}_t = \log(\frac{V_t}{V_{t-1}}) \tag{2.8}$$

where V_t the value of the portfolio at time t.

2.4 Data

As far as the data is concerned, we focus only on historical real data starting from January 2019 to June 2022. We collect the prices of nine crypto-currencies: BTC, ETH, XRP, LTC, LINK, MATIC, DOT, SOL and ADA by extracting 5:30 pm asset prices of each day using all the centralised and decentralized exchanges from the Kaiko Database between January 2019 to June 2022. However, for some crypto-curencies like MATIC for example, their prices only start at august 2020. This is not a bothering issue since we use one year of historical data to generate daily Value-at-Risks. For example: we use one year data from August 2020 to August 2021 and generate daily VaRs from September 2021 to June 2022.

On the other hand, we simulate two type of portfolios:

- Mono-asset: containing exactly one asset for example 1 BTC.
- Multi-asset portfolio: by putting equal weights on the chosen cryptos based on an initial investment, i.e. $w_i = \frac{1}{n}$ where n is the number of the assets of the portfolio and w_i is the weight of the asset i.

Moreover, it is crucial to point out that the portfolio is static and will never be rebalanced.

Examples of different portfolios:

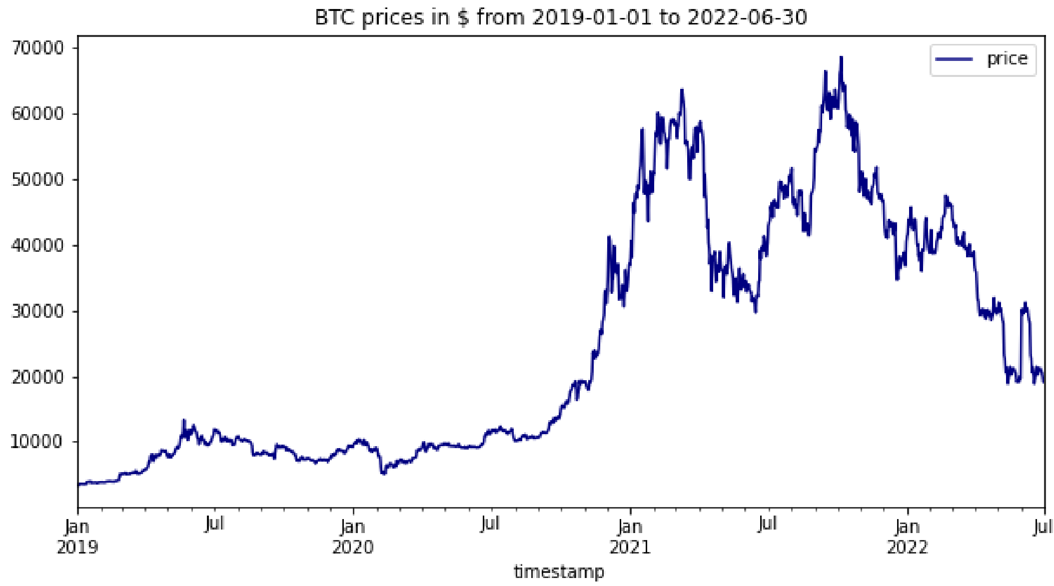

Figure 2.5: Mono-Asset Portfolio evolution

From the graphs, we notice different regimes: a still one from 2019 to the end of the year, a soaring one from the beginning of 2021 and the third and the plunging one from the end of January 2022. In fact, the crypto market is evolving very quickly and can switch from a

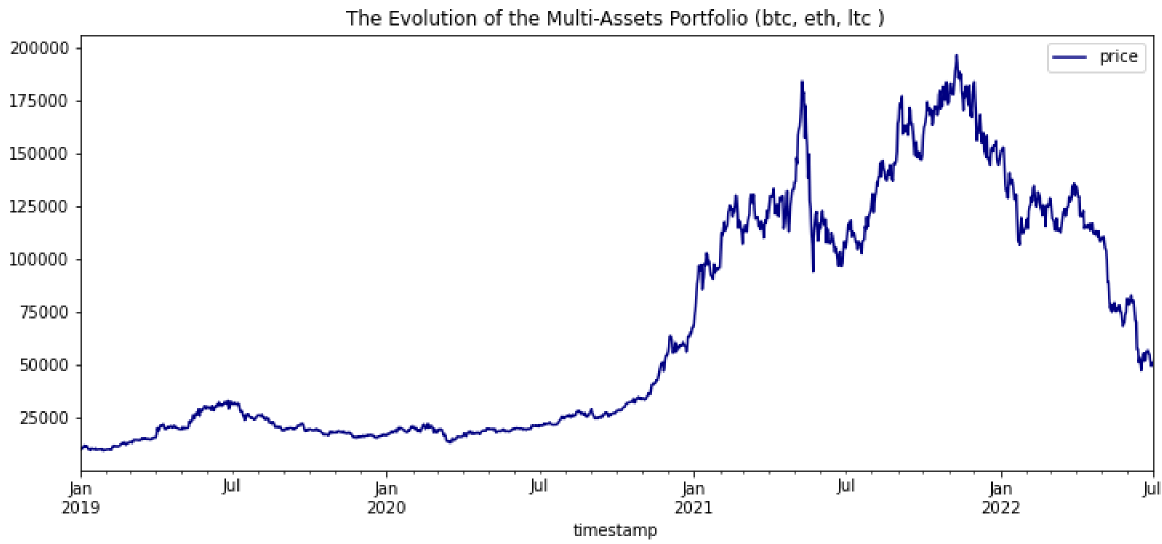

Figure 2.6: Multi-Asset Portfolio evolution

Coins	BTC	ETH	LTC
Quantities	0.895	24.70	109.56

Table 2.1: Quantities of the asset of the multi-asset portfolio

market regime to another in few weeks. One of the main goals of this study is to catch sight of the behaviour of different VaR methods moving from a regime to another. In other words, this point is very interesting for the computation of the different VaRs to see which approach captures the change in the portfolio's pattern.

Chapter 3

Methodologies

There are three kinds of methods traditionally used to estimate the Value-at-Risk: non-parametric methods, parametric methods and semi-parametric methods.

- The non-parametric approach uses actual historical data, it is simple and easy to use. There is no hypothesis about the distribution of the data.
- Parametric methods are calibrated on the observed data (usually two models, one for the mean and one for the variance). Closed-form formulas for the quantile are then derived, or Monte-Carlo simulations are used to estimate the quantile.
- Semi-parametric methods are a hybrid method that combines the two first ones. Only a part of the model is parametric and the rest is non parametric.

On the other hand, it's important to also distinguish two categories of approaches: the one that considers the entire value of a portfolio and the one that considers the portfolio's assets value. In this latter, we are interested in the correlation between the coins rather than just the actual worth of the entire portfolio.

To develop these methodologies, we start by listing some facts about crypto portfolios:

- We don't have much historical data. Sometimes on traditional finance, up to a decade is used to compute VaR and backtest it. Something that we don't have in the crypto finance. In fact, we use only three years of data.
- The market is evolving very quickly and can switch from a market regime to another in few weeks as seen in 2021-2022 when the value of most of cryptos lossed 70% of their value in two months.
- The volatility of cryptos is much higher than that of traditional assets.

3.1 Historical Simulation Method

Historical simulation is the most popular and also the simplest method among all the approaches in terms of implementation and understanding. It's a non-parametric approach that uses specific historical returns (daily variations, arithmetic yield and so on) to construct the cumulative distribution function. In fact, there is no assumptions made on the distribution of the historical returns unlike the parametric approaches. This method consists in replacing the theoritical quantile by an empirical quantile computed on the past data, over a rolling period. It is in fact an entirely data-driven method with no assumptions made about the model. In general, this approach requires a long history of returns in order to get a meaningful VaR but in our case we only use one year of returns data.

The first step lies in setting the time interval and then calculating the returns (log returns, daily variations and so on) of the portfolio. Then, compute the quantile related to the risk level based on the returns distribution. For example a 0.95 quantile if we want to compute the 95% confidence level value-at-risk.

The mathematical formulation of the VaR changes based on the choice of the returns.

- if we use the arithmetic returns:

  $$VaR_\alpha^t(X_{t+1}) = VaR_\alpha^t(V_t - V_{t+1}) = V_t q_t^\alpha(-\frac{V_{t+1} - V_t}{V_t}) = V_t q_t^\alpha(-\mathbf{AR}_{t+1}) = -V_t q^{1-\alpha}(\mathbf{AR}_{t+1}),$$

  because V_t is $\mathcal{F}_t$-measurable.

- if we use the losses $X_{t+1} = -P\&L_t$:

  $$VaR_t^\alpha(X_{t+1}) = q_t^\alpha(X_{t+1}) \text{ of the equation (2.4)}$$

  where V_t is the value of the portfolio at time t, q_t^α is the α-quantile calculated using the returns and X_{t+1} the portfolio loss between t and $t+1$.

3.2 ARMA-GARCH Models Method

For a better understanding of the ARMA and GARCH models, let's first introduce the location scale approach. A probability distribution can be characterized by location and scale parameters. Location and scale parameters are typically used in modeling applications, that can be defined as:

$$Y_{t+1} = \mu_t + \sigma_t Z_{t+1} \tag{3.1}$$

Where Y_{t+1} and Z_{t+1} are random variables, μ_t is the mean, σ_t the volatility process. All these variables follow series of different models.
For example, the probability density function for the standard normal distribution, which has the location parameter equal to zero and scale parameter equal to one.

3.2.1 Stationarity

Stationarity means that the statistical properties of a process generating a time series do not change over time. It does not mean that the series does not change over time, just that the way it changes does not itself change over time. A stationary process is mean-reverting, i.e, it fluctuates around a constant mean with constant variance. Besides, there exist two types of stationarity: the weak form and the strong form.

Weak stationarity is when the time-series has constant mean and variance throughout the time. It only requires the shift-invariance (in time) of the first moment and the cross moment (the auto-covariance). Formally, the process $\{X_t; t \geq 0\}$ is weakly stationary if:

- The first moment of X_t is constant; i.e. $\forall t,\ \mathbb{E}(X_t) = \mu$.

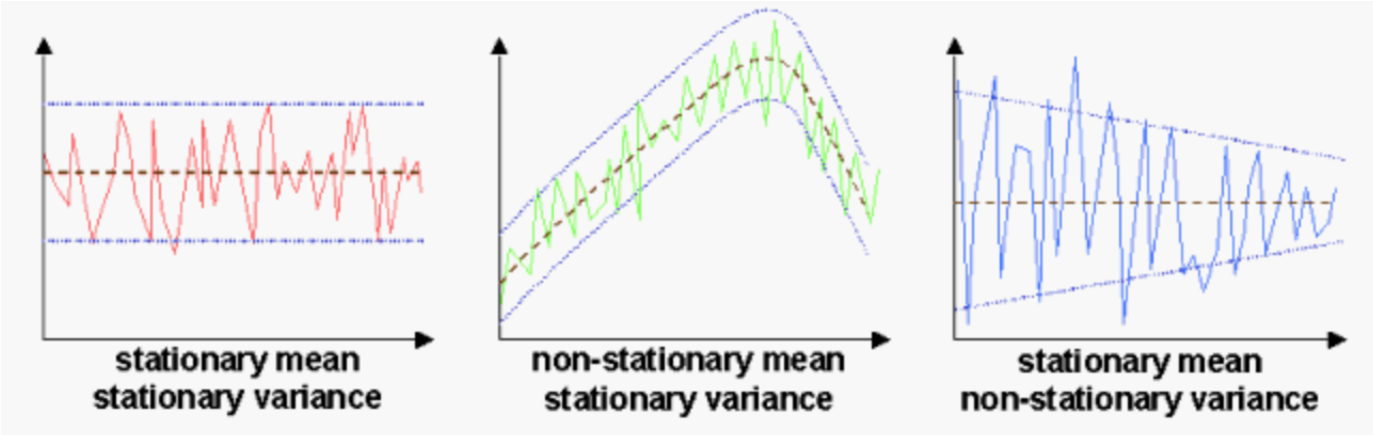

Figure 3.1: Illustration of stationary and non stationary time series

- The second moment of x is finite for all t; i.e.$\forall t,\ \mathbb{E}[X_t^2] < \infty$ (which also implies of course $\mathbb{E}[(X_t - \mu)^2] < \infty$; i.e. that variance is finite for all t).
- The cross moment — i.e. the auto-covariance depends only on the difference $u - v$; i.e. $\forall u, v, a, (X_u, X_v) = cov(X_{u+a}, X_{v+a})$.

Strong stationarity is when the distribution of a time-series is exactly the same trough time. requires the shift-invariance (in time) of the finite dimensional distributions of a stochastic process. This means that the distribution of a finite sub-sequence of random variables of the stochastic process remains the same as we shift it along the time index axis. For example, all i.i.d. stochastic process are stationary. Formally: $\{X_t; t \geq 0\}$ is stationary if:

$$F(X_{t_{1+\tau}}, ..., X_{t_{n+\tau}}) = F(X_{t_1}, ..., X_{t_n})$$

Why is stationarity needed ?

Stationarity is an important concept in the field of time series analysis with tremendous influence on how the data is perceived and predicted. When forecasting or predicting the future, most time series models assume that each point is independent of one another. The best indication of this is when the dataset of past instances is stationary. In other words, a stationary time series gives the property of time-independence. Thus, time series with trends or wit seasonality will affect their value at different time and therefore leads to unreliable forecasting.

The most basic method for stationarity detection rely on plotting the data, and visually checking for trend and seasonal components. This is mostly a dubious task. However, statistical tests provide a quick check and confirmatory evidence that the time series is stationary or non-stationary. They can only be used to inform the degree to which a null hypothesis can be rejected or fail to be rejected. One of the most popular tests is the Augmented Dickey-Fuller test which is used through the function $adfuller$ from *statsmodels* library in python. The value of p-value is used to determine whether there is stationarity. If the value is less than 0.05, the stationarity exists.

3.2.2 VaR Computation

Let Y_t be a strictly stationary time series representing daily observations of the portfolio's log returns and following a location scale model. Let Q_t^α the quantile on log-returns.

$$Y_{t+1} = \mu_{t+1|t} + \sigma_{t+1|t} Z_{t+1} \tag{3.2}$$

$$Q_t^\alpha = \mu_{t+1|t} + \sigma_{t+1|t} q_t^\alpha \tag{3.3}$$

where:

- Z_{t+1} is called the innovation process allowed to follow one of the three distributions: normal, t-Skewt and t-Student.
- $\mu_{t+1|t}$ is the conditional mean (could be zero, constant or an ARMA(p,q) model i.e. an auto-regressive moving average process).
- $\sigma_{t+1|t}$ is the conditional volatility process that is estimated using the most popular GARCH models: GARCH(p,q), EGARCH(p,q) and APARCH(p,q) existing in the well-known *arch_model* python library.
- q_t^α is a quantile of either the chosen innovation process distribution or the log-returns distribution.

We assume that $\mu_{t+1|t}$ and $\sigma_{t+1|t}$ are measurable with respect to $\mathcal{F}_t$ that holds information about the return process available up to time t.
Furthermore, q_α is computed using two different methods: parametric and semi-parametric approaches. In fact, we are interested in estimating quantiles in the tails of either the three distributions mentioned beforehand or using the one year log-returns data. This latter is similar to the historical simulation. Therefore, in this case the GARCH-VaR would be considered a semi-parametric method.

3.2.3 VaR Formulation

Finally, since we are using the log returns in this methodology, we need to reverse the VaR formula, i.e. we use the following equation: $q_t^\alpha(\phi(L)) = \phi(q_t^\alpha(L))$, if ϕ is C^0 and increasing and L a random variable.

$$\begin{aligned}
VaR_t^\alpha(-V_{t+1} + V_t) &= -q_t^{1-\alpha}(V_{t+1}) + V_t \\
&= -q_t^{1-\alpha}\left(e^{\log(V_{t+1})}\right) + V_t \\
&= V_t\left(1 - e^{q_t^{1-\alpha}(\mathbf{LR}_{t+1})}\right) \\
&= V_t\left(1 - e^{-q_t^{\alpha}(-\mathbf{LR}_{t+1})}\right).
\end{aligned}$$

Thus:

$$VaR_t^\alpha(X_{t+1}) = VaR_t^\alpha(-V_{t+1} + V_t) = V_t\left(1 - e^{Q_t^{1-\alpha}}\right). \tag{3.4}$$

Remark: We consider that $Q_t^\alpha = q_t^{1-\alpha}(\mathbf{LR}_{t+1})$ demonstrated in VaR decomposition section, equation (2.7).

3.2.4 Mean Models

Apart from a constant or zero mean, we can also use another approach to get a more accurate value of $\mu_{t+1|t}$. We cite three different models: AR(p), MA(q) and lastly ARMA(p,q). For general reference on ARMA models see [WIN].

The general ARMA model was described in the 1951 thesis of Peter Whittle, Hypothesis testing in time series analysis, and it was popularized in the 1970 book by George E. P. Box and Gwilym Jenkins. The ARMA model is a tool for understanding and, perhaps, predicting future values in this series. The AR part involves regressing the variable on its own lagged (i.e., past) values. The MA part involves modeling the error term as a linear combination of error terms occurring contemporaneously and at various times in the past. The model is usually referred to as the ARMA(p,q) model where p is the order of the AR part and q is the order of the MA part (as defined below).

Remark: For the sake of simplicity, we denote by μ_t the conditional mean $\mu_{t+1|t}$.

Autoregressive Model (AR): The autoregressive model specifies that the output variable depends linearly on its own previous values and on a stochastic term (an imperfectly predictable term). Thus the model is in the form of a stochastic difference equation. The order of an autoregression is the number of immediately preceding values in the series that are used to predict the value at the present time. The order q of the AR model can usually be estimated by looking at the ACF plot of the time series.

$$\mu_t = \sum_{i=1}^{p} \phi_i \mu_{t-i} + \epsilon_t$$

where: $(\phi_i)_{1\leq i\leq p}$ are parameters and ϵ_t is white noise, usually independent and identically distributed (i.i.d.) normal random variables.

In order for the model to remain stationary, the roots of its characteristic polynomial must lie outside of the unit circle.

Moving-average Model (MA): The moving-average model is essentially a finite impulse response filter applied to white noise. It is a time series model that accounts for very short-run autocorrelation. It basically states that the next observation is the mean of every past observation. The order of the moving average model, q, can usually be estimated by looking at the ACF plot of the time series.

$$\mu_t = c + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t$$

where: $(\theta_i)_{1\leq i\leq q}$ are the parameters of the model, c is the expectation of μ_t (often assumed to equal 0) and ϵ_t are again, i.i.d. white noise error terms that are commonly normal random variables.

ARMA Model:

The ARMA model is essentially an infinite impulse response filter applied to white noise, with some additional interpretation placed on it. The general ARMA model was described in the 1951 thesis of Peter Whittle, who used mathematical analysis (Laurent series and Fourier analysis) and statistical inference.[6][7] ARMA models were popularized by a 1970 book by George E. P. Box and Jenkins, who expounded an iterative (Box–Jenkins) method for choosing and estimating them.

$$\mu_t = \sum_{i=1}^{p} \phi_i \mu_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t$$

Finally, before fitting any of these mean models, we need to choose the appropriate lag for the AR model using the autocorrelations. The auto-correlation function is a statistical representation used to analyze the degree of similarity between a time series and a lagged version of itself. This function allows the analyst to compare the current value of a data set to its past value. In fact, it is determined by checking the partial autocorrelation plot. The '*plot_pacf* 'method is used to plot to assess the direct effect of past data on future data.

3.2.5 Volatility Models

The generalized autoregressive conditional heteroskedasticity (GARCH) process is a model to estimate the volatility of financial markets. GARCH aims to minimize errors in forecasting by accounting for errors in prior forecasting and enhancing the accuracy of ongoing predictions. (For more knowledge on GARCH models see [FZ10])

Remark: For the sake of simplicity, we denote byσ_t the conditional mean $\sigma_{t+1|t}$.

ARCH(p) Model:

ARCH (Auto-regressive Conditional Heteoskedastic Model) is the simplest model in stochastic variance modeling which was developed by Engle (1982).The model can be expressed as follows:

$$\sigma_t^2 = \omega + \textstyle\sum_{i=1}^{p} \alpha_i \epsilon_{t-i}^2,$$

To assure σ_t^2 is asymptotically stationary random sequence, we can assume that $\sum_{i=1}^{p} \alpha_i < 1$ and ω, α_i are parameters related to the model to be estimated.

GARCH (p, q) **Model:**

The GARCH Model, short for The Generalized Auto-Regressive Conditional Heteoskedastic Model, is based on an infinite ARCH specification. Standard GARCH models assume that positive and negative error terms have asymmetric effect on the volatility. The model can be expressed as follows:

$$\sigma_t^2 = \omega + \textstyle\sum_{i=1}^{p} \alpha_i X_{t-1}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-i}^2$$

where p is the order of the symmetric innovation and q is the order of the lagged (transformed) conditional variance and and must both satisfy $p, q \geq 0$ and $\omega, \alpha_i, \beta_i$ are parameters related to the model to be estimated.

EGARCH (p, q) **Model:**

The exponential GARCH is another form of the GARCH model. E-GARCH model was proposed by Nelson (1991) to overcome the weakness in GARCH handling of financial time series. In particular, to allow for asymmetric effects between positive and negative asset returns.

The E-GARCH model differs from GARCH in several ways. For instance, it used the logged conditional variances to relax the positiveness constraint of model coefficients. Another advantage, as pointed out by Nelson and Cao (1992), is that there are no restrictions on the parameters. The model can be expressed as follows:

$$log(\sigma_t^2) = \omega + \sum_{k=1}^{q} \beta_k g(Z_{t-k} + \sum_{k=1}^{p} \alpha_k log(\sigma_{t-k}^2))$$

where $g(Z_t) = \theta Z_t = \lambda(|Z_t| - E(|Z_t|))$, σ_t^2 is the conditional variance and $\omega, \beta, \alpha, \theta, \lambda$ are coefficients.

APARCH (p, q) **Model:**

The Asymmetric Power ARCH model is brought by Ding, Granger and Engle (1993). It attempts to capture asymmetric responses of volatility to positive and negative 'news shocks', the phenomenon known as the leverage effect. It can also well express the fat tails and excess kurtosis. The general structure is:

$$\sigma_t^\delta = \omega + \sum_{i=1}^{p} \alpha_i(|\epsilon_{t-i}| - \gamma_i \epsilon_{t-i})^\delta + \sum_{k=1}^{q} \beta_k \sigma_{t-k}^\delta$$

where ω, α_i, γ_j , β_i and δ_i are the parameters which are needed to be estimated and $\epsilon_t = \sigma_t Y_t$ and Y_t is a standard gaussian.

3.2.6 Model Selection

Before calculating the value-at-risk, we mainly use one metric to select the best GARCH model. It is called the Akaike's information criterion (AIC). This criterion was formulated by the statistician Hirotugu Akaike in 1973. It is a good measure for testing the goodness of how fit the model is mathematically. It measures the amount of information lost by training the GARCH model (or ARMA model).

The AIC must be as low as possible. In other words, A model with a minimum value of AIC is chosen to be the best fitting model among several competing models. Another crucial consideration is that AIC is not intended to find a valid model. The best-fitting model is not necessarily the actual model. Rather, it indicates that the model is superior than rival models in terms of providing the closest approximation to the genuine model or reality. Naturally, the best fitting model may change as a function of sample size, because model parameters may be calculated more reliably with a larger sample size. In the factor-analytic context, for example, a relevant inquiry might not be how many right factors there are, but how many factors can be safely recovered given the data set at hand.

The AIC value of the model is the following:

$$AIC = 2k - 2\log(\hat{L})$$

where k is the number of estimated parameters in the model and $\hat{L}$ is the maximum value of the likelihood function of the model.

Remark: We could also use another criterion for model selection which is called the Bayesian information criterion (BIC), $BIC = k\log(n) - 2\log(\hat{L})$. Likewise, the model with lower BIC values are generally preferred.

Example of GARCH VaR computation:
The first step in this methodology consists in fitting for example the model GARCH(1,1) to find the constants α_1 and β_1 to compute the volatility that we include in the VaR computation (2.5) along with a zero-conditional mean and a standard normal distribution quantile.

3.3 Variance-Covariance Method

The variance-covariance method uses the variances and covariances of assets and the arithmetic returns of each asset for VaR calculation. It is hence a parametric method as it depends on the parameters of the probability distribution of the returns.

The idea behind the variance-covariance is similar to the idea behind the historical. It uses historical price movements (standard deviation, mean price) of a given equity or portfolio of equities over a specified lookback period, and then uses probability theory to calculate the maximum loss within your specified confidence interval. It is, in fact, a parametric method. Variance refers to the spread of a data set around its mean value, while a covariance refers to the measure of the directional relationship between two random variables.

The variance-covariance method assumes that asset returns are normally distributed around the mean of the bell-shaped probability distribution. Assets may have tendency to move up and down together or against each other. This method assumes that the standard deviation of asset returns and the correlations between asset returns are constant over time. Let Y_{t+1} be the arithmetic returns of the portfolio and $\hat{Y}_{t+1}$.

The general formula for the variance-covariance Value-at-Risk:

$$Y_{t+1} = \sum_{i=1}^{n} w_t^i \mathbf{AR}_{t+1}^i \tag{3.5}$$

$$\mu_{t+1|t} := \mathbb{E}(Y_{t+1}|\mathcal{F}_t) = \sum_{i=1}^{n} w_t^i \; \mathbb{E}(\mathbf{AR}_{t+1}|\mathcal{F}_t)$$

$$\sigma_{t+1|t}^2 := \mathbb{V}(Y_{t+1}|\mathcal{F}_t) = \begin{pmatrix} w_1 \dots & w_n \end{pmatrix} \begin{pmatrix} \gamma_{11,t} & \gamma_{12,t} & \cdots & \gamma_{1n,t} \\ \ddots & \ddots & \dots & \vdots \\ \gamma_{i1,t} & \gamma_{ii,t} & \ddots & \gamma_{ij,t} \\ \vdots & \dots & \ddots & \vdots \\ \gamma_{n1,t} & \gamma_{n2,t} & \cdots & \gamma_{nn,t} \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \tag{3.6}$$

We have the following hypothesis:

$$Y_{t+1} := \mu_{t+1|t} + \sigma_{t+1|t} Z_{t+1}$$

$$Q_t^\alpha = \mu_{t+1|t} + \sigma_{t+1|t} . q_t^\alpha$$

Where:

- Y_{t+1} represents the arithmetic returns of the portfolio between t and $t+1$.
- Z_{t+1} is an innovation process that follows a standard normal distribution.
- $\mu_{t+1|t}$ is the mean of the arithmetic returns of the portfolio,
- $\gamma_{ij,t}$ is the covariance between arithmetic returns of the asset i and j and the asset j, i.e $\gamma_{ij,t} = Cov(\mathbf{AR}_t^i, \mathbf{AR}_t^j), \quad 0 \leq i \leq j \leq n$, where $\mathbf{AR}_t^i$ is the arithmetic return of the asset i at time t,
- q_t^α is the quantile of the standard normal distribution with α probability,
- w_t^i are the weights of the asset i at time t.

If a portfolio has multiple assets, its volatility is calculated using a matrix. A variance-covariance matrix is computed for all the assets. Otherwise, if it's a mono-asset one, i.e. $n = 1$, σ_t^2 is simply the variance of the portfolio's arithmetic returns.

For simplicity in our experiments, we assume that the covariance matrix $(\rho_{ij,t})_{i,j}$ is constant over time, this is estimated empirically over the data. We could have used a multidimensional GARCH modeling alternatively, this is left to future investigation.

3.3.1 VaR Formulation

Finally, since we are using the arithmetic returns, we need to inverse the VaR formula (2.4): Let w_t^i be the weight of the asset i at time t :

$$w_t^i = \frac{\delta_t^i P_t^i}{\sum_{j=1}^n \delta_t^j P_t^j} \tag{3.7}$$

where δ_t^i is the number of the asset i at time t and P_t^i its value at time t.

$$\begin{aligned}
q_t^\alpha(-V_{t+1}) &= q_t^\alpha(-V_{t+1} + V_t) - V_t \\
&= q_t^\alpha\left(-\sum_{j=1}^n \delta_t^j P_t^j \frac{P_{t+1}^j - P_t^j}{P_t^j}\right) - V_t \\
&= q_t^\alpha\left(-\sum_{j=1}^n w_t^j V_t \mathbf{AR}_{t+1}^j\right) - V_t \\
&= V_t\left(q_t^\alpha(-\sum_{j=1}^n w_t^j \mathbf{AR}_{t+1}^j) - 1\right) \\
&= V_t\left(-q_t^{1-\alpha}(\sum_{j=1}^n w_t^j \mathbf{AR}_{t+1}^j) - 1\right)
\end{aligned}$$

Thus:

$$VaR_t^{\alpha}(-V_{t+1}+V_t) = V_t\left(-q_t^{1-\alpha}(\sum_{j=1}^{n} w_t^j \mathbf{AR}_{t+1}^j)\right) = V_t \times Q_t^{\alpha} \tag{3.8}$$

Remark: We consider that $Q_t^{\alpha} = -q_t^{1-\alpha}(\sum_{j=1}^{n} w_t^j \mathbf{AR}_{t+1}^j)$ as demonstrated in the VaR decomposition section, equation (2.7).

3.3.2 Implementation Steps

The implementation steps to calculate daily VaRs of a portfolio are the following:

- Create a portfolio by fixing the quantities of the assets in the portfolio from January 2020 to June 2022.
- Fix a dataset of one year of prices from January 2020 to January 2021.
- Calculate the arithmetic returns of each asset in the portfolio.
- Create a covariance matrix based on the returns dataframe.
- Update the weights using the formula (3.7) that links the weights and the quantities of the assets.
- Calculate the portfolio mean $\mu_{t+1|t}$ and standard deviation $\sigma_{t+1|t}$.
- Calculate the quantile normal standard distribution with a specified confidence interval (99% and 95% risk level).
- Compute Q_T^{α}.
- Estimate the value at risk (VaR) for the portfolio using the formula (3.8).
- Do it all over again to compute the second day VaR by moving the one year arithmetic returns dataset to the next day, i.e $\{R_1, ..., R_{365}\} \longrightarrow \{R_2, ..., R_{366}\}$.

3.4 Kaiko Value-at-Risk

The Kaiko method is a proprietary method based on a non-parametric approach and is split into two major principles: a normalization and stabilization of the data through pre-processing the returns; and a stronger consideration of recent data compared to the past. The detail of which remains confidential.

Chapter 4

Backtesting

Once the VaR methodology is fixed, the challenging part lies in assessing the accuracy of the measure and by extension. In fact, Financial risk model evaluation or backtesting is a key part of the internal model's approach to market risk management as laid out by the Basle Commitee on Banking Supervision (1996). On the other hand, knowing the measure's accuracy is especially important for financial institutions, which use VaR to estimate how much cash they need to reserve to cover potential losses. Any errors in the VaR model could indicate that the institution is not holding enough reserves, which could result in significant losses not only for the institution but also for its depositors, individual investors, and corporate clients.

That's where backtesting comes in hand. It is in fact the most important part of the VaR study. It is a way to gauge the effectiveness and the accuracy of the model implemented. Backtesting in value-at-risk is used to compare the predicted losses with the ones that are actually realised. A core paper is [Chr98] which describes the goodness of VaR estimations.

What's more, we will refer to an event where the portfolio loss exceeds the VaR measure as a violation. In fact, what's important in backtesting is the clustering of violations. Each time t that a forecast is performed, one can observe, after one day, the realization of -$P\&L_{t+1}$ and compute the violation indicator. It is defined by :

$$I_{t+1} := \mathbb{1}_{-P\&L_{t+1} \leq \hat{V}aR_t^\alpha} \tag{4.1}$$

The sequence of I (hit function) is made of 0 and 1, where the probability of having 1 is α.

This is formalized below:

Definition A sequence $(\hat{V}aR_t^\alpha : t \geq 0)$ of forecasts of $VaR_t^\alpha(-P\&L_{t+1}|\mathcal{F}_t)$ is ideal if

$$\mathbb{P}(-P\&L_{t+1} \leq \hat{V}aR_t^\alpha|\mathcal{F}_t) = \alpha \tag{4.2}$$

for any t. In fact, ideally, we should have: $\frac{1}{m}\sum_{t=1}^{m} I_{t+1} = \alpha$ where m is the number of daily VaRs computed.

Christoffersen, the author of [Chr98] has an interesting characterisation of sequence of ideal forecasts: A sequence $(VaR_{\alpha,t} : t \geq 0)$ of forecasts is ideal if and only if the sequence $(I_t : t \geq 0)$ is i.i.d. with Bernoulli distribution with parameter α.

This leads to several ways to effectively backtest forecasts. This is performed over a data set D of available data for the sequence of predictor $\hat{V}aR_{\alpha t}$.

4.1 The Exceedance Ratio against confidence level

GenerallyGenerally speaking, the exceedance ratio is the frequency with which a random process exceeds some critical value. In this case is it the number of times our losses surpass the VaR.
For instance, we evaluate and expect that with 99% confidence, the worst daily loss will not exceed 1% or be around 5% for the 95% VaR confidence level.

Over and above that, in order to have a clearer outlook of this metric, we also compute the difference between the exceedance ratio and the theoritical one using a range of risk levels from 10% to 99% and plot it against the confidence interval of the quantile determined in our previous mathematical research to find the VaR model that stays within this interval and also to assess the effectiveness of the value-at-risk model in every risk level.

Mathematically, the exceedance ratio is defined by:

$$ER_D = \frac{1}{\#D} \sum_{t \in D} I_t \tag{4.3}$$

Under the *ideal* forecast assumption, it behaves (as $\#D$ is large) as a Gaussian distribution with mean α and variance $\frac{\alpha(1-\alpha)}{\#D}$. It implies that one has :

$$\sqrt{\frac{\alpha(1-\alpha)}{\#D}}(ED_D - \alpha) \notin [-1.96, 1.96] \tag{4.4}$$

with approximate probability 5%

Remark: D is a dataset containing the dates corresponding to all the computed daily VaRs. For instance, if we have computed 500 VaRs, we'd have 500 different dates.

4.2 The log-likelihood Test

For general reference on the log-likelihood tests see [LR].
Working under the assumption of a good forecast where I's are i.i.d., the sequence of $I_D = (I_t, t \in D)$ is a Bernoulli sequence with some fixed parameter: its likelihood (given the Bernoulli parameter is β) equals

$$L(\beta, I_D) := (1-\beta)^{n_{0,D}} \beta^{n_{1,D}}$$

with:

$$n_{1,D} = \textstyle\sum_{t \in D} I_t \ ; \ n_{0,D} = \#D - n_{1,D}$$

The loglikelihood ratio test writes:

$$LR := -2\log(\tfrac{L(\alpha, I_D)}{L(\hat{\alpha}, I_D)}),$$

where: $\hat{\alpha}$ is an estimate of the Bernoulli parameter α using the data set D, that is $\hat{\alpha} = \frac{n_{1,D}}{\#D}$.
As the sample size $\#D$ gets large, LR follows a χ^2-distribution, under the assumption of good forecast.

On the other hand, how do we verify that the value of the likelihood is a $\chi^2(1)$?
Let S a random variable that follows a $\chi^2(1)$ i.e $S = X^2$ where X is a standard normal distribution. Let $a \geq 0$ and $\beta \in (0,1)$ be the risk level of the confidence interval and ϕ the CDF of the standard normal distribution.

$$\mathbb{P}(S \in [0,a]) = \mathbb{P}(X \in [-\sqrt{a}, \sqrt{a}]) = \phi(\sqrt{a}) - \phi(-\sqrt{a}) = 2\phi(\sqrt{a}) - 1 = 1 - \beta$$

$$\mathbb{P}(S \in [0,a]) = 1 - \beta \iff a = (\phi^{-1}(1 - \frac{\beta}{2}))^2 \tag{4.5}$$

In fact, if the value of the log-likelihood is within the interval $[0,a]$ where $a = \phi^{-1}(1-\beta/2)^2 = 6.63489$ and $\beta = 0.01$, we can accept that it follows a $\chi^2(1)$ distribution, i.e *we can accept that the VaR forecast is an ideal forecast.*

Chapter 5

Results

One important point to mention is that all the portfolios tested have the same evolution and the same changing regimes.

The different VaRs results show that:

- All the mono-asset portfolios have the same backtesting results.
- Each multi-asset portfolio has a different output in the backtests.

Consequently, we choose three different portfolios to represent the results:

- \$10k mono-asset portfolio containing the crypto-currency ETHER (ETH).
- \$10k multi-equally weighted- assets portfolio contaning Bitcoin (BTC), Ether (ETH), Ripple (XRP) and Litecoin (LTC)
- \$10k multi-equally weighted- assets portfolio contaning Polkadot (DOT), Chainlink (LINK), Solana (SOL).

5.1 Legend

Let's define some data displayed in the graphs and tables:

- **Kaiko VaR** is the VaR computed using the Kaiko approach.
- **HS EGARCH ARX t** is the VaR computed using an empirical quantile, EGARCH(1,1) model on the volatility, AR model on the mean and innovation process that follows a t-distribution.
- **PM EGARCH ARX t** is the VaR computed using a theoritical t-student quantile, EGARCH(1,1) model on the volatility, AR model on the mean and innovation process that follows a t-distribution.
- **Arithmetic VaR** is the VaR computed using historical simulation on arithmetic returns.
- **Daily Variation VaR** is the VaR computed using historical simulation on P&L.
- **Variance covariance** VaR is the VaR computed using variance-covariance method.

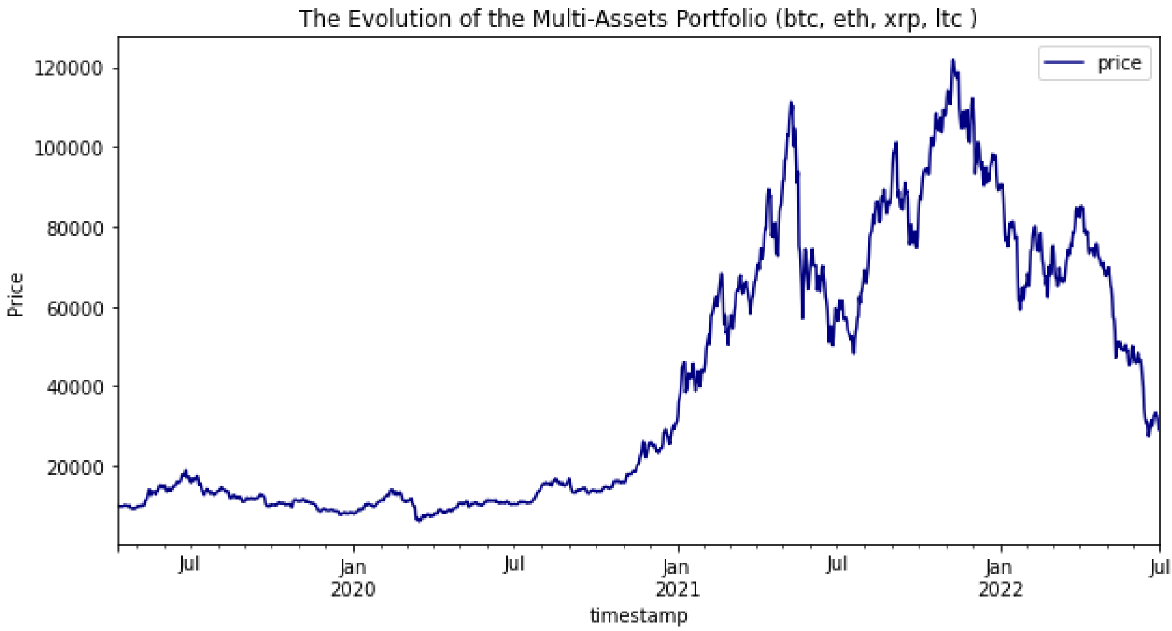

Figure 5.1: Evolution of an equally weighted multi-assets portfolio (BTC, ETH, XRP and LTC)

- Let ENQ be the exceedance ratio normalised quantity, i.e.:

$$ENQ = \sqrt{\frac{\alpha(1-\alpha)}{\#D}}(ER_D - \alpha)$$

Let's recall that ENQ $\in [-1.96, 1.96]$ and the loglikelihood $\in [0, 6.63]$

5.2 GARCH Implementation Results

First, we notice that the returns are stationary from the auto-correlation graph which is pretty expected since one of the ways to correct non-stationarity is principally by differentiating the data, i.e. go from V_t to $V_{t+1} - V_t$, $\forall t \geq 0$ or applying logarithms, i.e. go from V_t to $log(\frac{V_{t+1}}{V_t})$. Generally speaking, to evaluate the order of the lags (p, q) of either an ARMA or a GARCH model, we define and estimate regression models with ARMA(p, q) and garch(p,q) errors by varying p = 1,..,n and q = 1,...,n. Store the optimized loglikelihood objective function value for each model fit and calculate AIC for each model fit. The best fitting model is the regression model corresponding to the lowest AIC. However, one could also examines the plot 5.2 to find the lag after which the partial autocorrelations function (PACF) are all within the confidence interval. In this case, given any portfolio, none exceeds the interval. Therefore the lag parameters of the ARMA and all the GARCH models are fixed to $p = q = 1$.

Secondly, the outputs of the $'arch\_model'$ library mean model ARX (auto-regressive model already implemented in the python library) and the ARMA mean model implemented externally are constant after few days and very close and this given any portfolio. Besides, there is no big difference between the constant mean model where we compute the mean of the returns training data and and AR, MA and ARMA mean models. Moreover, in the graph 5.4, we plot the two GARCH VaRs with the losses. One that uses the empirical quantile and the second that uses the theoritical. Giving any GARCH model, the two quantile approaches give the same result: the parametric and semi-paramteric GARCH VaR coincide or almost coincide. This means that the model of the quantile is irrelevant to the VaR computation.

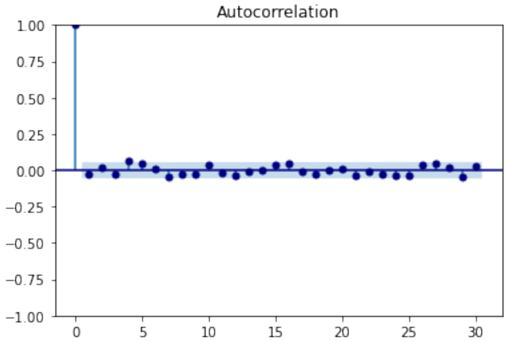

Figure 5.2: Auto-Correlations of the log-returns

Thirdly, given any portfolio, after testing all the combinations of the GARCH models with the three distributions and conditional means, we conclude that based on the AIC results, the model that suits all the portfolios the best is the EGARCH(1,1) with Autoregressive (AR) mean and t-distribution. On the other hand we also notice on 5.3 that the GARCH model with a normal distribution has the highest value of AIC, which means that the Gaussian approach does by no means encapsulate the crypto-market.

EGARCH_ARX_t	-1469.349077
EGARCH_Constant_t	-1469.349077
EGARCH_ARX_skewt	-1467.586356
EGARCH_Constant_skewt	-1467.586356
EGARCH_zero_t	-1465.681600
EGARCH_zero_skewt	-1464.487909
Garch_ARX_t	-1460.932069
Garch_Constant_t	-1460.932069
Garch_ARX_skewt	-1458.939406
Garch_Constant_skewt	-1458.939406
Garch_zero_t	-1457.675673
aparch_Constant_t	-1457.377081
aparch_ARX_t	-1457.377081
aparch_ARX_skewt	-1455.921196
aparch_Constant_skewt	-1455.921196
Garch_zero_skewt	-1455.675673
aparch_zero_t	-1453.883497
aparch_zero_skewt	-1451.883497
aparch_zero_Normal	-1404.683182
aparch_ARX_Normal	-1404.306390
aparch_Constant_Normal	-1404.306390
...	
EGARCH_zero_Normal	-1401.515777
Garch_ARX_Normal	-1394.394278
Garch_Constant_Normal	-1394.394278
Garch_zero_Normal	-1394.311086

Figure 5.3: AIC values for different GARCH models

5.3 Historical Simulation VaR Results

The historical simulation is well-known to be unable to capture the change in regimes of the traditional market, let alone a crypto-market that has a massive ups and down movements. Even if doesn't make assumptions about distribution of returns (uses empirical distribution),

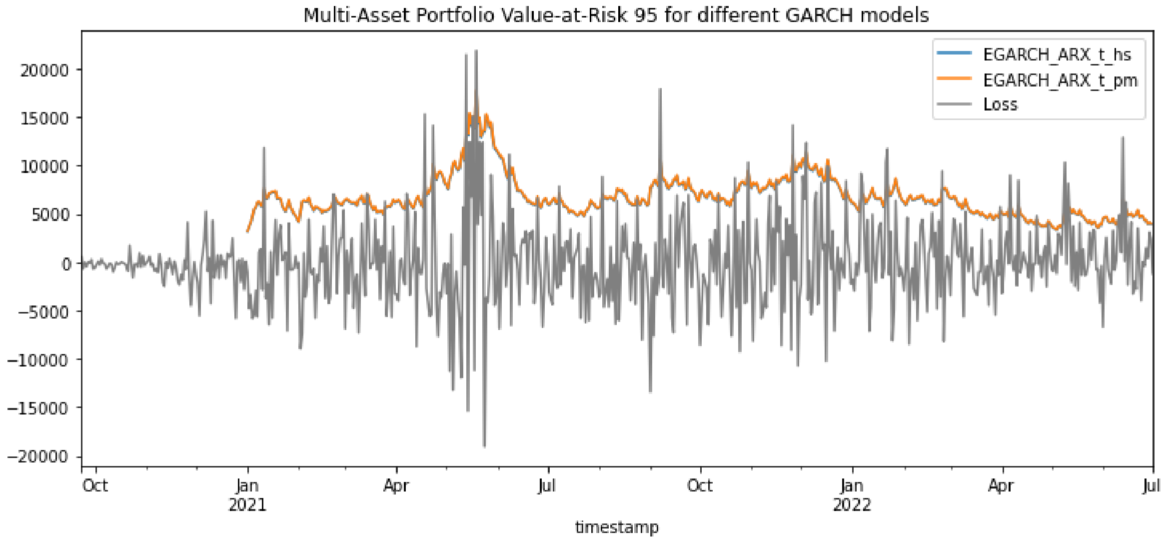

Figure 5.4: Parametric and semi-parametric GARCH VaR with losses

it assumes the past will repeat itself. From the results in **??**, 5.9 and 5.8, despite the cons, we noticed that it doesn't work on the daily variations but it works quite well on the arithmetic returns as this latter stays within the confidence level of the quantile as seen in 5.2, 5.1 and 5.3. This, led us to think that the more stable data we have the more accurate the VaR will be.

5.4 Variance-Covariance VaR Results

The assumptions of return normality and constant covariances and correlations between assets in the portfolio may not hold true in real life. Moreover, the fact that the standard deviation is computed using historical returns of a certain period does not respond appropriately or even sufficiently to changing market conditions just like historical simulation method.

On the other hand, generally speaking, the method variance-covariance has always been implemented with the standard normal distribution. However, the AIC results from the GARCH models led us to test the student distribution since the gaussian distribution seems to be the one that has the highest AIC value. After testing it on multiple portfolios, the results were all the same and not promising as we can see in the graphs 5.6 and 5.5. In fact, other than the incapacity to capture the change in the regimes of the portfolio, the VaR with t-distribution does not perform well compared to the norm-VaR as the daily losses are displayed quite far from the daily VaR during 2021 for example.

5.5 Backtesting Results

From the Losses-VaR graphs of ETH portfolio **??** and multi-asset portfolio 5.9 and 5.8, we see that the variance-covariance doesn't perform as expected especially. In fact, it flattens after 2021. The end of this period marks the change of a regime. On the other hand, we see in the graphs 5.12, 5.10 and 5.14 that in terms of the exceedance ratio of the variance - covariance VaR is very far from the theoritical one for all risk levels whereas the GARCH-VaR and the Arithmetic VaR ones stays within the confidence interval for every risk level starting from 40% for the three portfolios. However, the Daily variation VaR performs very badly and is out of

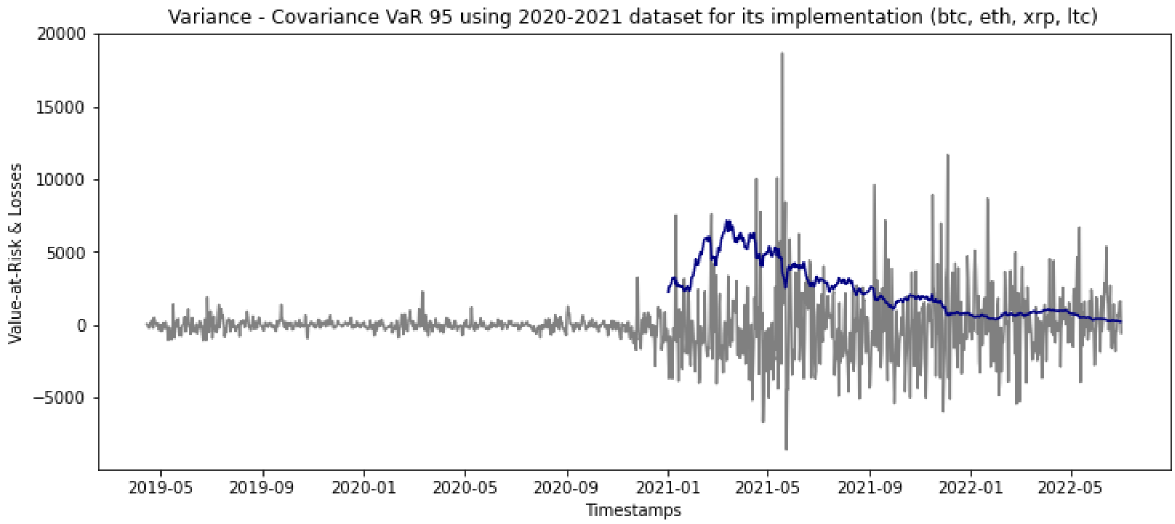

Figure 5.5: Variance-Covariance Value-At-Risk with standard Gaussian distribution

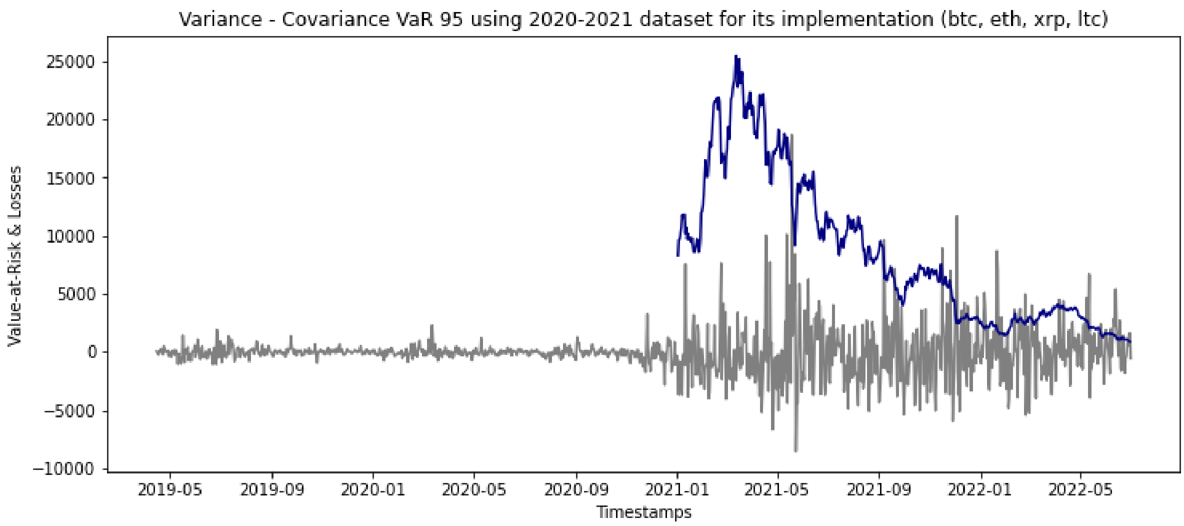

Figure 5.6: Variance-Covariance Value-At-Risk with t-distribution

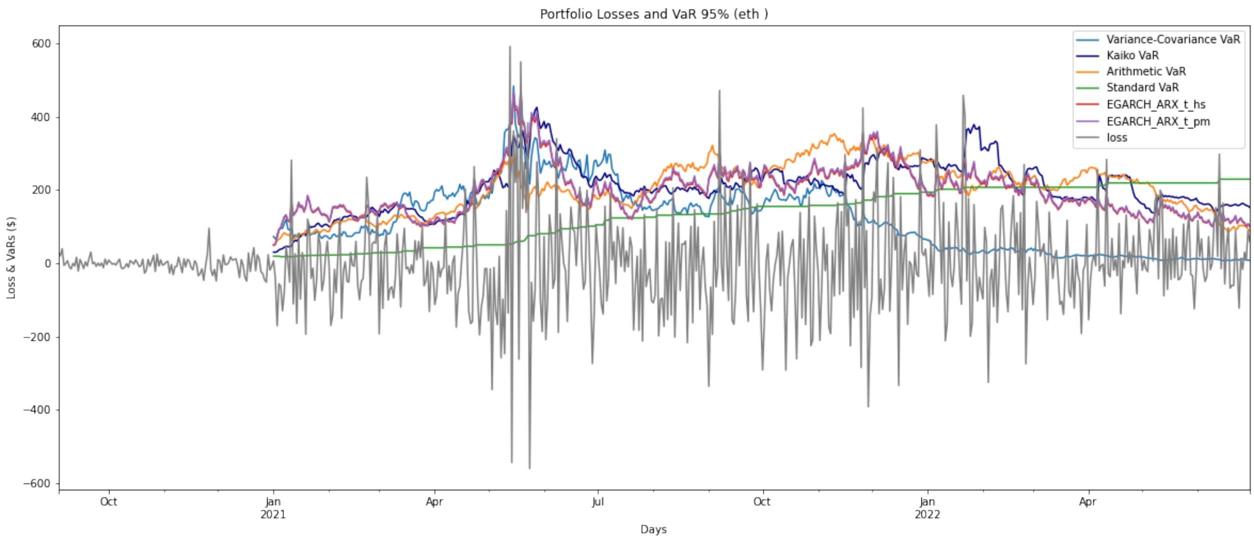

Figure 5.7: Graph of all the different Value-at-Risks displayed on the losses of the mono-asset portfolio (ETH)

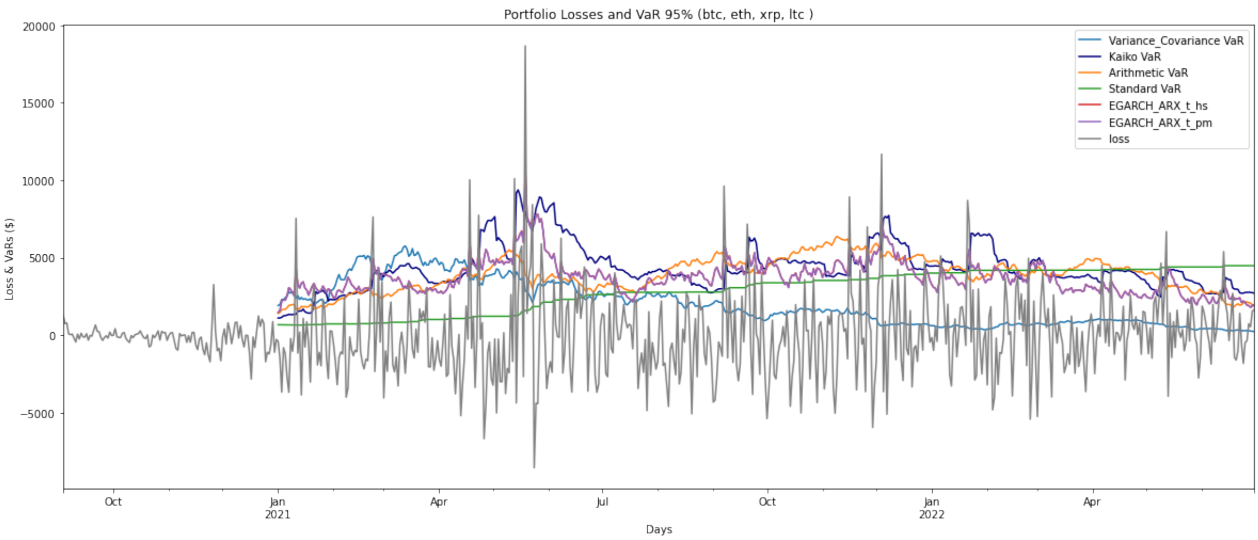

Figure 5.8: Graph of all the different Value-at-Risks displayed on the losses of the multi-asset portfolio (BTC, ETH, XRP and LTC)

VaR 95	**Exceedance Ratio** (%)	**ENQ** ($\times 10^{-5}$)	**Log-likelihood**
Kaiko VaR	5.56	4.30	0.52
Arithmetic VaR	6.04	7.97	1.74
Daily variation VaR	13.8	67	91
HS EGARCH ARX t VaR	4.19	-6.18	1.17
PM EGARCH ARX t VaR	4.19	-6.18	1.17
Var Covar VaR	15.90	83	132

Table 5.1: Numerical backtesting results for different VaR 95 methods on a multi-asset portfolio (BTC, ETH, LTC and XRP)

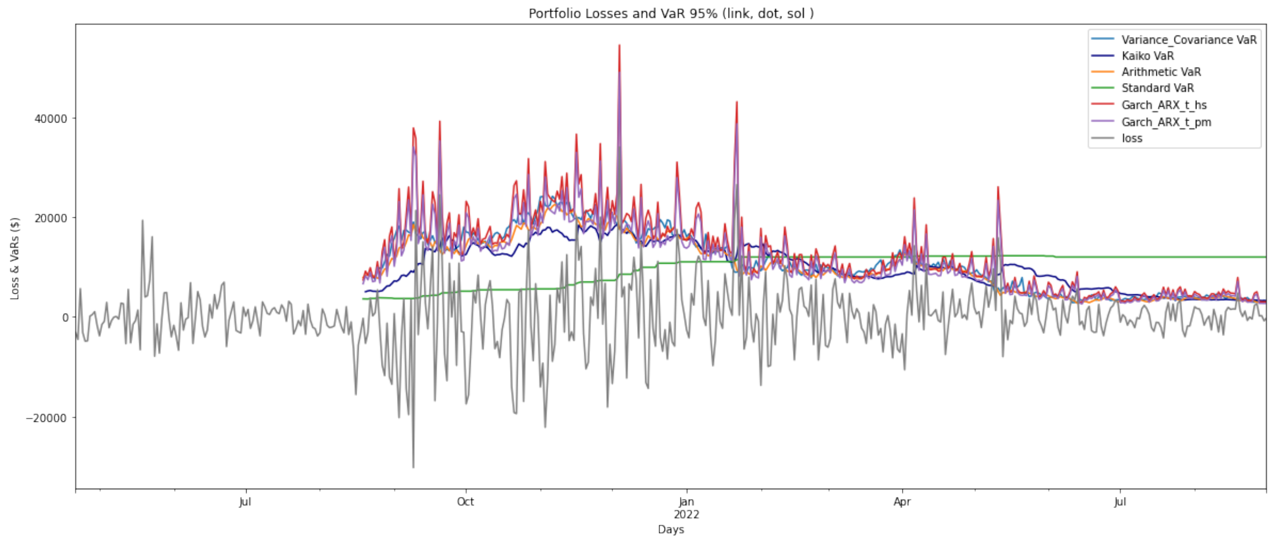

Figure 5.9: Graph of all the different Value-at-Risks displayed on the losses of the multi-asset portfolio (LINK, DOT and SOL)

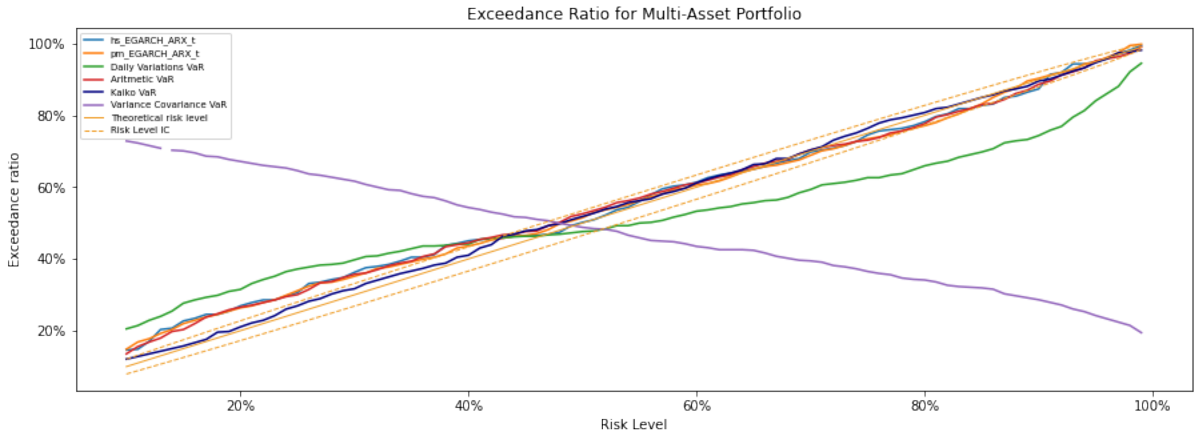

Figure 5.10: Exceedance Ratio on multi-assets portfolio (BTC, XRP, LTC and ETH)

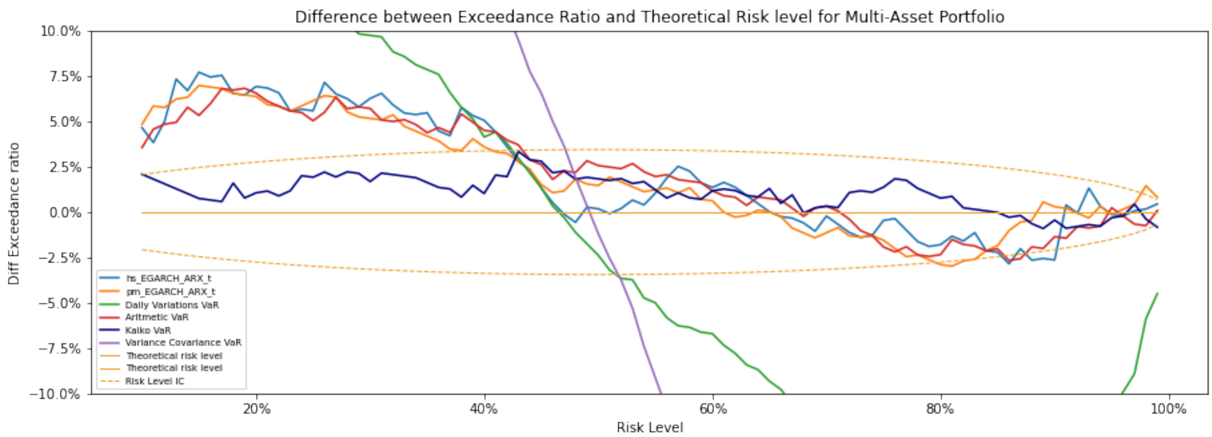

Figure 5.11: Difference between exceedance Ratio and theoretical risk level on multi-assets portfolio (BTC, XRP, LTC and ETH) using 2020 - 2021 training data.

the interval almost all the time. More importantly, the Kaiko VaR performs significantly well for every risk level.

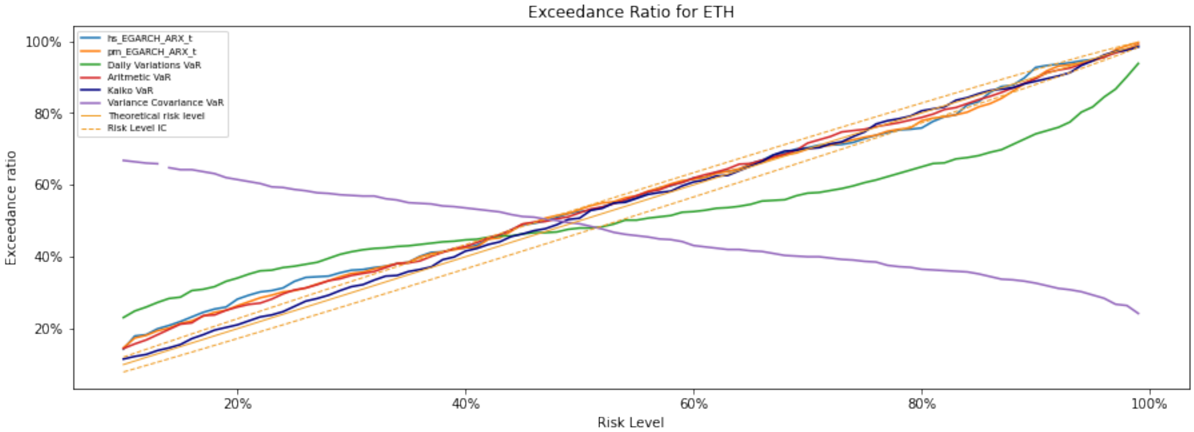

Figure 5.12: Exceedance Ratio on ETH portfolio Backtesting using the 2020-2021 training data

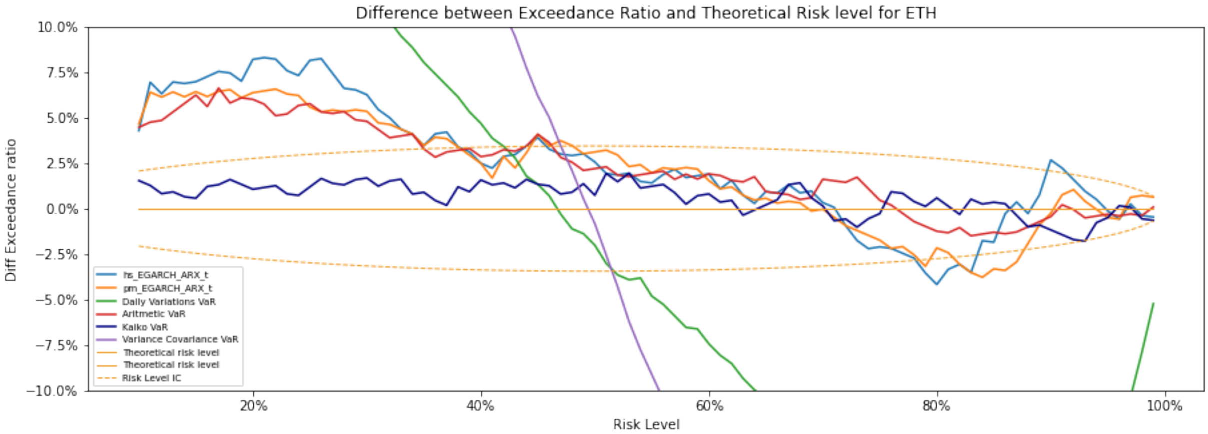

Figure 5.13: Difference between Exceedance Ratio and the theoretical risk level on ETH portfolio Backtesting using the 2020-2021 training data

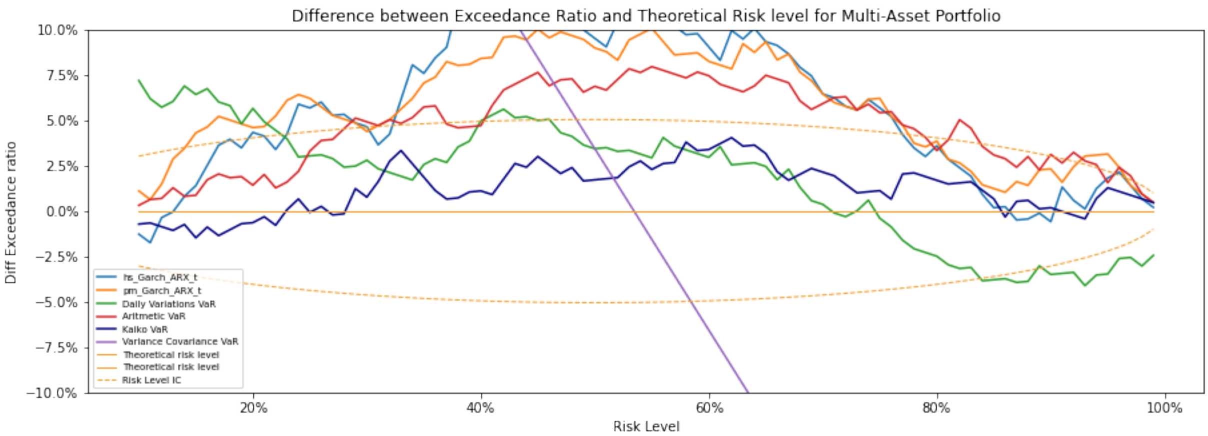

Figure 5.14: Difference between exceedance Ratio and theoritical risk level on multi-asset portfolio (DOT, LINK and SOL) Backtesting using the 2020-2021 training data

Besides, as the log-likelihood test results should be within the interval $[0, 6.63]$ to have an ideal forecast of the VaR. As we can see in the tables 5.2, 5.1, and 5.3, all the VaRs respect this condition except the daily variation and the variance covariance Value-at-Risks.

VaR 95	Exceedance Ratio (%)	ENQ ($\times 10^{-5}$)	Log-likelihood
Kaiko VaR	7.14	19	4.68
Arithmetic VaR	6.59	14	2.66
Daily variation VaR	15.75	100	86.87
HS EGARCH ARX t VaR	6.41	14	2.1
PM EGARCH ARX t VaR	6.04	9.73	1.17
Var Covar VaR	19.23	132	139.7

Table 5.2: Numerical backtesting results for different VaR 95 methods on a mono-asset portfolio (ETH).

VaR 95	Exceedance Ratio (%)	ENQ ($\times 10^{-5}$)	Log-likelihood
Kaiko VaR	5.03	0.4	125
Arithmetic VaR	4.23	-8	0.49
Daily variation VaR	14.65	52	27.86
HS EGARCH ARX t VaR	3.17	23	3.70
PM EGARCH ARX t VaR	3.70	-14	1.463
Var Covar VaR	24.603	0.0021	163

Table 5.3: Numerical backtesting results for different VaR 95 methods on a multi-asset portfolio (DOT, SOL and LINK)

Important result:

From the graph 5.15 where we only test the multi-assets portfolio containing BTC, ETH, XRP and LTC, it is extremely interesting to notice that when we use the one year training returns data from 2019 to 2020 to compute the different VaR instead of 2020 to 2021, the Kaiko approach seems to be the only one to catch the changing regimes since 2019. Besides, we can clearly see that even the Arithmetic VaR that seemed to work quite well is not adequate anymore whereas the Kaiko VaR remains around the exceedance ratio confidence level.

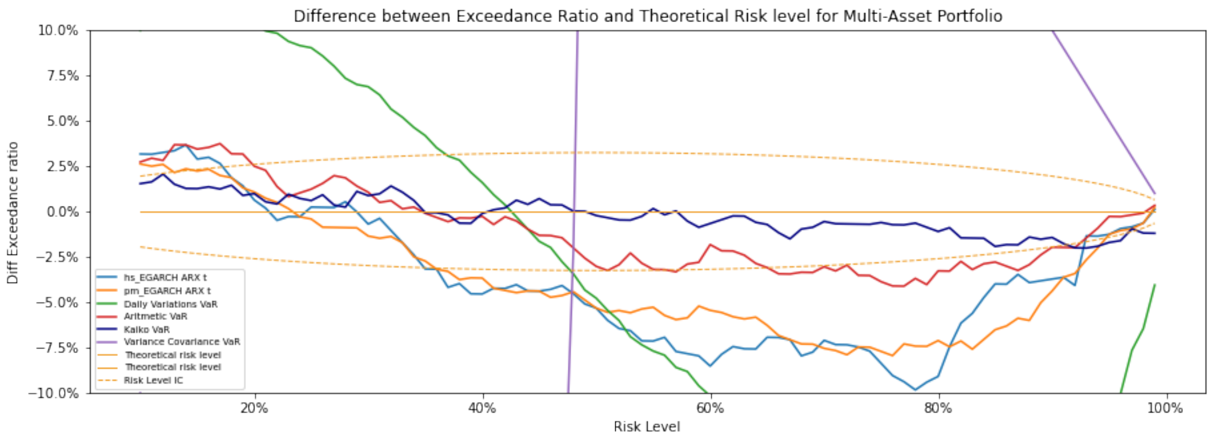

Figure 5.15: Difference between exceedance Ratio and theortical risk level on multi-asset portfolio (BTC, ETH, XRP and LTC) backtesting using the 2019-2020 training data

Chapter 6

Use Cases

6.1 Introduction

We use the Kaiko methodology to compute the daily 95% risk level Value-At-Risk from September 2021 to September 2022 because it captures an accurate value of the VaR compared to the other methods. Besides, we use nine crypto-currencies ETH, LTC, BTC, MATIC, XRP, DOT, SOL, ADA and LINK to create different portfolios.

In the crypto world, perhaps always, one of the biggest challenges facing investors is the choice overload. In fact, when we find ourselves in a situation where we are faced with a myriad of assets, investing becomes overwhelming and paralysing. Hence, in order to prevent a situation of doubt and bad feelings when picking them, the best solution is to let the math do the job for us. That's where the value-at-risk comes in hand.

Notation: We denote by P_j the j-th portfolio.

6.2 Use Case 1: Minimizing the Value-at-Risk for a better portfolio

Let's suppose we set the risk budget for a \$100k portfolio, and don't want to lose more than \$10k one in every 20 days on average. We would set the VaR risk level to 95%, charted in 6.1 with the daily losses, and monitor when the portfolio exceeds that \$10k level. When the portfolio VaR exceeds that level we would look to de-risk it so that the Value-at-Risk is under that \$10k threshold.

An interesting use of VaR would be to add different crypto-currencies on at a time to a potential portfolio with an initial investment of \$100k equally weighted between BTC/ETH for example, except now we're looking to add an extra asset but keeping the same initial investment of \$100k. The goal would be to identify the asset that allows us to minimise the value-at-risk without much change in the return rate defined in formula 6.2.

The steps are simple:

- We consider $100K$ invested in an equally weighted portfolio containing BTC and ETH.
- We fix a start and an end for daily VaRs and let m be the number of days between these two dates. In this case, m=365.

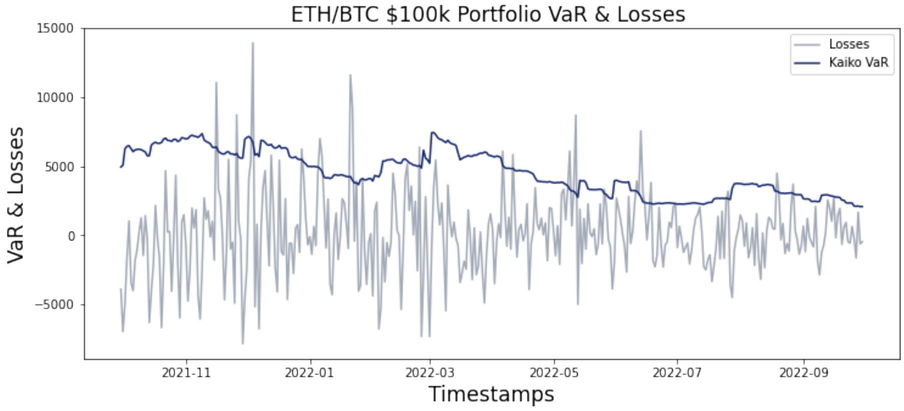

Figure 6.1: $100K Portfolio

- We compute the average VaR with $\alpha = 95$ risk level for the portfolio P_j (in this case a BTC/ETH equally weighted portfolio) and the average returns R^j_m of the portfolio P_j, i.e:

$$\hat{V}aR^{\alpha,j}_m := \frac{1}{m}\sum_{t=1}^{m} VaR^{\alpha,j}_t \tag{6.1}$$

$$r^j_t := \frac{V^j_t - V^j_{t-1}}{V^j_{t-1}} + 1$$

$$R^j_m := \frac{1}{m}\sum_{t=1}^{m} r^j_t \times 100 \tag{6.2}$$

  where $\hat{V}aR^{\alpha,j}_m$ is the average VaR of the portfolio P_j, $VaR^{\alpha,j}_t$ is the VaR of the portfolio P_j, R^j_m is the average return rate of the portfolio P_j.

- We create other $100k three-assets portfolios containing one new asset, additionally to BTC and ETH. Since we have seven cryptos left, we'll have seven other portfolios in addition to ETH/BTC portfolio.

- We compute the average VaR and the return rate of all three-assets portofolios.

- We box plot the average VaR-Portfolio graph and display the dataframe containing the average VaR and the return rates. See the graphs 6.3 and 6.4.

Firstly, in the graph 6.2, we see that all the new portfolios are correlated. Secondly, in the chart 6.3, we test the seven other assets in addition to our BTC/ETH portfolio to see which portfolio minimises the value-at-risk without changing the return rates. Thirdly, since the return rates are very close, we only focus on the average VaR. Finally, the latter drops from 7641 to 5958 as shown in the dataframe 6.4 for the portfolio containing the crypto ADA. Therefore, the portfolio ETH/BTC/ADA seems to be the best choice.

6.3 Use Case 2: What is the best portfolio?

In this use-case, instead of adding assets to an existing portfolios, a procedure that sometimes may not add much value in the risk management, we compute the average VaR and the returns rates of all the possible combination of multi-assets portfolios in order to get the best one. Furthermore, one can always think that for every return, there is a portfolio that

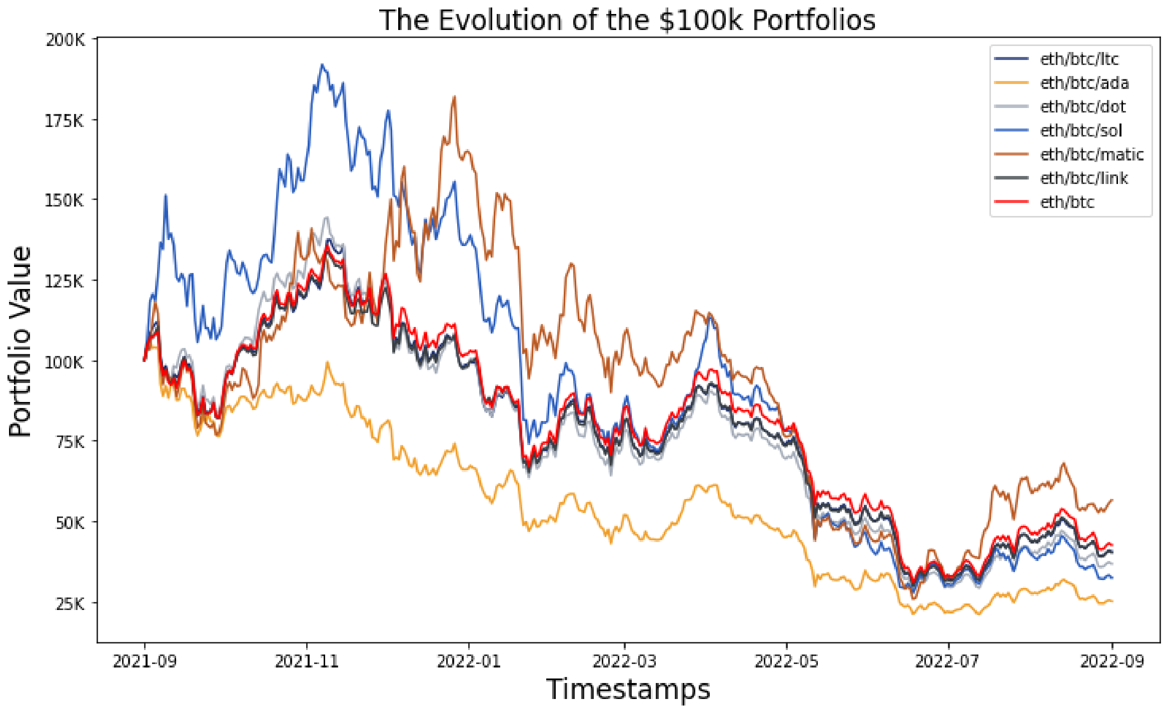

Figure 6.2: ETH/BTC \$ 100k portfolio evolution along with the new potential \$ 100k portfolios.

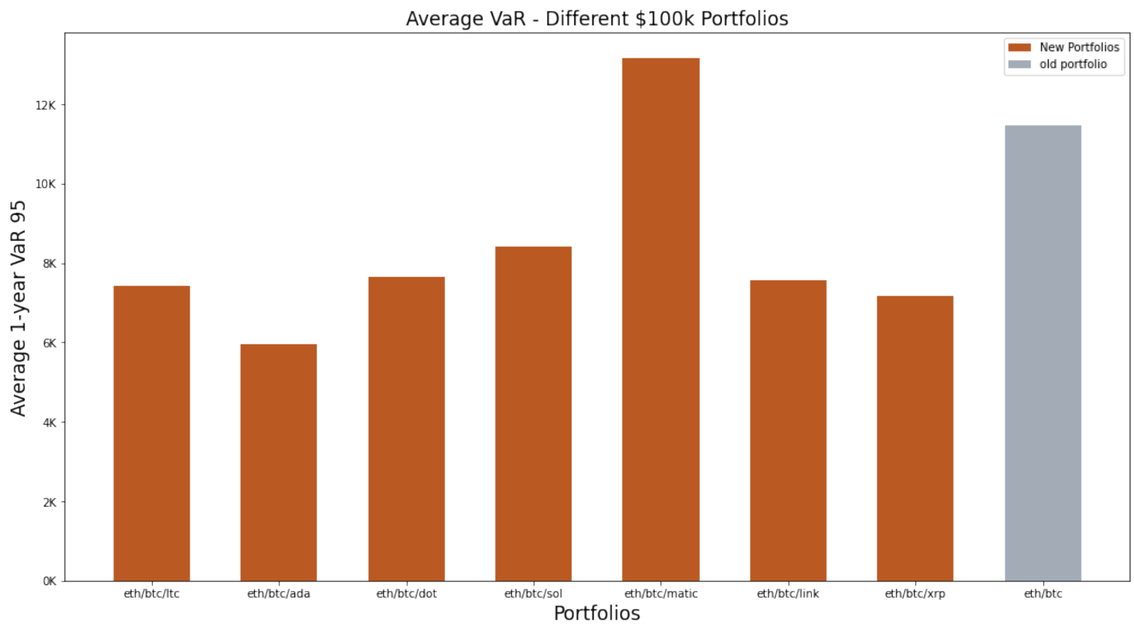

Figure 6.3: Impact of new assets on the $\$100k$ initial portfolio's VaR and return rate.

minimizes risk. Conversely, for each level of risk, one can find a portfolio that maximizes the expected return. This notion reminds us of the Markowitz efficient frontier in the traditional finance.

6.3.1 Equally weighted Portfolios

Let P be a set of all the possible combinations of the portfolios based on a list of nine cryptocurrencies. Thus, $\#P = 2^9 - 1 = 511$.

	return_rate	avg_var
eth/btc/ltc	99.839651	7409.158867
eth/btc/ada	99.738673	5958.782522
eth/btc/dot	99.806745	7659.141909
eth/btc/sol	99.759681	8413.466248
eth/btc/matic	100.012766	13159.716042
eth/btc/link	99.840441	7567.208773
eth/btc/xrp	99.856282	7177.940765
eth/btc	99.852164	7641.664663

Figure 6.4: DataFrame containing the return rates and the average VaR for each equally weighted portfolio.

Definition: Let the optimal surface be the set of the portfolios $(P_j)_{1\leq j\leq P}$ that verifies the following two inequalities:

$$\hat{V}aR_m^{\alpha,j} \leq VARLIMIT$$

$$R_m^j \geq RETURNLIMIT$$

$$VARLIMIT = \min_{P_j\in P}(VaR_m^{\alpha,j}) + \sigma_{\hat{V}aR_m} \tag{6.3}$$

$$RETURNLIMIT = \max_{P_j\in P}(\hat{R}_m^j) - \sigma_{\hat{R}_m^j} \tag{6.4}$$

where $\sigma_{\hat{V}aR_m}$ and $\sigma_{\hat{R}_m}$ are the standard deviation of respectively the average VaR and the return rates calculated using all the average var and average returns rates of all the combination of the portfolios.

In this use case, the steps are simple:

- We consider $100K$ invested in all possible combinations of portfolios based on a range of the nine crypto-currencies.
- We fix a portfolio P_j from the list of all the portfolios (511 portfolios).
- We fix a start and an end for daily VaRs and let m be the number of days between these two dates.
- We compute the average VaR $\hat{V}aR_m^{\alpha,j}$ with $\alpha = 95$ risk level in 6.1 and the average returns R_m^j of the portfolio P_j in formula 6.2.
- We scatter plot an average VaR-return rate graph 6.5.
- We extract the portfolios within the optimal surface.

Definition: We denote by V the couple $(VaR_m^{\alpha,j}, R_m^j)_j$ containing the average VaRs and the returns rates of all portfolios $(P_j)_{j\in\{1,\ldots,502\}}$. Let *Regulation on* V be the action of setting the optimal surface V. Two *Regulations* on V is doing one *Regulation* on V and then another one on $\tilde{V}$ a subset of V. This would minimise the optimal surface and have fewer optimal portfolios.

The challenging part about this study is the choice of the optimal surface and it all depends on the investor, whether they are adopting a risk prevention or a risk taking mentality. In other words, is the investor more interested on how much he could gain and thus focus more on the return rates ? Or is he more interested in how much he could loose by paying more attention on the value-at-risk ?

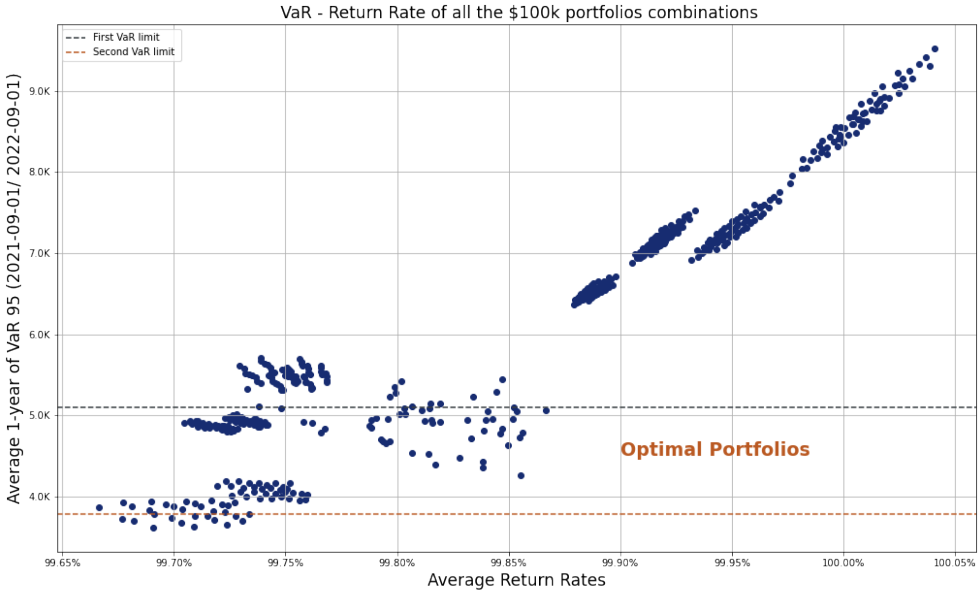

Figure 6.5: First view of the average VaR - Return rates scatter plot of all the possible $100k equally weighted portfolios (one *Regulation* was not enough as we have a lot of portfolios underneath the first var limit).

From the chart 6.6, we notice that all return rates are very close which makes this indicator uninteresting in terms of choosing the best portfolio. Besides, cryptos like DOT, ETH, SOL and MATIC are not included in the portfolios of the optimal surface. Finally, we conclude that the best portfolio is the equally weighted portfolio containing BTC and ADA that have the lowest average VaR that equals $3610.91.

6.3.2 Optimised Weighted Portfolios

In this use case, we optimise the weights of the portfolios instead of using equally weighted portfolios. In fact, one would think about minimising the VaR on the weights. However, the VaR is not convex (see the graph 6.7) with respect to the weights of the portfolio. Therefore, we can't choose the best weights of its assets by minimising the value-at-risk, i.e. the following expression 6.3.2 is difficult to assess:

$$\min_{w^i\in(0,1)} VaR_t^\alpha(X_{t+1}) = \min_{w^i\in(0,1)} V_t\left(q_t^\alpha\left(-\sum_{i=1}^n w^i \mathbf{AR}_{t+1}\right)+1\right), \quad \sum_{i=1}^n w^i = 1$$

However, we can think about optimising the asset's weight of a given portfolio using the Markowitz approach where we, instead, minimise the variance of the portfolio's arithmetic

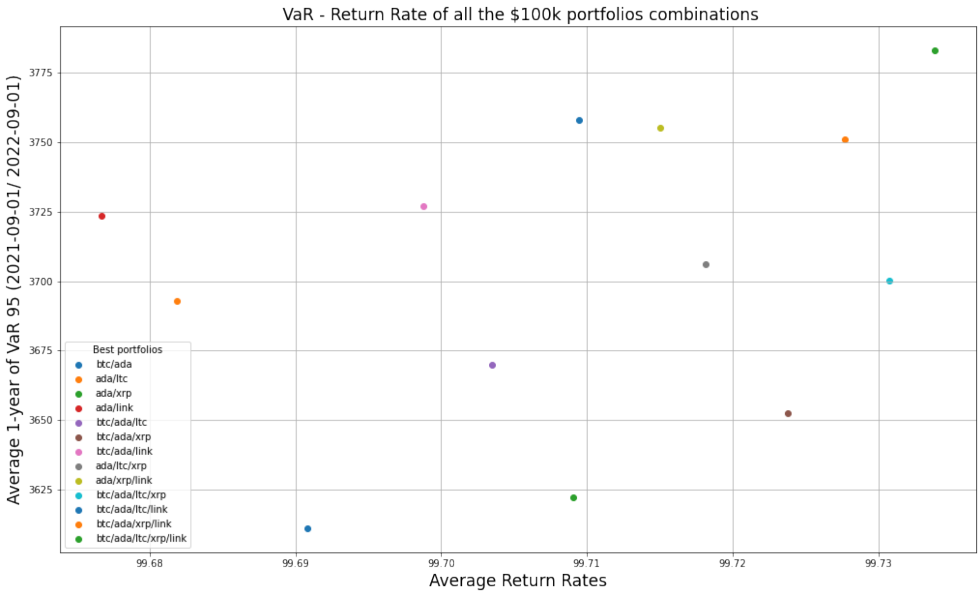

Figure 6.6: Close view of the optimal surface of the graph 6.5 after two regulations on the dataframe of the average VaRs.

returns $\mathbf{AR}_t$, i.e. $\min_{w_i \in (0,1), \sum_{i=1}^{n} w_i = 1} \mathbb{V}(\mathbf{AR}_t)$.

$$\mathbf{AR}_t = \sum_{i=1}^{n} w_i \mathbf{AR}_t^i \ , \qquad \sum_{i=1}^{n} w_i = 1, \qquad w_i \in (0,1) \tag{6.5}$$

$$\mathbb{V}(R_t) = \begin{pmatrix} w_1 \dots & w_n \end{pmatrix} \begin{pmatrix} \gamma_{11,t} & \gamma_{12,t} & \cdots & \gamma_{1n,t} \\ \ddots & \ddots & \cdots & \vdots \\ \gamma_{i1,t} & \gamma_{ii,t} & \ddots & \gamma_{ij,t} \\ \vdots & \cdots & \ddots & \vdots \\ \gamma_{n1,t} & \gamma_{n2,t} & \cdots & \gamma_{nn,t} \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \tag{6.6}$$

where $\mathbf{AR}_t^i$ is the arithmetic returns of the asset i, w_i its weight and n the number of the asset in the portfolio, $\gamma_{ij,t}$ is the covariance between arithmetic returns of the asset i and the asset j, i.e $\gamma_{ij,t} = Cov(\mathbf{AR}_t^i, \mathbf{AR}_t^j), \quad 0 \leq i \leq j \leq n$.

Remarks:

- We use $(0.001, 1)$ instead of $(0, 1)$ as the boundary of the weights to make sure to include all the possible portfolios. In fact, it is possible to be unable to minimise on the constraint $w^i \in (0, 1)$ and thus have $w^i = 0$ for a certain asset i. Obviously, the question that may be raised is whether or not we can consider an asset that has a weight of 1% negligible comparing the other assets. This is a probable approah left to the trader. Here the goal is only to cover every possible portfolio.

- In this case we only use the multi-assets portfolios because of the weight optimisation. Hence, we have $2^n - 1 - n = 502$, where n is the number of the existing assets (here $n = 9$).

- The steps are similar to the previous study. The only thing that is different is the weights of the portfolios. Here, instead of equally weighted portfolios, each portfolio P_j has its own optimised weight vector.

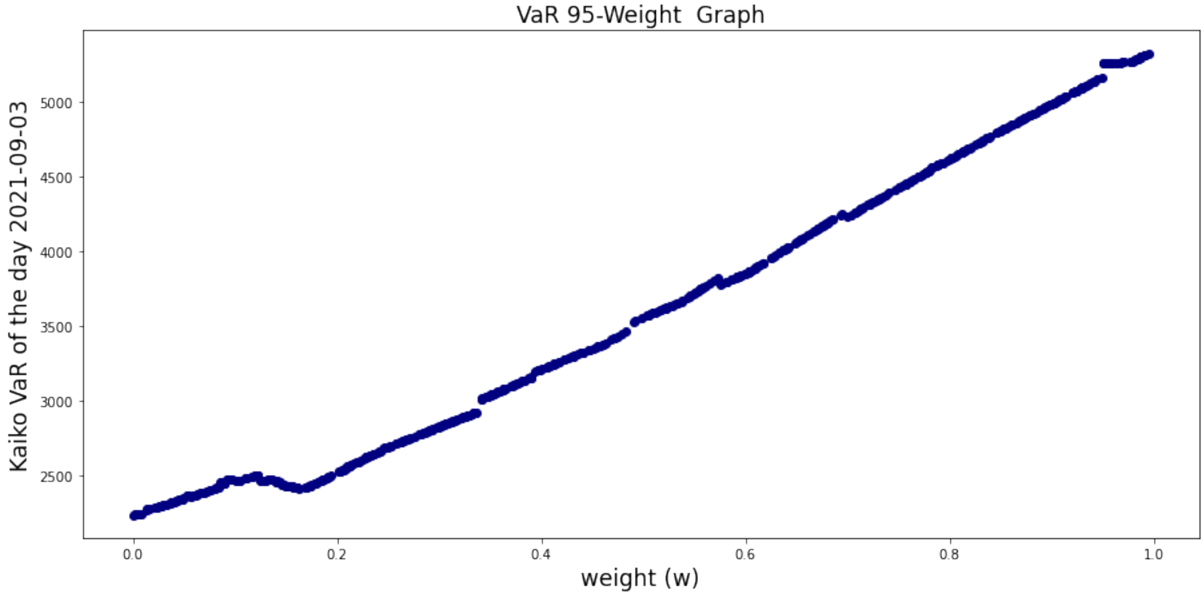

Figure 6.7: Here we display a Kaiko VaR with 95% risk level of the day 2021-09-03 using multi-asset portfolios containing ETH and BTC generated with different weights $(w, 1-w)$ where $w \in (0, 1)$. In fact, this is a counterexample for the non-convexity of the VaR with respect to the weight w.

From the numerical results of the returns rates in the dataframe 6.4 and in the graph 6.5, we conclude that the returns rates are not very appealing to the use case. Thus, we only use the $VARLIMIT$ in the formula 6.3 to set the optimal surface. Besides, it is necessary to do two or three *Regulations* to extract the best portfolios since we have a myriad of them (maximum of 30 out of 511 portfolios).

In the graph 6.10, we see that LTC, XRP and LINK are the most common cryptos in the optimal surface. Moreover, we notice that SOL is nowhere in the optimal portfolios and the crypto ETH is present in only one portfolio. Finally, the portfolio that has the lowest VaR is the one containing XRP (weight of 26%) and LINK (weight of 74 %). This is the best one.

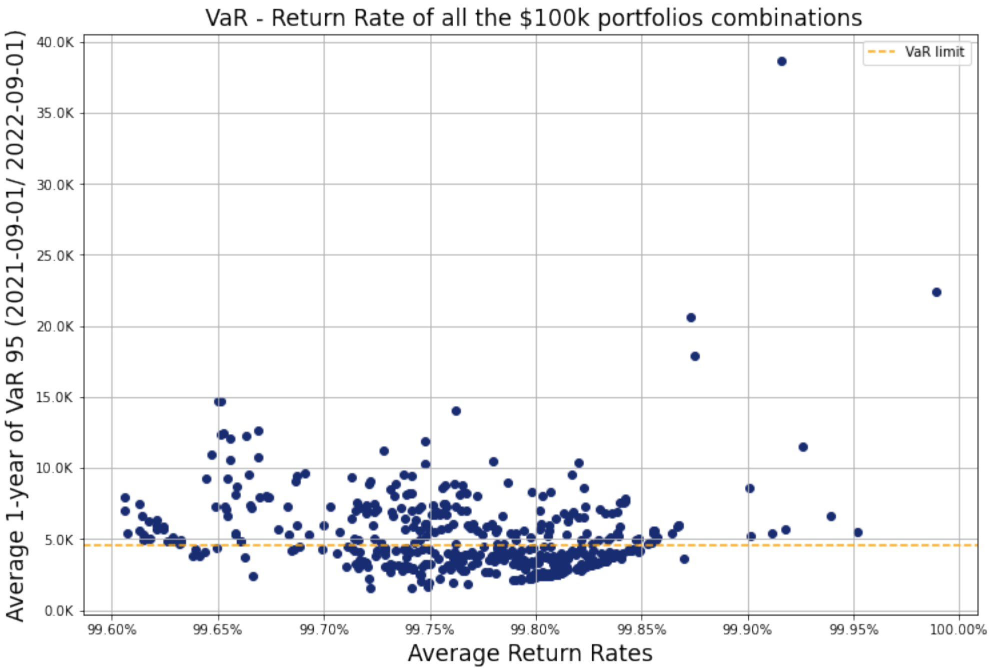

Figure 6.8: First view of the average VaR - Return rates graph of all the possible $100k weight optimised portfolios

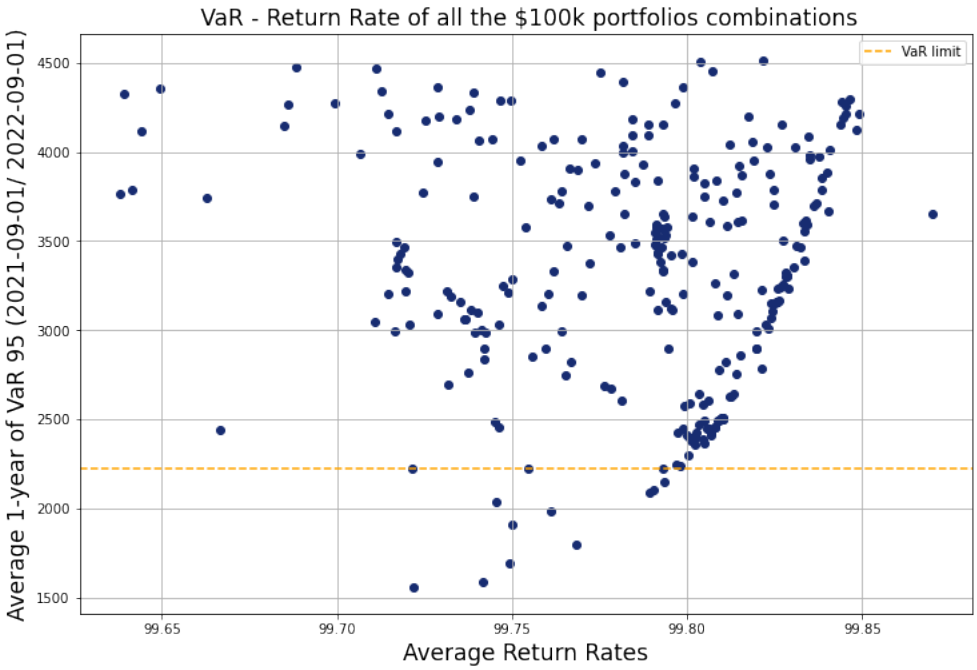

Figure 6.9: Second view of the average VaR - Return rates graph of all the possible $100k portfolios with optimised weights after another average VaR limit regulation.

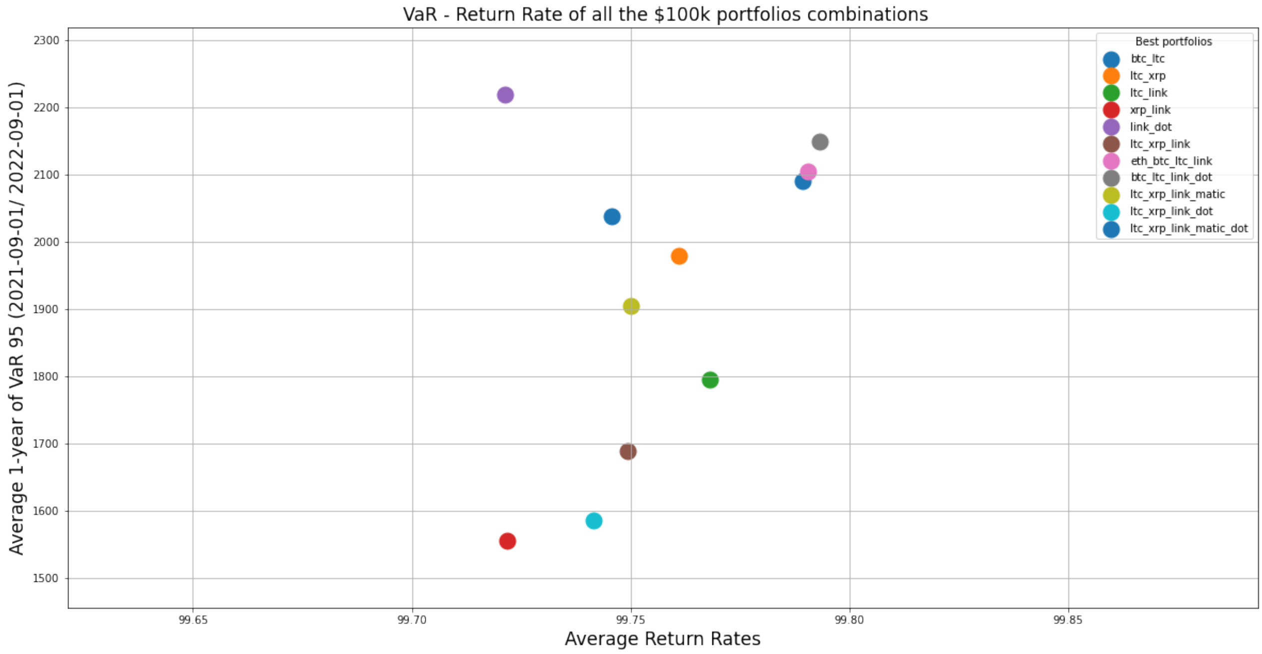

Figure 6.10: Close view on the optimal portfolios of the graph 6.9 in the average VaR - Return rates graph

	btc	ltc	xrp	link	dot	eth	matic
Portfolios							
btc_ltc	0.496211	0.503789	NaN	NaN	NaN	NaN	NaN
ltc_xrp	NaN	0.843694	0.156306	NaN	NaN	NaN	NaN
ltc_link	NaN	0.816825	NaN	0.183175	NaN	NaN	NaN
xrp_link	NaN	NaN	0.259512	0.740488	NaN	NaN	NaN
link_dot	NaN	NaN	NaN	0.639914	0.360086	NaN	NaN
ltc_xrp_link	NaN	0.532938	0.096279	0.370783	NaN	NaN	NaN
eth_btc_ltc_link	0.527052	0.470948	NaN	0.001000	NaN	0.001	NaN
btc_ltc_link_dot	0.606786	0.391214	NaN	0.001000	0.001000	NaN	NaN
ltc_xrp_link_matic	NaN	0.530570	0.237298	0.231131	NaN	NaN	0.001
ltc_xrp_link_dot	NaN	0.368378	0.094518	0.536103	0.001000	NaN	NaN
ltc_xrp_link_matic_dot	NaN	0.503263	0.150907	0.243477	0.101353	NaN	0.001

Figure 6.11: The optimised weights for the portofolios within the optima surface.

Chapter 7

Conclusion & Future Work

All things considered, we can say that despite the lack of the historical data for the back-tests, among all the VaR methodologies implemented, where most of them are slow to react to changing market environments, the Kaiko approach is the only one that captures an accurate value of the value-at-risk.

Furthermore, the study of different Value-at-Risk methodologies led us to conclude that the stabilisation of the returns is a mandatory step to have an ideal forecast of the VaR. In fact, a deep study of the returns of the crypto-currencies would allow us to better evaluate their risk and find a model to simulate them and eventually try another method to compute the VaR such as the Monte Carlo approach or the copula method. This latter has been studied during a certain period of my internship but was not prioritized and could be something to do in the future. In addition to that, another possible future work would be to compute other risk metrics such as the expected shortfall also so-called the conditional value-at-risk (CVaR).

References

[AJME15] Rudiger Frey Alexander J. McNeil and Paul Embrechts. Quantitative risk management: concepts, techniques and tools. page 408., 2015.

[Chr98] Peter F Christoffersen. Evaluating Interval Forecasts. *International Economic Review*, 39(4):841–862, November 1998.

[FZ10] C. Francq and J.M.. Zakoian. ARCH Models: Structure, Statistical Inference and Financial Applications. John Wiley Sons Ltd., Chichester. page 504., 2010.

[LR] Lehmann and Romano. Testing Statistical Hypothesis. page 794.

[Sto11] Pavel Stoimenov. Philippe Jorion, Value at Risk, 3rd Ed: The New Benchmark for Managing Financial Risk. *Statistical Papers*, 52(3):737–738, August 2011.

[WIN] Olivier WINTENBERGER. Time Series Analysis. page 67.