UCLouvain

LOUVAIN School of Management

Louvain School of Management

Extreme risk in portfolio selection: model comparison and analysis

Author: Walter Delava
Supervisor: Prof. Nathan Lassance
Academic year 2023-2024
Dissertation for the master 120 in Business Engineering
Financial Engineering
Daytime schedule

Declaration

During the preparation of this master's thesis, the author(s) utilized OpenAI/ChatGPT for the following purpose:

1.Code implementation: The AI tool was used to gain efficiency and time when implementing the code in R software for this master's thesis.

2. Layout: The AI tool was used to save time on the layout of graphs and tables.

3. Examples: The AI tool was used only as a basis for inspiration.

After using OpenAI/ChatGPT the author(s) diligently reviewed and edited the content produced by the tool. We take full responsibility for the final content presented in this thesis.

By signing this declaration, we affirm that the content of this master's thesis reflects our original work, augmented by the responsible use of AI.

Signed by: Walter Delava (19/05/2024)

Abstract

This master's thesis applies Extreme Value Theory (EVT) to risk management in financial markets, comparing traditional Value-at-Risk (VaR) models with EVT-based models. Using major stock indices such as BEL20, CAC40, S&P500, and MSCI World, the analysis examines the models' ability to predict extreme losses. Findings indicate that EVT models are more robust in identifying extreme events and estimating maximum losses. The results support the use of diverse financial data and multiple quantitative approaches for more accurate risk predictions. This study highlights the practical implications of EVT in enhancing financial risk management.

Acknowledgments

Above all, I would like to thank my parents for always supporting me and giving me the best possible environment to progress in my university studies.

I would also like to thank my promoter Prof. Nathan Lassance for his availability and his invaluable advice which guided me throughout the writing of this master's thesis.

Contents

List of Abbreviations

- VAR : Value-at-Risk
- ES : Expected Shortfall
- EVT : Extreme Value Theory
- GEV : Generalized Extreme Value
- GPD: Generalized Pareto Distribution
- BMM Block-Maxima Model
- POT : Peak-Over Threshold
- EWMA : Exponential Weighted Moving Average
- MLE : Maximum Likelihood Estimation

List of Tables

List of Figures

1. Introduction

Risk is a concept that affects all of us, in one way or another, in our daily lives, and can take an infinite number of forms. In the meantime, extreme risk, is a form of risk whose frequency is small but whose consequences can be enormous. Wars, climatic catastrophes, political tensions, etc. are all extreme risks that are almost impossible to predict and have strong effects on our economy. Financial markets, which are reflecting the state of our society severely react to those extreme events. In this master's thesis, we will reflect on the prediction of these uncommon risks, trying to quantify their potential negatives impacts by presenting the tools at our disposal to deal with these risks.

One question we might ask ourselves is whether traditional methods of estimating risk perform as well on the financial markets as the so-called "extreme prediction models? As we have seen in history, financial markets are very sensitive to extreme situations. This has motivated the research in this thesis to find out whether extreme prediction models do better than conventional methods. What are the limits of these models when applied to real data? How can we simply apply these models and understand the results obtained to apply them to financial management?

In the first part of this master's thesis, we will present the theory needed to understand the application of these models to real-life situations. We'll start by presenting risk in general, so that we have a common grounding in the subject before diving more specifically into financial risk. Next, we will define the notion of risk measures, which are essential to understand risk in finance and their associated properties. We will also introduce the extreme values occurring in the tails of distributions as this concept will help understand that these distributions zones will play in the various models studied. To end the first section, the theory of Value-at-Risk (VaR) will be covered with its associated Expected Shortfall (ES).

The second part will present the main theme of this master's thesis, namely the Extreme Value Theory (EVT), which was first introduced by Emil Gumbel in 1938. This theory aimed to explain extreme events observed during a certain defined period and to predict future extreme events supposed to occur in a certain sample size (Gumbel, 1958). We will also present the distributions associated with the EVT, the Generalized Extreme Value distribution (GEV) and the Generalized Pareto Distribution (GPD). The two models that are applied thanks to these distributions will be covered. On one hand, the GEV-based model, known as the Block-Maxima Model (BMM), involves extracting the maximum values from sub-periods of similar time, which are pre-defined. This extraction process is used to estimate the GEV distribution. On the other hand, the second model, called the Peak-Over Threshold (POT), relies on the GPD. In the POT model, extreme events are identified as data points exceeding a certain predefined threshold. In both models, VaR is estimated based on historical data to provide the best possible forecasts.

In this third and final part, we are going to apply these models to our four selected indices (BEL20, CAC40, S&P500 and MSCI World) to extract the desired results for the best possible analysis. This analysis will be based firstly on the historical VaR method using the normal distribution. Subsequently, we will explore the Exponential Weighted Moving Average Method (EWMA) model and the two EVT models: BMM and POT.

Following this analysis, the results obtained show that the models based on Extreme Value Theory perform better overall than the traditional models. The objective for these models was to come as close as possible to the percentage of excess theoretical VaR in relation to the VaR confidence level. For both levels, the Peak Over Threshold model performed best, being close to 1.2% and 5% for the 99% and 95% VaR respectively, which is desirable. Behind this, we had the Block-Maxima model with very different results compared with POT, for an estimate with a 99% confidence level gives results of around 2%, which is twice more than the theory. Next to this, the classical methods gave surprising good results with the normally distributed VaR close to 5% for the 95% VaR but fails for the 99% confidence level with results close to 2.5%. As for the last method, the Exponential Weighted Moving Average, it was not at all convincing, with results twice or even four times higher than the theoretical results, which completely excludes it from this ranking. These results show that at the 95% confidence level, normally distributed VaR estimation performs as well as the POT model, but at the 99% confidence level, EVT models appear to be useful.

This project aims to contribute to a comprehensive comparison of traditional and EVT-based models and to highlight the practical implications of models predicting extreme events. It also aims to allow people to project themselves by integrating these models with the financial data that everyone may encounter in their everyday life and to develop their understanding of the subject.

2. Fundamentals of Financial Risks

This chapter explores the broad and complex topic of portfolio selection, incorporating the nuanced issue of extreme risks. Over recent years, this subject has captivated a lot of arrays of professionals from the financial sector as well as mathematicians. It presents a delicate yet fascinating challenge that has stimulated extensive debate and research. As we explore this theme, we aim to engage with various perspectives and analyses to understand better the complexities involved in portfolio selection under conditions of extreme financial risks. This introduction aims to not only enrich our theoretical knowledge but also to enhance practical approaches to managing financial portfolios in unpredictable markets.

2.1 Introduction to risk

In the realm of finance, the concept of risk is intrinsically linked to the uncertainty surrounding the future outcomes of decisions made by individuals and organisations. This uncertainty primarily concerns the potential for outcomes that result in financial losses. Actors within the financial sector are particularly focused on strategies to minimise this risk as much as possible. Their aim is to limit exposure to situations where decisions could lead to significant financial detriment.

According to McNeil et al. (2005), risk can be defined from two perspectives. One definition view risk as "any event or action that may adversely affect an organisation's ability to achieve its objectives and execute its strategies." This encapsulates a broad spectrum of risks that can derail an organisation from its planned path. Alternatively, risk is described as "the quantifiable likelihood of loss or less-than-expected returns." This definition emphasises the measurable aspect of risk, allowing organisations to quantify their risk exposure in financial terms.

However, these authors also note a critical limitation in these definitions: they do not comprehensively address all potential contexts. This suggests that while the provided definitions are robust in many scenarios, they may fall short in capturing the nuances of risk in every financial situation. Such an acknowledgment points to the complexity and multifaceted nature of risk in financial environments.

2.2 Classification of Financial Risks

Understanding various types of financial risks is crucial before delving into the methodologies used for measuring these risks. Financial risks, as described in Measuring Market Risk Dowd (2005), refer to the potential for financial loss—or gain—stemming from unexpected changes in underlying risk factors. In the financial sector, institutions encounter several primary types of risk throughout their operational lifespan:

- **Market risk:** as defined by Kevin Dowd in Measuring Market Risk, is the potential for loss (or gain) that arises from unexpected changes in the prices of securities or market rates, such as interest rates or exchange rates. This type of risk is an inherent aspect of participating in financial markets. Due to its inevitability, market participants invest considerable efforts to predict and mitigate its impacts as accurately as possible.

- **Credit risk:** can be succinctly defined as the possibility that a counterparty will fail to meet its obligations as per the agreed terms, (Brown & Moles, 2014). A prominent example of the severe impact of credit risk is the subprime mortgage crisis in the United States. During this crisis, many borrowers found themselves unable to repay their loans, leading to widespread defaults. This situation escalated as the inability of borrowers to meet loan obligations resulted in significant losses for financial institutions and ultimately triggered a market crash. The subprime crisis exemplifies how high credit risk can lead to profound economic disruptions, underscoring the importance of robust credit risk management practices.

- **Operational risk:** is defined by the Basel Committee on Banking Supervision as "the risk of loss resulting from inadequate or failed internal processes, people, and systems, or from external events."(BIS, 2021) This encompasses a broad range of potential issues that can occur within an organisation, including but not limited to human error, system failures, fraud, and disruptions caused by external events such as natural disasters or cyber-attacks.

- **Liquidity risk:** Liquidity risk involves the difficulty of converting assets into cash quickly without significant loss in value. It encompasses three key dimensions, as outlined in the Financial Stability Review (Alentorn & Markose, 2008):

    1. Instrument Liquidity: The ease with which financial instruments can be exchanged for cash while preserving their value.
    2. Market Liquidity: The market's capacity to absorb large volumes of trades without significantly affecting asset prices.

3. Monetary Liquidity: The amount of readily available liquid assets in the economy, typically measured against nominal GDP.

This type of risk can critically affect an entity's ability to meet its financial obligations, particularly in scenarios where liquidating assets at fair market values is challenging. Effective liquidity management is therefore essential to maintain financial stability and ensure continuous operations under various market conditions.

2.3 Quantitive Risk Measures

Now that we have defined risk, we will define risk measures and their mathematical foundations. We will review the properties that allow a risk measure to be considered coherent, including translation invariance, subadditivity, positive homogeneity, and monotonicity. For robust and reliable financial risk management, these properties must be respected. In addition, we will present tail risk, which models the severity of portfolio losses over specific periods. Understanding these concepts is crucial for effective risk management and prepares us to discuss more extreme risk scenarios.

2.4 Defining Risk Measures

We explore the definition of a risk measure and its underlying mathematical principles. Initially, let us assume that while the set of all possible states of the world at the end of a given period is known, the probabilities of events occurring within this set may not be known or universally agreed upon. To draw a parallel with market risk, imagine that the state of the world is represented by a list of prices for all securities and exchange rates. We also assume that the full range of potential outcomes for this list is known.

In the study of financial risk measures, we start by defining key mathematical elements as follows:

- Ω: A set that can be infinite, representing different states of nature, particularly when dealing with a continuous random variable that can assume any value within a specific range.
- X: A random variable describing the final net worth of a financial position taking values in the set Ω.
- G: A set of all risk measures, expressed as real-valued functions on Ω.

Definition of a Risk Measure: Formally, a risk measure is a mapping from the set G into the real numbers R. In financial terms, the value $\rho(X)$ assigned by the measure ρ to a risk X represents the minimum additional cash that an agent must add to their risk position X to make it viable for proceeding with their investment plan. When $\rho(X)$ is negative, it indicates a possible withdrawal of cash from the position.

2.5 Fundamentals Properties

A risk measure must satisfy certain properties to be considered "coherent." These properties ensure that the risk measure is robust and reliable for financial risk management. The axioms, as presented in the paper (Artzner et al., 1999), align closely with the axioms of the acceptance set. Here, we detail these axioms with their mathematical definitions and practical implications:

- Translation Invariance:

  Definition: For any portfolio X in G and any real number α, the equation $\rho(X + \alpha) = \rho(X) - \alpha$ holds.

  Implication: Adding cash to a portfolio reduces its risk by the same amount, reflecting the risk-free nature of cash. This property ensures that the risk measure accurately adjusts for cash infusions or withdrawals.

- Subadditivity:

  Definition: For any two portfolios X_1 and X_2 in G, the inequality $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$ is satisfied.

  Implication: The risk of a combined portfolio should not exceed the sum of the risks of the individual portfolios. This property encourages diversification, as combining portfolios does not increase overall risk.

- Positive Homogeneity:

  Definition: For any portfolio X in G and any positive scalar λ, $\rho(\lambda X) = \lambda\rho(X)$.

  Implication: If an investment's position is scaled up, the risk measure scales proportionally. This ensures that risk measures are consistent across different sizes of the same investment.

- Monotonicity:

  Definition: For any two portfolios X and Y in G where $X \leq Y$, the inequality $\rho(Y) \geq \rho(X)$ holds.

  Implication: If one portfolio consistently incurs greater or equal losses compared to another, its risk measure should also be greater. This axiom assures that the risk measure respects the relative performance of portfolios.

These axioms are fundamental to ensuring that any risk measure is coherent and can be effectively applied in risk management scenarios observed in the financial markets. By understanding these principles, we lay the groundwork for discussing more specific and extreme cases of risk in portfolio management.

2.6 Tail Risk

When addressing risk measurement, it's essential to define a statistic that assesses the severity of holding a portfolio over a specified period based on the loss distribution $F_L(l)$ (McNeil et al., 2005). Portfolio returns can extend significantly into the tails, which necessitates that any effective risk measure should account for tail-risk and the potential skewness of the distribution.

2.7 Conclusion

This section has outlined the fundamental properties that a risk measure must satisfy to be considered coherent, namely translation invariance, subadditivity, positive homogeneity, and monotonicity. These properties ensure that the risk measure is robust and reliable for financial risk management, providing a solid theoretical foundation for assessing and managing financial risks.

Understanding these principles is crucial because they form the basis for more advanced risk measures. For example, VaR, a widely used risk measure, is not coherent because it fails to satisfy subadditivity, which is a widely risk to underestimating risk when diversifying portfolios. This limitation has led to the adoption of Expected Shortfall, which is coherent and provides a more accurate reflection of extreme risks.

The insights gained from this section are essential for the subsequent analysis and application of coherent risk measures in portfolio management, ensuring more reliable and effective risk assessment in financial markets.

3. Risk Indicators in Finance

Quantifying potential losses is crucial for making informed investment decisions and ensuring regulatory compliance. Two primary measures dominate this field: VaR and ES. Both tools serve to estimate potential financial losses under normal and extreme market conditions, yet they differ significantly in their approach and implications. This section goes further into the definitions, methodologies, advantages, and limitations of VaR and ES, illustrating their roles in contemporary risk management practices. By examining these measures, we aim to understand their practical applications and their importance in managing financial portfolios effectively.

3.1 Value at Risk

We are now going to tackle one of the best-known forms of risk measurement, VaR. This measure is crucial in the management of financial crises and will therefore interest us in the measurement of extreme events. VaR has been widely adopted to assess the risk of potential losses in portfolios. For example, the Basel Committee on Banking Supervision uses VaR as a basis for its regulations on banks' capital requirements, specifically within the frameworks of Basel II and Basel III. These agreements require banks to maintain a minimum capital based on a VaR estimate to cover potential losses due to market fluctuations (Basel Committee on Banking Supervision, 2004, 2011). The VaR measurement aims to estimate the maximum potential loss of a portfolio over a specified period, given a certain confidence level. However, in this context, we are analysing the distribution of returns rather than losses.

The probability density of returns allows us to visualise the frequency of different return levels observed. With the density plot of returns, we can identify the VaR threshold as the quantile below which a certain proportion of returns falls. For example, if we consider a VaR at a 95% confidence level, it means that 95% of the returns will be above this threshold, and only 5% of the returns will be below this value. This approach enables us to understand the distribution of returns and assess the risk of negative returns within the portfolio. (Holton, 2003)

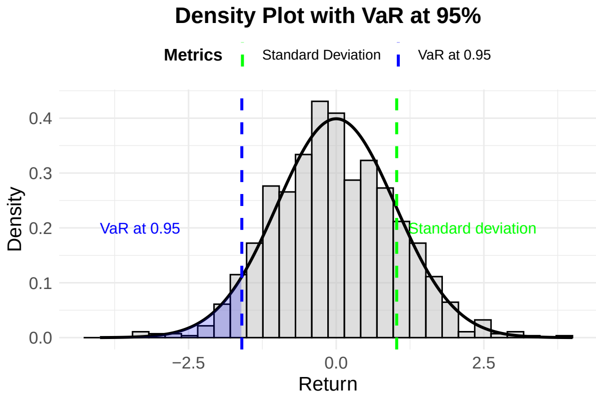


Figure 3.1: Density Plot with VaR at 95% Confidence Level

If we look mathematically to the VaR, the probability formula associated with it is given by:

$$P(L \leq \mathrm{VaR}_\alpha) = \alpha, \tag{3.1}$$

where:

- L represents the portfolio loss.
- VaR_α is the Value-at-Risk at the confidence level α.
- α is the confidence level (e.g., 95% or 99%).

In other words, the VaR at the confidence level α is the loss threshold such that the probability of the loss L being less than or equal to this threshold is α.

VaR is a straightforward extension of maximum loss, which takes these criticisms into account. The idea is simply to replace "maximum loss" by "maximum loss which is not exceeded with a given high probability", the so-called confidence level (McNeil et al., 2005).

As we saw in the last section with the coherence points for a risk measure, we can apply them to the VaR. VaR is a coherent risk measure if it satisfies all the following conditions: Translation Invariance, Homogeneity, Monotonicity, and Subadditivity (Artzner et al., 1999).

VaR satisfies the first three conditions: Translation Invariance, Homogeneity, Monotonicity. However, VaR fails the subadditivity criterion. This means that the risk of a combined portfolio can be greater than the sum of the risks of individual portfolios, which is particularly problematic for diversified portfolios. VaR also "completely ignores the severity of losses in the far tail of the loss distribution" (Emmer et al., 2015), making it less reliable for comprehensive risk management.

The VaR can be calculated in various ways, and the following are some common methods:

1. **Historical Simulation**: Uses historical market data to simulate potential losses.
2. **Variance-Covariance Approach**: Assumes normal distribution of returns and uses the mean and standard deviation of portfolio returns.
3. **Monte Carlo Simulation**: Uses random sampling to generate a wide range of possible outcomes based on the statistical properties of the portfolio.

For our analysis, we will use the variance-covariance method. Comparing other models will enable us to see if there is any difference with a method that assumes a normal distribution for returns. We might think, then, that this method will return poor results because it assumes no extreme distribution in the distribution tails, which a priori will not happen with our data. Parametric methods, such as variance-covariance, can underestimate these risks due to their distribution assumptions.

3.2 Expected Shortfall (ES)

ES, also known as Conditional Value-at-Risk (CVaR), is a risk measure that addresses some of the limitations of VaR. The ES benefits from the fact that it respects all the axioms of coherence. It can therefore be used to estimate the total average loss in the worst-case scenario where the VaR threshold is exceeded (BCBS, 2019b)

To have an accurate estimation of the ES, the distribution of the tail must be precise and reflect the reality of the situation. As the threshold of the VaR is defined, the ES can be the expected loss beyond that threshold. We can define the ES as:

$$ES_\alpha(L) = E(L \mid L \geq VaR_\alpha(L)) \tag{3.2}$$

where L is the amount of money and α is the confidence level.

ES is given by the integral and for a loss L and a CDF F_L, the expected shortfall at confidence level α is defined as:

$$ES_\alpha(L) = \frac{1}{1-\alpha} \int_\alpha^1 VaR_\gamma(L)\, d\gamma \tag{3.3}$$

This plot shows the link between VaR and ES:

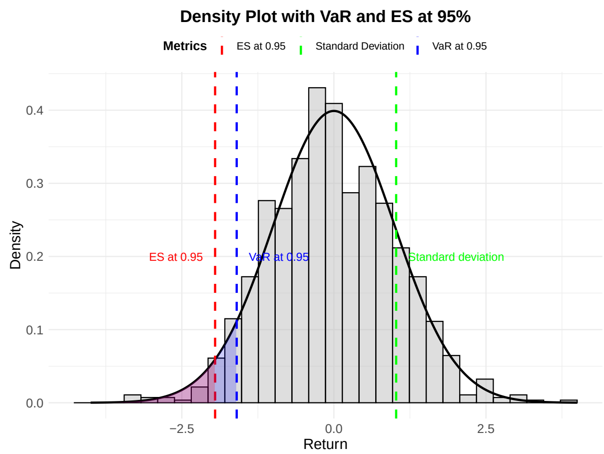


Figure 3.2: Density Plot with VaR and Expected Shortfall (ES) at 95% Confidence Level

3.3 Conclusion

In this section, we have focused on two fundamental risk measures: VaR and ES. These measures are essential for quantifying potential financial losses and play a crucial role in risk management.

VaR estimates the maximum potential loss over a specified period at a given confidence level. It is widely used for its simplicity and intuitive appeal; it has notable limitations. VaR fails the subadditivity criterion, meaning it can underestimate the risk of diversified portfolios. It also ignores the severity of losses beyond the VaR threshold, making it less reliable for comprehensive risk management.

ES, also known as CVaR, addresses these limitations by considering the average loss in the worst-case scenarios beyond the VaR threshold. ES is a coherent risk measure, satisfying all necessary properties such as subadditivity, which ensures it provides a more accurate and reliable assessment of risk.

Linking these measures with EVT, which we will see in the next chapter, will provide a robust framework for modeling and managing extreme risks, ensuring a better prediction of extreme events with our financial data.

4. Extreme Value Theory (EVT)

4.1 Classic EVT and different methods

The aim of this section is to explain the fundamental theory of EVT. We will review the asymptotic model of Block Maxima, the theory of how to estimate the distribution of maxima, and their return levels. We will then review the Peak Over Threshold model with an explanation of the theory. This chapter is based on the work of Stuart Coles in "An Introduction to Statistical Modeling of Extreme Values", Chapter 3.

The basic model of extreme value theory developed in this chapter focuses on the statistical behavior of extreme values.

The EVT model involves analysing the maximum values of a sequence of independent and identically distributed (i.i.d.) random variables. Specifically, for a sequence $X_1, X_2, \ldots, X_n$, the maximum value M_n is defined as:

$$M_n = \max\{X_1, \ldots, X_n\} \tag{4.1}$$

From equation 4.1, the distribution of M_n can be derived exactly for all values of n (Coles, 2001):

$$\begin{aligned}\Pr\{M_n \le z\} &= \Pr\{X_1 \le z, \ldots, X_n \le z\} \\ &= \Pr\{X_1 \le z\} \times \cdots \times \Pr\{X_n \le z\} \\ &= F(z)^n.\end{aligned} \tag{4.2}$$

The block maxima approach divides the dataset into blocks of equal size and takes the maximum value within each block. According to the Fisher-Tippett-Gnedenko theorem, the distribution of these maxima converges to a Generalized Extreme Value (GEV) distribution as the block size increases, which can take the form of Gumbel, Fréchet, or Weibull distributions (Coles, 2001).

In most cases, the distribution F is not known, and this equation is only valid if the distribution F is known. We accept that the distribution of F is unknown and instead arrive at a form that can be estimated based solely on the extreme data. This approach is analogous to approximating the distribution of sample means by the normal distribution (Coles, 2001), which is an extreme analogue of the central limit theorem (Embrechts et al., 1997; de Haan & Ferreira, 2006).

By using two constants $a_n > 0$ and b_n, the form of the maxima can be linearly reformulated. These constants are sequences that normalise the distribution of maxima, ensuring that the scaled maxima converge to a non-degenerate distribution. Specifically, the constants a_n and b_n are chosen such that:

$$M_n^* = \frac{M_n - b_n}{a_n}, \quad a_n > 0 \tag{4.3}$$

where M_n^* is the normalized maxima. The sequences $\{a_n\}$ and $\{b_n\}$ depend on the underlying distribution F and are determined through the Fisher-Tippett-Gnedenko theorem, which guarantees that the normalised maxima converge in distribution to a Generalized Extreme Value (GEV) distribution (Coles, 2001).

Theorem. If there exist sequences of constants $\{a_n > 0\}$ and $\{b_n\}$ such that:

$$\Pr\left\{\frac{M_n - b_n}{a_n} \leq z\right\} \to G(z), \quad as \quad n \to \infty \tag{4.4}$$

where G is a non-degenerate distribution function, then G belongs to one of the following families:

I. **Gumbel Distribution**

$$G(z) = \exp\{-\exp[-(az - b)]\}, \quad -\infty < z < \infty \tag{4.5}$$

II. **Fréchet Distribution**

$$G(z) = \begin{cases} 0, & z \leq b \\ \exp\{-(az - b)^{-\alpha}\}, & z > b \end{cases} \tag{4.6}$$

III. **Weibull Distribution**

$$G(z) = \begin{cases} \exp\{-[(az - b)^{\alpha}]\}, & z < b \\ 1, & z \geq b \end{cases} \tag{4.7}$$

This approach to extreme value theory aims to use asymptotic models to estimate the M_n^* distribution functions and to determine which of the three families of distribution functions is the closest. The Gumbel, Fréchet, or Weibull distributions are essential for modeling the behavior of the tail. Having these three families of distributions at one's disposal makes it possible, in many situations, to choose the one that is most appropriate to the situation. In the early days of EVT, it was common to adopt one of these three families and estimate the parameters of that distribution. This technique has its weaknesses, because you must choose the right one at the outset and then assume that this choice is the best. This is why the GEV distribution combined these three distributions (Coles, 2001; Embrechts et al., 1997; de Haan & Ferreira, 2006).

4.2 Generalized Extreme Value Distribution (GEV)

The GEV distribution is a unified model that encompasses the Gumbel, Fréchet, and Weibull distributions. By combining these three distributions into a single framework, the GEV distribution allows for more flexibility and avoids the need to make a preliminary choice of the specific distribution family. This makes it more robust and widely applicable in different scenarios (Coles, 2001).

In particular, GEV makes it possible to identify three distinct types of limits for estimating the behaviour of the tails of distributions represented by the values X_i of function F. As we saw earlier, these are illustrated by the families of distributions (Gumbel, Frechet, and Weibull).

This distribution is formulated as follows:

$$G(z) = \exp\left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\} \tag{4.8}$$

with the parameters:

- μ is the location parameter,
- σ is the scale parameter,
- ξ is the shape parameter.

This equation, which unifies the three families, is defined on the set:

$$\text{z} : 1 + \xi(\text{z - } \mu)/\sigma \quad 0$$

where the parameters satisfy: $-\infty < \mu < \infty$, $\sigma > 0$, and $-\infty < \xi < \infty$.

The value of ξ determines which type of distribution the GEV corresponds to:

- When $\xi = 0$, the GEV distribution reduces to the Gumbel distribution more centered.
- When $\xi > 0$, the GEV distribution corresponds to the Fréchet distribution and assume a fat tail on the right side of the distribution (positive values).
- When $\xi < 0$, the GEV distribution corresponds to the Weibull distribution and assume a fat tail on the left side of the distribution (negative values).

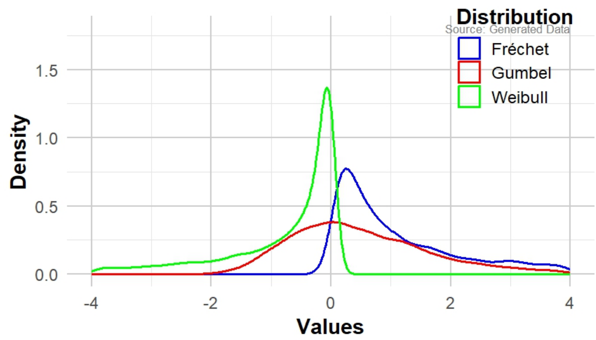


Figure 4.1: Distribution Functions of Gumbel, Fréchet, and Weibull Families

On the Figure 4.1, we can clearly see the different shape of the families distribution regarding the ξvalue.

This formulation provides a flexible model that can represent different types of tail behaviors in a unified manner. The GEV distribution provides a unified framework for modelling extreme values, simplifying the modelling and estimation processes in various fields such as hydrology, finance, and engineering. The parameters μ, σ, and ξ allow the distribution to be tailored to the empirical data, making the approach both flexible and robust for different practical applications.

4.3 Block Maxima Model (BMM)

The first model we will present is the BMM, which is useful for estimating extreme values from a set of independent and identically distributed (i.i.d.) observations. To run this model, the first step is to choose equally sized blocks by splitting the data and then fitting the (GEV) distribution to the maxima of these blocks. The number of blocks chosen depends on the size and nature of the data. For financial data, it is common to align the block size with the period of interest (e.g., monthly, quarterly, yearly). For other types of data, different block sizes may be more appropriate (Coles, 2001).

Assuming that our observations are independent, a series of independent stochastic variables $X_1, X_2, \ldots$ are monitored. With the block maxima model, these variables are grouped into blocks of a chosen size n, from which the maxima will be extracted as $M_{n,1}, \ldots, M_{n,m}$ and then fitted to the GEV distribution (de Haan & Ferreira, 2006). Selecting the appropriate block size is a significant challenge in this model. As the number of blocks increases, the number of data points available for analysis decreases. Minimizing the number of blocks can complicate the estimation of the maxima distribution, making the process less practical.(Embrechts et al., 1997). In the analysis we shall see the impact of those block size variations.

4.4 Maximum likelihood

This part goes through the estimation of the maximum likelihood. ξ will play a role in the estimation of the maximum likelihood.

- For $\xi > -0.5$: Regular estimation and asymptotic properties hold. This range ensures that the GEV distribution remains within the realm where maximum likelihood estimation (MLE) yields reliable and consistent results.
- For $-1 < \xi < -0.5$: No asymptotic properties. In this range, while estimation can still be performed, the lack of asymptotic properties means that the estimates might not converge to the true parameter values as the sample size increases.
- For $\xi < -1$: No estimation is possible. This bound ensures that the GEV distribution remains valid. When ξ is less than -1, the distribution can produce non-physical results (e.g., negative probabilities), making the estimation process infeasible.

Let's assume that $Z_1, \ldots, Z_m$ are independent and follow the GEV distribution. In the case where $\xi \neq 0$, the log-likelihood can be defined as:

$$\ell(\mu, \sigma, \xi) = -m \log \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{m} \log \left[1 + \xi \left(\frac{Z_i - \mu}{\sigma}\right)\right] - \sum_{i=1}^{m} \left[1 + \xi \left(\frac{Z_i - \mu}{\sigma}\right)\right]^{-1/\xi} \tag{4.9}$$

subject to the constraint:

$$1 + \xi \left(\frac{Z_i - \mu}{\sigma}\right) > 0 \quad i = 1, ..., m. \tag{4.10}$$

These constraints ensure that the GEV distribution parameters are correctly estimated within a feasible range. The constraints on ξ are critical to maintaining the theoretical properties of the GEV distribution, including its asymptotic behavior and the validity of the maximum likelihood estimates.

If the parameters violate the condition, it corresponds to at least one data point falling beyond the endpoint of the distribution. In such cases, the likelihood would be equal to 0, and the log-likelihood would tend to negative infinity.

When $\xi = 0$, the log-likelihood equation is:

$$\ell(\mu, \sigma, 0) = -m \log \sigma - \sum_{i=1}^{m} \frac{Z_i - \mu}{\sigma} - \sum_{i=1}^{m} \exp \left(-\frac{Z_i - \mu}{\sigma}\right) \tag{4.11}$$

This equation is derived using the Gumbel limit for the GEV distribution. By optimizing this equation while respecting the three parameters of the GEV distribution, we obtain the maximum likelihood. Care must be taken not to violate the parameter conditions, as this could lead to invalid results. When ξ is near zero, it is preferable to use the second equation instead of the first one.

For the case where $\xi \neq 0$, the return level z^p is given by:

$$z^p = \mu + \frac{\sigma}{\xi} \left((-\log(1-p))^{-\xi} - 1 \right) \tag{4.12}$$

Where:

- $\hat{\mu}, \hat{\sigma}, \hat{\xi}$ are the maximum likelihood estimates of the GEV parameters.
- $y_p = -\log(1-p)$.

We are going to look at the use of these parameters obtained with ML estimations, as they'll be useful for estimating VaR under the BMM.

4.5 VaR with BMM

Thanks to the parameters obtained in the previous section: (μ), (σ) and (ξ). The maximum of the log-likelihood of the GEV density gives us the following formula for VaR:

$$\text{VaR} = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}} \left[1 - (-\log(q))^{-\hat{\xi}} \right], \quad \hat{\xi} \neq 0 \tag{4.13}$$

For $\hat{\xi} = 0$:

$$\text{VaR} = \hat{\mu} - \hat{\sigma} \log \left[\log \left(\frac{1}{q} \right) \right], \quad \hat{\xi} = 0 \tag{4.14}$$

where:

- $\hat{\mu}$ is the location parameter,
- $\hat{\sigma}$ is the scale parameter,
- $\hat{\xi}$ is the shape parameter,
- q is the confidence level.

4.6 Peak over Treshold model (POT)

Let's move on to the second model presented in this master's thesis. We will go through the theory by highlighting the key points and equations. The section below is based on the work of (Coles, 2001), known for its thorough and precise analysis.

The Peak Over Threshold (POT) model operates differently from the block maxima approach. POT defines a threshold u and considers all values exceeding this threshold as extreme. The distribution representing the excess above a certain threshold is defined as follows:

$$F_u(y) = P(X - u \leq y \mid X > u) \tag{4.15}$$

for, $0 \leq y < x_0 - u$ où $x_0 \leq \infty$, represented the right endpoint of F. This distribution represents the probability of a loss exceeding the threshold u by an amount y. (McNeil, 1999)

The probability function derived from this model is as follows, and will form the basis for the rest of this section:

$$\Pr\{X > u + y \mid X > u\} = \frac{1 - F(u + y)}{1 - F(u)}, \quad y > 0. \tag{4.16}$$

According to this conditional probability, composed of independent and identically distributed variables $X_1, X_2, \ldots$, the extreme events of the different variables are those that exceed the threshold u.

In an ideal scenario, if the distribution F were known, then the distribution of thresholds would also be known. However, like the previous model, this distribution is typically unknown. Therefore, we use the GPD to approximate the distribution when the distribution F is not known.

To approximate the distribution of the excess term, we use the GPD, which is defined as follows:

$$G(y) = 1 - \left(1 + \frac{\xi y}{\sigma}\right)^{-1/\xi}, \quad y > 0, \sigma > 0, \xi \in R. \tag{4.17}$$

The equation above defines a new family called the Generalized Pareto family. In comparison, if BMM help to approximate the G distribution, then in this new family POT also has an estimate of this same G distribution. GPD parameters are determined solely by GEV parameters associated with BMM.

4.6.1 Selection of the best threshold

The Mean Residual Life (MRL) plot, introduced by (Davison & Smith, 1990), is a tool for determining the threshold u using the expectation of the GPD excesses. The expectation of the excesses over the threshold u is given by:

$$e(u) = E(x - u \mid x > u) = \sigma_u + \frac{\xi_u}{1 - \xi}, \tag{4.18}$$

where ξ is the shape parameter, and σ_u is the scale parameter corresponding to the threshold u. The condition $\xi < 1$ ensures the existence of the expectation. This equation shows that the expectation of excesses is linear in u, with a gradient of $\frac{\xi}{1-\xi}$ and an intercept of $\frac{\sigma_u}{1-\xi}$.

For a set of samples $(X_1, X_2, \ldots, X_n)$, the empirical estimate of the mean of excesses is calculated as:

$$e_n(u) = \frac{1}{N_u} \sum_{i=1}^{N_u} (X_i - u), \quad X_i > u, \tag{4.19}$$

where N_u denotes the number of exceedances over u. The data set $\{u, e_n(u)\}$ forms the MRL plot. Generally, the value of u above which the plot becomes approximately a straight line is selected as the optimal threshold.

4.7 Maximum likelihood

The goal of this section is the same as for the first model: parameter estimation using the maximum likelihood technique. This involves estimating the parameters and replacing the actual parameters with their estimates. For the parameters σ, and ξ, this means applying the MLE. Additionally, an estimation of the number of values exceeding the threshold u is needed through the parameter C.

The number of values exceeding the threshold u follows a binomial distribution $\mathrm{Bin}(n, C)$, where $\hat{C}$ is the maximum likelihood estimation of the parameter C. A standard property of the binomial distribution states that:

$$\hat{C} = \frac{N_u}{n}, \tag{4.20}$$

where N_u is the number of exceedances and n is the total number of observations. Therefore, the variance-covariance for all the estimated parameters can be computed.

The application of the maximum likelihood method in our analysis allows us to obtain accurate and robust estimates of the parameters of the extreme value model, calculate reliable return levels, and manage dependencies in the data. These aspects are crucial for effective modeling of extreme events and for making informed risk management decisions.

4.7.1 Estimation

Given a defined threshold, the parameters of the GPD can be estimated using the MLE method. Suppose that the values $y_1, \ldots, y_n$ are the y_n values that exceed the threshold

u. For $\xi \neq 0$, the log-likelihood equation is

$$\ell(\sigma, \xi) = -n \log \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{n} \log\left(1 + \frac{\xi y_i}{\sigma}\right),$$

provided that $\left(1 + \frac{\xi y_i}{\sigma}\right) > 0$ for $i = 1, \ldots, n$; otherwise, $\ell(\sigma, \xi) = -\infty$.

Numerical techniques are used to maximize the log-likelihood while avoiding instability when ξ is close to 0 and staying within the parameter-defined intervals. We will use this technique to calculate the parameters of our GPD during our analysis, making this technique highly useful for our purposes.

4.8 VaR with POT

Knowing how to estimate the best parameters for our distribution, we can now, with the parameters, estimate the VaR and the equation with the POT model derives a probability ($\alpha > F(u)$), (McNeil, 1999)

$$\mathrm{VaR}_\alpha = u - \frac{\hat{\sigma}}{\hat{\xi}}\left(\left(\frac{n}{N_u}(1-q)\right)^{-\hat{\xi}} - 1\right) \tag{4.21}$$

where:

- VaR_α is the VaR at confidence level α.
- σ is the scale parameter of the GDP
- ξ is the shape parameter of the GDP with
- N_u is the number of observations over the threshold u
- n is the number of observations

4.9 Conclusion

This section on EVT is crucial for understanding and analyzing rare events across various fields, especially in finance and environmental studies. It lays the theoretical foundation by explaining BMM and POT models, essential for applying EVT. Understanding the Gumbel, Fréchet, and Weibull distributions, and their unification under the GEV distribution, is fundamental for modeling extreme events. The methodology of dividing data into blocks and applying the GEV distribution is critical for estimating extreme values. Additionally, the POT model provides an alternative approach useful when data is plentiful but extreme events are rare. The use of MLE for determining GEV parameters offers a rigorous method for estimating extreme risks.

This knowledge is vital for practical applications in extreme risk analysis in finance but also climate science, and other fields where extreme events have significant impacts. These models are crucial for risk management and decision-making and the flexibility of the GEV distribution allows adaptation to various situations and data types without rigid preliminary choices. The POT methodology enables a more refined analysis of extreme events, focusing on threshold exceedances, often more relevant in real-world applications. This section also paves the way for advanced analyses, such as MLE, essential for refining risk estimates parameters and providing accurate predictions.

5. Empirical Analysis

The objective of this second major part of this master's thesis is to apply the theory seen above to real values, to extract the results and to produce a quantitative analysis, which will subsequently enable us to draw conclusions about the various models applied. We will undertake an empirical analysis based on EVT, presented in Chapter 4. The goal is to understand the tail distribution within the data, which helps to determine the prediction levels of the different models. Firstly, we'll consider two historical methods: the VaR normally distributed and the exponential Weighted Moving Average (EWMA) method. Secondly, the BMM, a division of the data into equal temporal blocks to extract the maximum values to estimate the parameters of the GEV. From that distribution, we will perform a VaR estimation. Then the POT model is designed to identify extreme values by setting a specific threshold. Any values exceeding this threshold are categorised as "extreme". These "extreme" values will then be used to represent the GPD and estimate the VaR.

Based on these analyses, we will draw conclusions on the effectiveness and accuracy of different models in predicting the risk performance of financial indices. These conclusions are crucial in financial risk management, providing players in this sector with the necessary tools to better understand, analyze risks more effectively and attempt to predict them as accurately as possible.

5.1 Methodology

The **data** used to compile the various models in this analysis consist of observations from the largest national and global financial indices. In fact, the choice of this data was based on the geography of these indices. Our aim was to compare these well-known and high-performing indices with what we define as 'local' indices on our scale. These indices are portfolios of shares, which take, according to which country they are attached, a certain number of shares of national companies with the largest capitalisation.

This analysis and comparison of models will therefore focus on the prices of the Belgian BEL20 index (20 largest Belgian capitalisations), the CAC40 (40 largest French capitalisations), the S&P500 (500 largest American capitalisations) and finally the MSCI World index, the largest capitalisations worldwide. All these indices are based on daily observation of their price movements. These indices have been extracted from the Bloomberg platform.

These raw daily data have therefore been processed to extract the information we need for our future analyses. Logarithm returns, that have been used here, are useful for comparing indices on the same basis. In fact, indices can be larger or smaller than others and therefore prices can also follow this logic. For this reason, we have transformed the data into log-returns to understand equivalent things. Another reason for this, is that the logarithm returns can be summed over several periods to give the average return over the entire period studied. They will also enable us to give more stable and interpretable metrics. Let's assume that p_t is the price at time t and p_{t-1} is the price at time $t-1$. Then the log return would be:

$$r_t = \log\left(\frac{p_t}{p_{t-1}}\right) \tag{5.1}$$

To carry out our analysis of the performance of our prediction models, we had to divide our data into **samples**: in-sample and out-of-sample. Additionally, we had to play with sliding windows over time. The 'in-sample' data is used to set the model parameters to make the most accurate predictions possible. Out-of-sample data is used to test these parameters on new data and evaluate the model's performance. The principle of sliding windows is simple. We estimate parameters using the 'in-sample' data and test them on the next 'out-of-sample' period, which is the following month or quarter, then we repeat the process with a time shift. The use of out-of-sample data allows us to cover a very wide range of data while making predictions based on actualised data. For our analysis, we decided to work with estimation periods of 2 years for monthly out-of-sample periods and periods of 4 years for quarterly out-of-sample periods, to reduce the difference in the number of data items in the in-sample. The table below shows the number of periods and the total daily data estimated for the monthly and quarterly.

Out-of-sample	Start Date	End Date	Periods	Observations
Monthly	2001/02/04	2024/05/07	275	6050
Quarterly	2003/01/21	2024/04/22	86	5525

Table 5.1: Out-of-sample Monthly and Quarterly Data: Start and End Dates, Number of Periods, and Observations

Let's move on to the **analysis method** we'll be applying to test the various models covered in this thesis. First, we'll analyse the raw data to extract the information we're interested in, to learn more about the behavior of our data. These primary analyses are based on the first methods of statistical analysis. This will cover mean of returns, variance, skewness, kurtosis, Jarque-Bera and Dickey Fuller tests. As we want to work in an extreme environment, we need to ensure that our data behaves in an extreme way over time, and these different metrics and tests will help us to define this.

Secondly, we're going to compare extreme models with 'classic' models, so we'll see if the former perform better. The two classic models based on historical data assume a normal distribution of returns, with a difference for the EWMA, which adds a coefficient of weight to the most recent returns, and this coefficient decreases as the return gets older. This makes it possible to give more importance to the most recent returns and adapt more quickly(MathWorks, n.d.). These two methods will be applied to out-of-sample periods in Monthly and Quarterly periode on the four indices.

Finally, we'll apply the BMM and POT models to our four indices to see how they perform and compare them with historical models.

5.2 Results

Now that we have the information explained on the data and the methodology, let's analyse our indices. The first Figure 5.1, shows the price of all the indices with their daily log returns below.

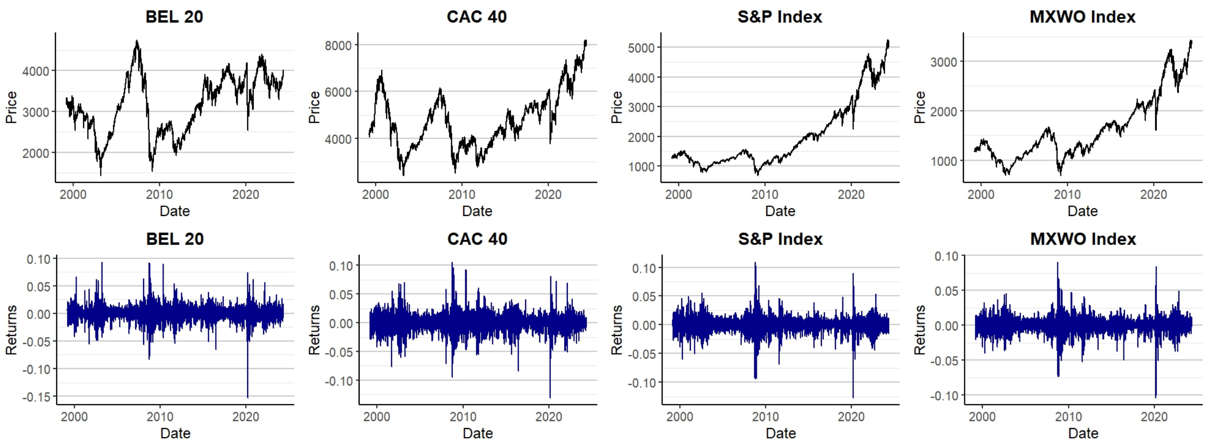
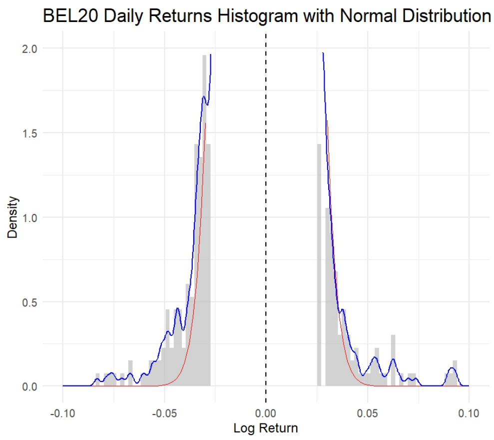


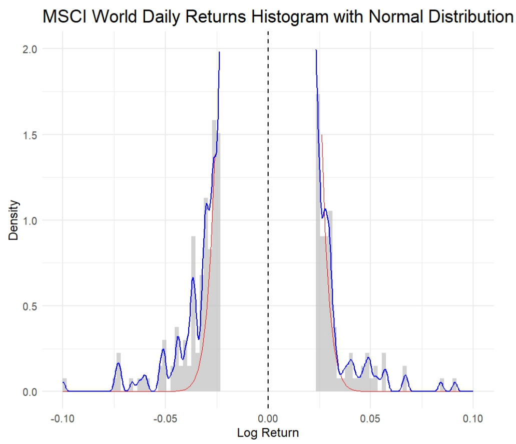
Figure 5.1: Prices and log daily returns for BEL 20, CAC40, S&P500, and MSCI World Indices

Over the **long term** for the CAC40, S&P500 and the MSCI World, we can see a strong upward trend for all these price charts despite the various crises that have occurred. By contrast, the BEL20 is performing less well, with a major correction around 2008 because of the crisis, followed by a weaker uptrend. In terms of **volatility**, it increased during periods of global economic crisis. On the one hand, the two smallest indices, BEL20 and CAC40, are highly sensitive due to their lesser diversification and on the other hand, we can see that the MSCI World is less volatile due to its high degree of diversification. As we can see from the return graphs, after a period of small fluctuations often comes a period of large fluctuations. In view of the **recurring crises**, it is vital to have the tools to analyse these periods of high volatility and recession.

To understand the phenomenon we are interested in, let's look at the histograms of the distributions of the daily returns of the BEL20 and MSCI World index compared with the



(a)



(b)

Figure 5.2: Histograms of Daily Returns for (a) BEL20 and (b) MSCI World Indices Compared with Normal Distribution

normal distribution in figure 5.2 (Same for CAC40 and S&P500 in figure A.1). We can see that the left tails in all indices are thicker and go further on the left side. We can also see that the distributions of returns deviate more quickly from the normal distribution on the left of the graphs than on the right, indicating strongly negative return values during certain periods. Due to the thicker shape of the BEL20 and CAC40 left return distribution, we can conclude to a stronger volatility for those two indices.

From the Table 5.2 below, we extracted the most important metrics in terms of pre-analysis and discuss about the results.

Table 5.2: Statistical Metrics for BEL 20, CAC 40, S&P 500, and MSCI World Indices

Metrics	**BEL 20**	**CAC 40**	**S&P 500**	**MSCI World**
Mean	0.00744	0.02601	0.05374	0.040967
Standard deviation	0.19307	0.21819	0.1914	0.16034
Skewness	-0.3773	-0.2112	-0.3758	-0.5736
Kurtosis	9.7520	6.6756	10.6201	10.668
Minimum	-0.1532	-0.1309	-0.1276	-0.1044
Maximum	0.0933	0.1059	0.1095	0.0909
Jarque-Bera test	26203***	12255***	31045***	31535***
ADF values	-19.749**	-19.445**	-19.016**	-18.784**
Observations	6568	6568	6568	6568

- **Mean**: When we compare mean returns, we see that the BEL20 and the CAC40 have lower mean returns than the S&P500 MSCI World, which means that the expected performance of the European indices over this period are lower.

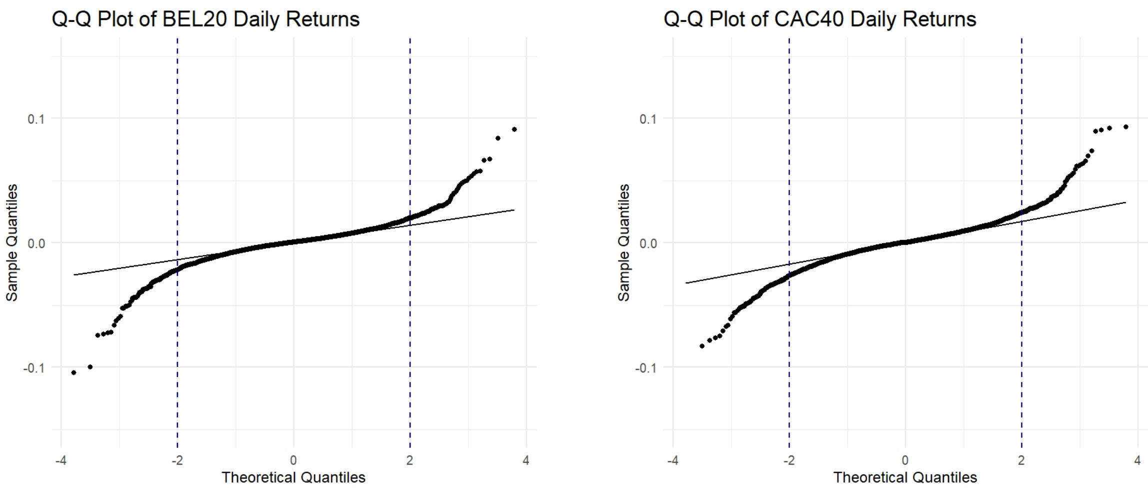
- **Skewness**: The four indices have negative skewness values, which refers to an asymmetric distribution of returns, more pronounced in the negative.

- **Kurtosis**: This value is of particular interest to us because it confirms, when it is above 3, that the data contains a phenomenon called "fat-tail", which reflects extreme events in the negative tail.

- **Jarque-Bera Test**: This test is useful for confirming whether the distributions are normally distributed, in our case with the JB Test all significant (***). P-value is under 1%, which means that the null hypothesis can be rejected because there is less than a 1% chance that the data is normally distributed. This confirms an asymmetric distribution. This test is particularly effective when the samples are large, which is our case.

- **ADF Values**: Using this test to identify unit roots, we can reject the null hypothesis and conclude that the return series are non-stationary.

With the Q-Q plots in figure 5.3 and A.2 , useful for comparing the distribution of a series of data with the normal distribution, we can see that for all the indices, analysis of the JB Test and skewness confirm that there is a "fat-tail" phenomenon in the distributions. In fact, we can see that the distribution of our data deviates more from the normal distribution in the distribution of negative returns.



(a) (b)

Figure 5.3: Quantile-Quantile Plots for BEL 20 and CAC40 Indices

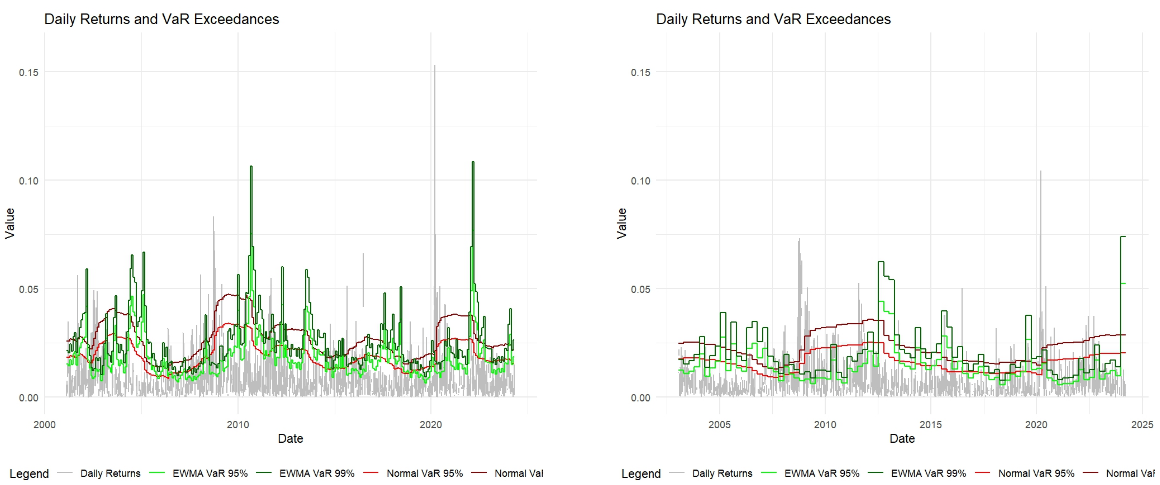
Thanks to the analyses in this section, we can conclude that we're going to face data with extreme behavior in some cases. We'll be able to apply our analysis in the next section and see how effective these models are. Now let's apply our historical and extreme models and compare them.

5.2.1 Historical Methods

For our first models, we are tackling the most basic ones, which will give us a solid foundation for comparison. These models are based on the normal distribution and on the past, while applying an evolutionary coefficient to each return in terms of the weight it carries in estimating the future, so that we cannot assume any extreme behavior or make any assumptions other than these.
The historical method is based on the normal distribution and assumes that returns follow this distribution. It uses the standard deviation and the mean of the in-sample returns to estimate the VaR for each out-of-sample, leading to a full estimation. This method provides a simple approach to estimation but as mentioned earlier, assumes that returns follow a normal distribution and we have seen that this is not the case.

As for the EWMA method, unlike the previous model which assumed that each return weighed the same in the estimate, this time the weight of each return decreases exponentially with the passage of time. This calculation method allows the model to react more quickly than the first.

For the analysis of these models, we have estimated these VaR over the full period of the out-of-sample monthly as well as quarterly, in order to be able to compare similar periods with the other models. The Figure A.5, shows, on the left, the daily returns of the BEL20 compared to the VaR estimation curves of the two historical methods and on the right for quarterly periods for the MSCI World index. (For other charts, see appendix)



(a) (b)

Figure 5.4: VaR Estimates Using Normal Distribution Method and EWMA Method for (a) monthly periods for BEL20 and (b) quarterly periods for MSCI World

In our case, we speak of VaR exceedance when the actual return over a period, in this case a day, exceeds the VaR estimate for that period. It is thanks to this number of exceedances that we are going to estimate the efficiency of our estimation models.

This method is called backtesting and is used to determine the efficiency of a model. In our case, estimating the efficiency of VaR is done as follows: we calculate the number of overshoots over the period and divide it by the number of times in that period. This will give us the overshoot percentage which can be compared, in the case of VaR to the confidence level. In fact, for a VaR with a confidence level of 95%, the theoretical percentage of overruns expected is equivalent to 5% (1% for 99% VaR). In the case of a 95% confidence level, several exceedances below this 5% will indicate an overestimate of the loss, while a higher figure will indicate an underestimate of the loss. The aim of our models is to get as close as possible to this theoretical percentage.

The tables 5.3 and 5.4 below show the excess rate for the two historical VaR models over monthly and quarterly estimation periods.

Monthly	**Normal Distribution VaR**		**EWMA VaR**	
	Exceedances (95%)	**Exceedence (99%)**	**Exceedances (95%)**	**Exceedances (99%)**
BEL20	322/6050 5.32%	154/6050 2.55%	505/6050 8.35%	293/6050 4.84%
CAC40	324/6050 5.36%	143/6050 2.36%	485/6050 8.02%	250/6050 4.13%
S&P500	320/6050 5.29%	153/6050 2.53%	470/6050 7.77%	276/6050 4.56%
MSCI World	318/6050 5.26%	168/6050 2.78%	487/6050 8.05%	279/6050 4.61%

Table 5.3: Monthly VaR Exceedances Using Normal Distribution and EWMA Methods for BEL20, CAC40, S&P500, and MSCI World Indices at 95% and 99% Confidence Levels

Quarterly	**Normal Distribution VaR**		**EWMA VaR**	
	Exceedances (95%)	**Exceedances (99%)**	**Exceedances (95%)**	**Exceedances (99%)**
BEL20	255/5525 4.62%	121/5525 2.19%	466/5525 8.43%	266/5525 4.81%
CAC40	247/5525 4.47%	109/5525 1.97%	415/5525 7.51%	235/5525 4.25%
S&P500	267/5525 4.83%	139/5525 2.52%	459/5525 8.31%	268/5525 4.85%
MSCI World	259/5525 4.69%	140/5525 2.53%	473/5525 8.56%	262/5525 4.74%

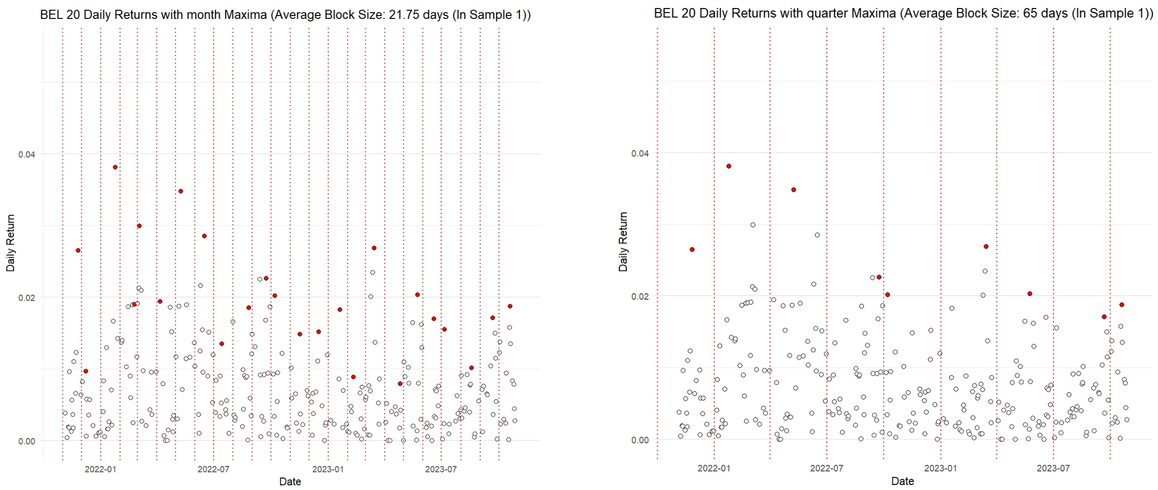
Table 5.4: Quarterly VaR Exceedances Using Normal Distribution and EWMA Methods for BEL20, CAC40, S&P500, and MSCI World Indices at 95% and 99% Confidence Levels

From the two tables, we see that the normally distributed VaR outperforms the EWMA VaR for all 4 indices. Indeed, the EWMA VaR has all his exceedances almost twice as large as the normally distributed VaR, and this for the 95% and 99% VaR. We can say that the EWMA model has a strong tendency to underestimate the risk and is not reliable when estimating the potential risk incurred. The model based on the normal distribution performs better on the quarterly estimates, being slightly closer to the theoretical percentage than for the monthly estimates. We can see that: on all figures seen before, that the EWMA retains extreme events to a greater extent over the estimated period because it places more importance on the most recent returns. That leads to a big number of exceedances.

To take this a step further into our extreme analysis, we now turn to the BMM, it focuses on temporal block maxima, enabling us to better model extreme risks in financial markets.

5.2.2 Block Maxima Model (BMM)

To understand the basic principle of the BMM, let's take a look at Figure 5.6 an "In-Sample" of the daily returns of the BEL20 index (extract from one in-sample period as example for monthly and quarterly). The sample on each chart is divided into equal parts: monthly on the left and quarterly on the right. In each of these "blocks", the maximum value is selected, creating a new sequence of maximum values. In our case, we analyze two years of "In-Sample" data for the monthly periods, divided into months and quarters, to see how this impacts the results obtained. These maxima are highlighted in red on the graphs.



(a) (b)

Figure 5.5: Block Maxima Model - In-Sample Analysis (a) monthly and (b) quarterly for BEL20 Index

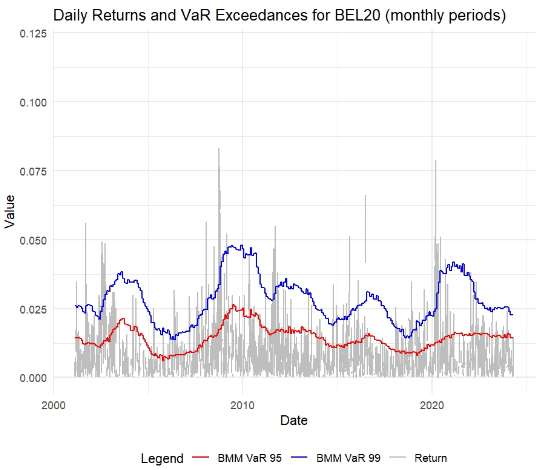
From this sequence of maxima, the parameters required to estimate the GEV distribution are calculated using the MLE method described in section 4.4. After estimating these parameters for all periods, we estimate the VaR based on these parameters and using the specific VaR function for the BMM, see section 4.5. We calculate the number of excesses on the corresponding out-of-sample periods to know the reliability of the model. All these results are shown in the table below for all monthly and quarterly indices.

Out-of-Sample	**Monthly - VaR Exceedances**		**Quarterly - VaR Exceedances**	
	95%	**99%**	**95%**	**99%**
BEL20	550/6050 (9.09%)	140/6050 (2.31%)	985/5525 (17.83%)	159/5525 (2.88%)
CAC40	486/6050 (8.03%)	122/6050 (2.02%)	808/5525 (14.62%)	151/5525 (2.73%)
S&P500	484/6050 (8%)	132/6050 (2.18%)	858/5525 (15.52%)	175/5525 (3.17%)
MSCI World	521/6050 (8.61%)	149/6050 (2.46%)	885/5525 (16.02%)	191/5525 (3.46%)

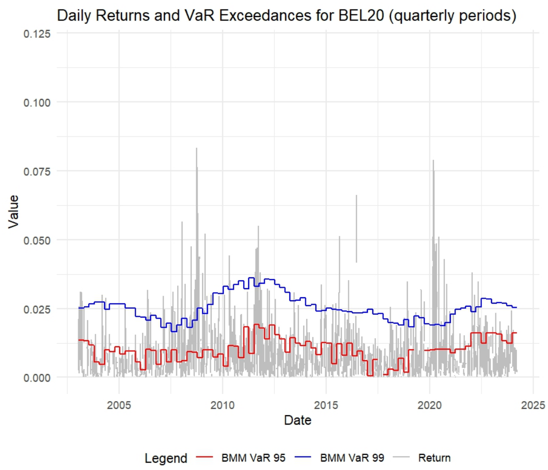
Table 5.5: Summary of VaR Exceedances Using Block Maxima Model for BEL20, CAC40, S&P500, and MSCI World Indices at 95% and 99% Confidence Levels

In this table 5.5 of VaR estimates for all the indices, we see that the BMM for estimation over monthly periods performs better than the two previous methods in the case of VaR at 99%. However, for estimation over a three-month period we see a direct drop in efficiency compared with monthly periods. In this exercise, estimation using the normal distribution works better than the block maxima. As for the VaR at a confidence level of 95%, we see that the results are to be excluded, in fact having to deal with a model based on extreme events estimating VaR at such a large confidence level is not optimal and therefore a comparison with these results is not possible.

To visualise the excesses, the 95% and 99% VaR curves have been plotted in the figures for all the indices (See appendix). These plots in figure 5.6 confirm that the BMM is effective for predicting over one-month-periods. We can see that the 99% VaR curve adapts well to market movements.



(a)



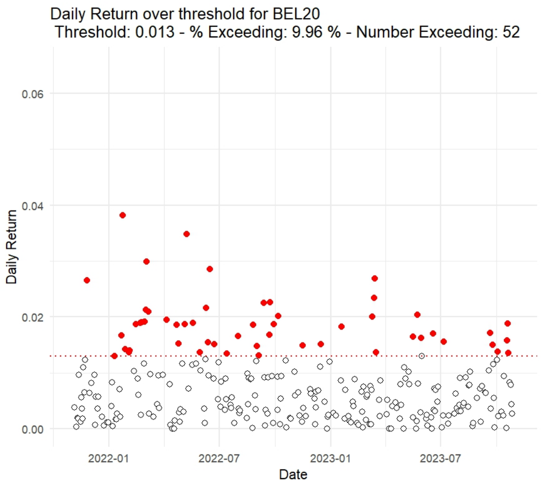
(b)

Figure 5.6: Block Maxima Model VaR Estimations for BEL20 at 95% and 99% Confidence Levels, (a) Monthly and (b) Quarterly

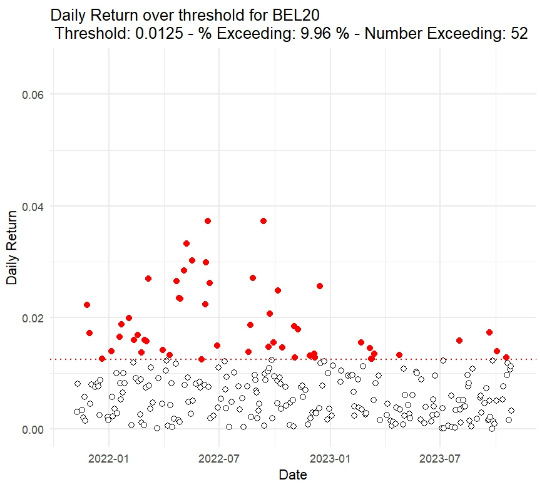
In conclusion, we are dealing with a model that performs better than the historical models over monthly periods and with small confidence levels. However, the model is worse when we try to estimate over longer periods and with larger confidence levels. As the BMM takes the maximum per block, it omits the extreme potentials in second and third place in the blocks, which limits its estimates and therefore leads to potential errors. We will now look at our second EVT model, the POT model who may not deal with this issue.

5.2.3 Peak Over Threshold Model (POT)

Based on the logic of the BMM model, in the POT model, extreme values are selected in relation to a predefined threshold which defines values as extreme if they exceed it. As illustrated in the figure 5.7 showing the daily performance of the BEL20 index and the MSCI World index, values exceeding this threshold are highlighted in red and, for this example, the threshold is defined if extreme values represent 10% of the total sample. For (Coles, 2001) , a threshold that captures 10% of the data in the sample is a good threshold and we will base ourselves on this assumption. This 10% represents the proportion of data considered as extreme values. Based on these extreme values, like the BMM, the POT model uses these values to estimate GDP and the parameters of this distribution are obtained using the MLE method described in the section 4.7.



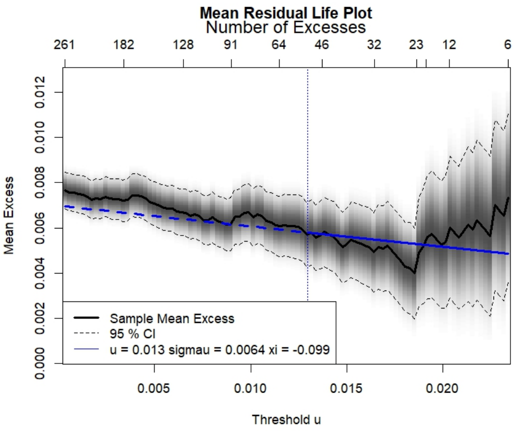
(a)



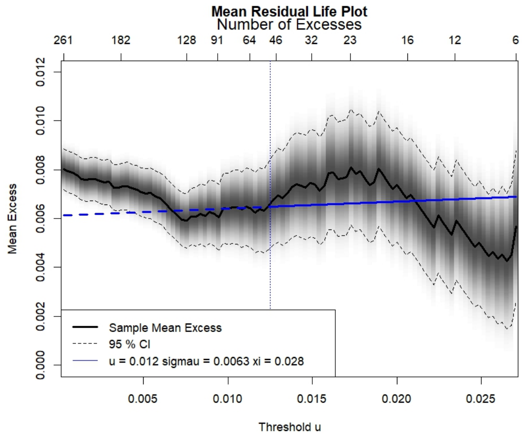
(b)

Figure 5.7: Maxima Over Threshold - In Sample monthly periods for (a) BEL20 and (b) MSCI World Indices

To confirm this assumption of the 10% proportion, the average residuals plot is a good graphical tool for defining the best possible threshold to best model the extreme values above this threshold. This graph plots the average value of the surpluses above the various possible thresholds against these thresholds, i.e. it shows the averages of these observations as a function of this threshold. When a linear region appears, an ideal threshold is often present in this zone because it suggests that beyond this threshold the values will follow an exponential distribution and this is desirable for modelling a GPD. In the example below we have plotted mean residuals life plot to visualise the phenomenon. We see that the desirable thresholds are defined around the 50-excess data for the two data sets, given that we are working with samples of 520. This confirms that the ideal threshold is around the 10%. (See figure 5.9)



(a)



(b)

Figure 5.8: Mean Residual Life Plot for (a) BEL20 and (b) MSCI World Indices

Now that we have an optimal solution for finding our thresholds, we will estimate the GPD for each in-sample to estimate the VaR thanks to this distribution on the corresponding out-of-sample. To have optimal results we will create a confidence interval [-0.1%, +0.1%] around the threshold and see how the estimates react. In the tables 5.6 and 5.7 below, we have VaR estimates for the monthly and quarterly estimation periods for VaR at 95% and 99% and confidence level for the threshold.

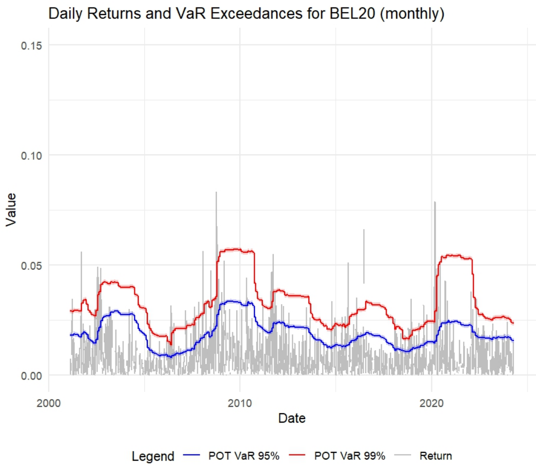
Out-of-Sample	**VaR 95%**			**VaR 99%**		
Monthly	**Lower Thresold**	**Thresold 10%**	**Upper Thresold**	**Lower Thresold**	**Thresold 10%**	**Upper Thresold**
BEL20	376/6050 (6.21%)	329/6050 (5.43%)	285/6050 (4.71%)	107/6050 (1.77%)	97/6050 (1.60%)	89/6050 (1.47%)
CAC40	358/6050 (5.91%)	319/6050 (5.27%)	286/6050 (4.72%)	80/6050 (1.32%)	88/6050 (1.46%)	100/6050 (1.65%)
S&P500	347/6050 (5.74%)	296/6050 (4.89%)	271/6050 (4.48%)	110/6050 (1.82%)	104/6050 (1.72%)	93/6050 (1.54%)
MSCI World	363/6050 (6.00%)	303/6050 (5.01%)	264/6050 (4.36%)	116/6050 (1.92%)	99/6050 (1.64%)	89/6050 (1.47%)

Table 5.6: Monthly VaR Exceedances Using Peak Over Threshold model for BEL20, CAC40, S&P500, and MSCI World Indices at 95% and 99% Confidence Levels with Threshold capturing 10% of the data and a confidence interval of [-0.1%, +0.1%]

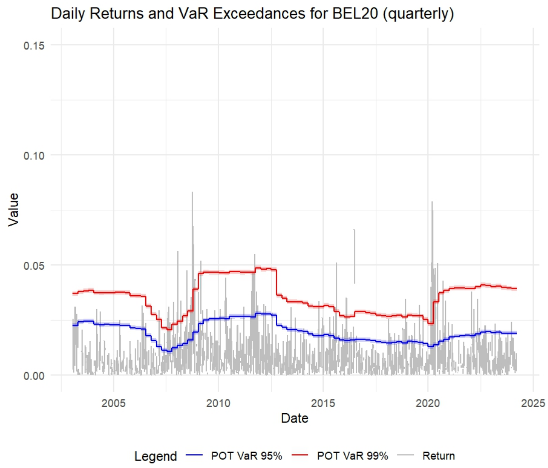
Out-of-Sample	**VaR 95%**			**VaR 99%**		
Monthly	**Lower Thresold**	**Thresold 10%**	**Upper Thresold**	**Lower Thresold**	**Thresold 10%**	**Upper Thresold**
BEL20	246/5525 (4.45%)	273/5525 (4.94%)	311/5525 (5.63%)	66/5525 (1.19%)	69/5525 (1.25%)	73/5525 (1.32%)
CAC40	286/5525 (5.17%)	257/5525 (4.65%)	236/5525 (4.27%)	59/5525 (1.07%)	65/5525 (1.18%)	72/5525 (1.30%)
S&P500	306/5525 (5.54%)	280/5525 (5.07%)	259/5525 (4.69%)	94/5525 (1.70%)	82/5525 (1.48%)	76/5525 (1.38%)
MSCI World	312/5525 (5.65%)	277/5525 (5.01%)	249/5525 (4.51%)	87/5525 (1.57%)	73/5525 (1.32%)	65/5525 (1.18%)

Table 5.7: Quarterly VaR Exceedances Using Peak Over Threshold model for BEL20, CAC40, S&P500, and MSCI World Indices at 95% and 99% Confidence Levels with Threshold capturing 10% of the data and a confidence interval of [-0.1%, +0.1%]

With the POT models, we have a clear improvement in the results obtained with each time the percentage of excess around the desired percentage. The POT does the best job of estimating VaR over quarterly periods at 95% confidence with the 10% threshold and at 99% confidence with the upper threshold, even though this is still very close to the reference threshold. We find values very close to 5% for the 95% VaR with the reference threshold, with an estimate that gives 5.01% excess and the furthest value is 4.65%, which is still a very good result. As for the 99% VAR, based on the results the upper threshold can be defined as better but with such a small spread it is difficult to conclude. As with the previous models, here is an illustration of the 95% and 99% VAR estimation curves in relation to the daily returns for the out-of-sample period for BEL20.



(a)



(b)

Figure 5.9: Daily return for BEL20 index with VaR 95% and 99% curve estimation on (a) monthly and (b) quarterly periods

We can therefore conclude that the POT is the best model in terms of future prediction for VaR values. It has outperformed all other models by far with consistency across different confidence levels and estimation periods.

5.2.4 Conclusion

In this part of our empirical analysis, we applied Historical VaR methods and EVT methods to real data from major financial indices. The aim was to understand the distribution of the tails of these data to accurately predict risk levels. We used the two best-known EVT models: historical VaR methods normally distributed and EWMA, as well as the BMM and the POT.

This analysis was crucial in comparing the effectiveness of historical and extreme models in predicting the performance of financial indices. Using real data and in-sample and out-of-sample test periods, we were able to assess the robustness of the different models. After all our analysis, we can conclude that EVT-based models do better than simple VaR estimation models. Although the BMM has difficulty estimating low confidence levels, it performs well at higher confidence levels. Of these four models, the POT is clearly the most effective, with results for each index very close to the theoretical ones, over the two confidence levels and the two time horizons. The results obtained are very close to 5% for the VaR (5.01%) and very close to 1% for the 99% VaR (1.18%).

6. Conclusion

Through this master's thesis, we were able to explore a part of risk management, focusing on extreme events using different models. We have described in detail the traditional VaR measure and, alongside this, models based on EVT. This has enabled us to understand how they work in predicting financial losses.

The theory part lays the groundwork by defining risk in financial context, highlighting their natures and the real challenges in estimating them.

The historical risk methods, normally distributed and EWMA VaR, which are much simpler to implement, show us some difficulties in predicting extreme events, in view of the results, which are a long way from the expected theoretical results wich is 1% for VaR at 99% (EWMA method has a percentage error of around 4.5% for monthly estimates at 99% VaR, while POT performs at around 1.30%). In the meantime, the VaR method estimated based on the normal distribution gives good results on the 95% confidence level. Its results are close to 5% performing better in quarterly periods. The block maxima model has shown its limitations, with values completely out of range at around 17% over the 95% confidence level and around 2% on the 99% confidence level. This can be explained by the way it selects extreme values, only one per block, which means that it omits a large number of them. One thing has become clear: so-called classic historical method based on the normal distribution is reliable at confidence levels as high as 95%, but extreme models are indispensable when confidence levels rise drastically. Despite the fact that this is an extreme model, it is impressive to see that the POT model performs just as well at 95% confidence levels as at 99%.

The results above confirmed our initial hypotheses on the impact of the periods studied, which influenced our results in our case. At the same time, we have succeeded in presenting a method that is both effective and reliable for predicting maximum losses over a given period. Applying this to the Belgian market gives us an idea of the kind of data that anyone could be confronted with in their financial life.

Through this master's thesis, we have been confronted to limitations, as the data selection, although the best-known financial indices, do not capture all future market dynamics and can be paired with other data to make more realistic predictions. Arbitrary choices had to be made for the models, such as the choice of period length for the blocks of BMM and the choice of in-sample and out-of-sample for our analyses, which may impact our results.

Future research should consider expanding the dataset to include more varied market conditions and integrating additional financial metrics to enhance the robustness of the models. This would involve longer time frames and a greater diversity of financial instruments to better capture the dynamic nature of markets. Combine other risk assessment metrics such as Expected Shortfall to work with coherent risk measure. Further study of the POT model, which has produced excellent results, coupled with other models could perform even better and try to extract even more robust estimates.

Taken as a whole, this work offers a good basis for future research, combining both the theoretical aspects developed and the application of models to best protect ourselves from the extreme losses to which we are exposed as financial players. In the hope that the methods explored here will contribute to better risk management, enabling us to navigate more serenely on these volatile and uncertain markets.

Bibliography

Alentorn, A., & Markose, S. (2008). Generalized Extreme Value Distribution and Extreme Economic Value at Risk (EE-VaR). Retrieved from https://www.researchgate.net/publication/226073080_Generalized_Extreme_Value_Distribution_and_Extreme_Economic_Value_at_Risk_EE-VaR

Artzner, P., Delbaen, F., Eber, J. M., & Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, 9(3), 203-228. Retrieved from https://www.researchgate.net/publication/228281201_Coherent_Measures_of_Risk

Basel Committee on Banking Supervision. (2004). *International Convergence of Capital Measurement and Capital Standards*. Retrieved from https://www.bis.org/publ/bcbs128.htm

Basel Committee on Banking Supervision. (2005). *International Convergence of Capital Measurement and Capital Standards: A Revised Framework*. Retrieved from https://www.bis.org/publ/bcbs118.htm

Basel Committee on Banking Supervision. (2011). *Basel III: A global regulatory framework for more resilient banks and banking systems*. Retrieved from https://www.bis.org/publ/bcbs189.htm

Basel Committee on Banking Supervision. (2019). *Minimum Capital Requirements for Market Risk*. Retrieved from https://www.bis.org/bcbs/publ/d457.htm

Bank for International Settlements. (2021). *Minimum Capital Requirements for Market Risk*. Retrieved from https://www.bis.org/bcbs/publ/d457.htm

Brown, K., & Moles, P. (2014). *Credit Risk Management*. Wiley. Retrieved from https://www.wiley.com/en-us/Credit+Risk+Management-p-9781119995632

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics. Springer, London.

Davison, A. C., & Smith, R. L. (1990). Models for Exceedances over High Thresholds. *Journal of the Royal Statistical Society. Series B (Methodological)*, 52(3), 393-442. Retrieved from https://rss.onlinelibrary.wiley.com/doi/10.2307/2348100

Dowd, K. (2005). *Measuring Market Risk*. Wiley. Retrieved from https://www.amazon.com/Measuring-Market-Risk-Kevin-Dowd/dp/0470013036

Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Springer. Retrieved from https://www.springer.com/gp/book/9783540609315

Emmer, S., Kratz, M., & Tasche, D. (2015). What is the Best Risk Measure in Practice? A Comparison of Standard Measures. *Journal of Risk*, 17(2). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2706348

European Central Bank. (2008). *Financial Stability Review*. Retrieved from https://www.ecb.europa.eu/pub/pdf/other/financialstabilityreview200812en.pdf

Gumbel, E. J. (1958). *Statistical Theory of Extreme Values and Some Practical Applications*. Columbia University Press.

Holton, G. A. (2003). *Value-at-Risk: Theory and Practice*. Elsevier.

MathWorks. (2024). *Value at Risk estimation and backtesting*. Retrieved May 27, 2024, from https://nl.mathworks.com/help/risk/value-at-risk-estimation-and-backtesting.html

McNeil, A. J. (1999). *Extreme Value Theory for Risk Managers*. Department of Mathematics, ETH Zurich. Retrieved from https://people.math.ethz.ch/~embrecht/ftp/EVT_risk_1999.pdf

McNeil, A. J., Frey, R., & Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press. Retrieved from https://press.princeton.edu/books/hardcover/9780691122557/quantitative-risk-management

de Haan, L., & Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer. Retrieved from https://www.springer.com/gp/book/9780387344713

A. Appendix

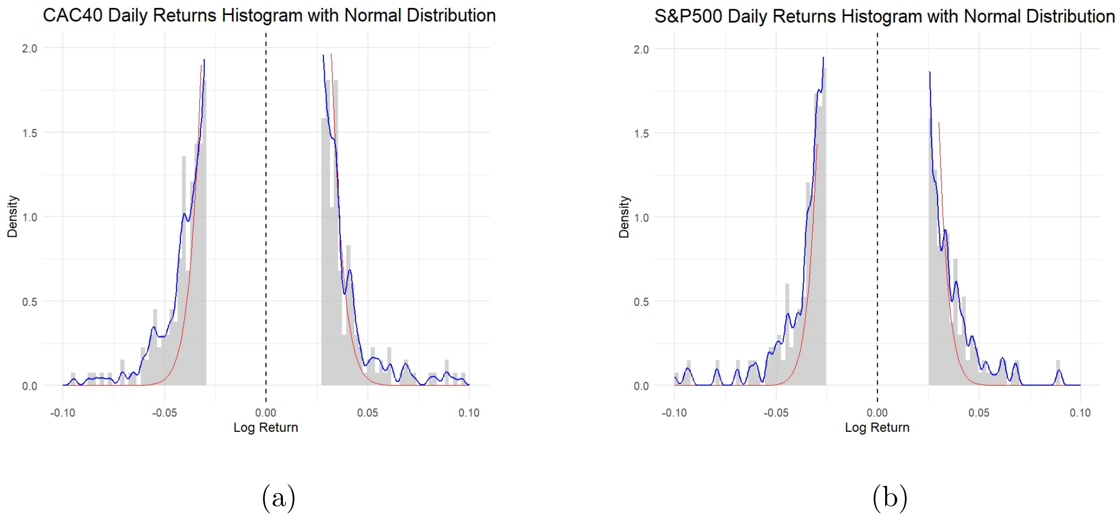


Figure A.1: Histograms of Daily Returns for (a) CAC40 and (b) S&P500 Indices Compared with Normal Distribution

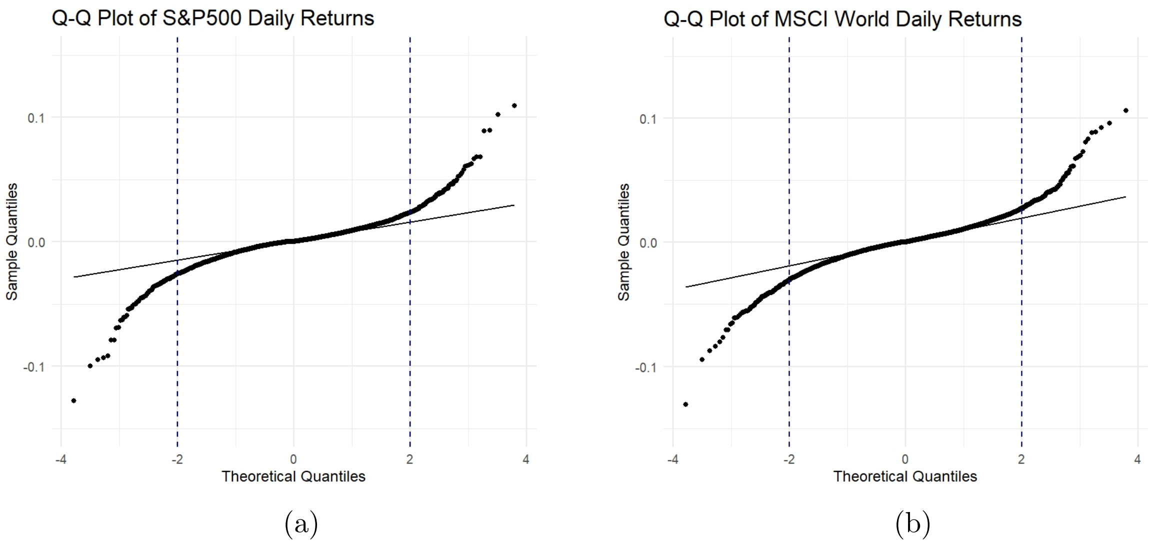
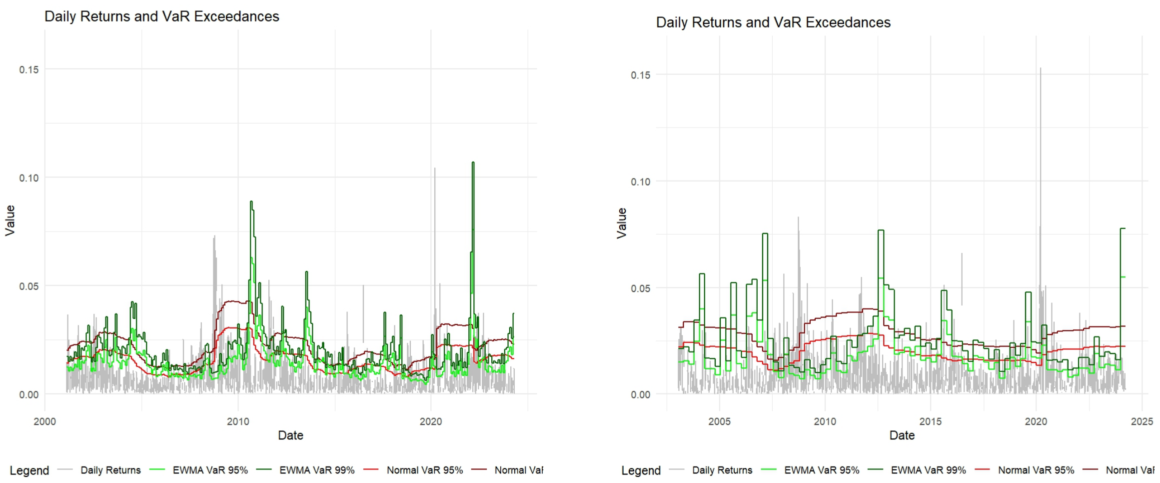
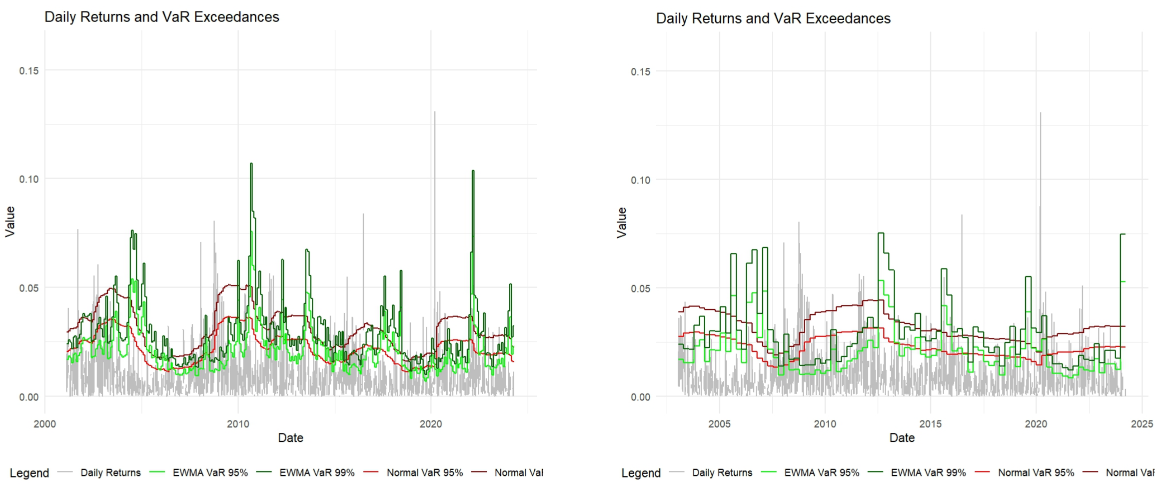


Figure A.2: Quantile-Quantile Plots for S&P500 and MSCI World Indices



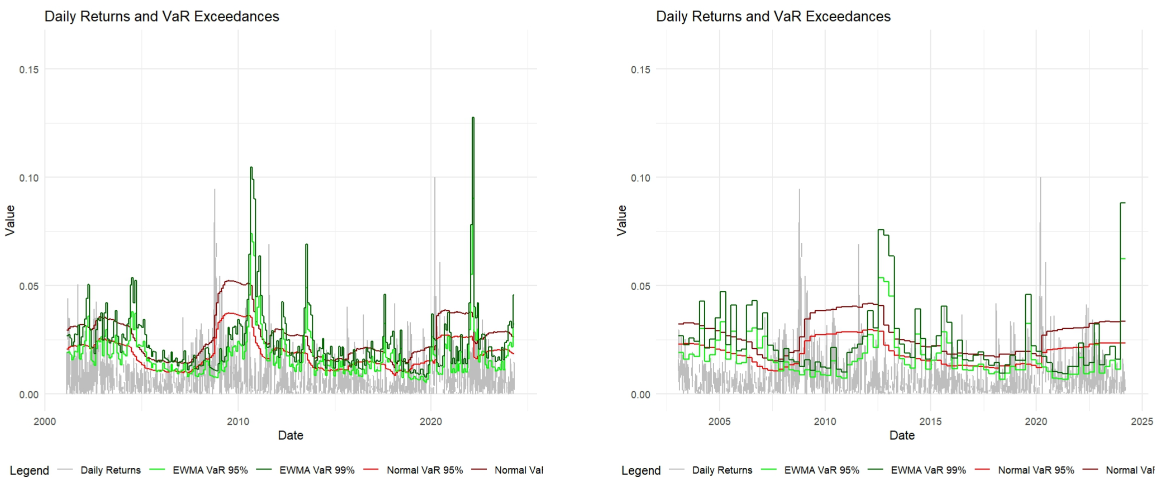
(a) (b)

Figure A.3: VaR Estimates Using Normal Distribution Method and EWMA Method for (a) monthly periods for MSCI World and (b) quarterly periods for BEL20



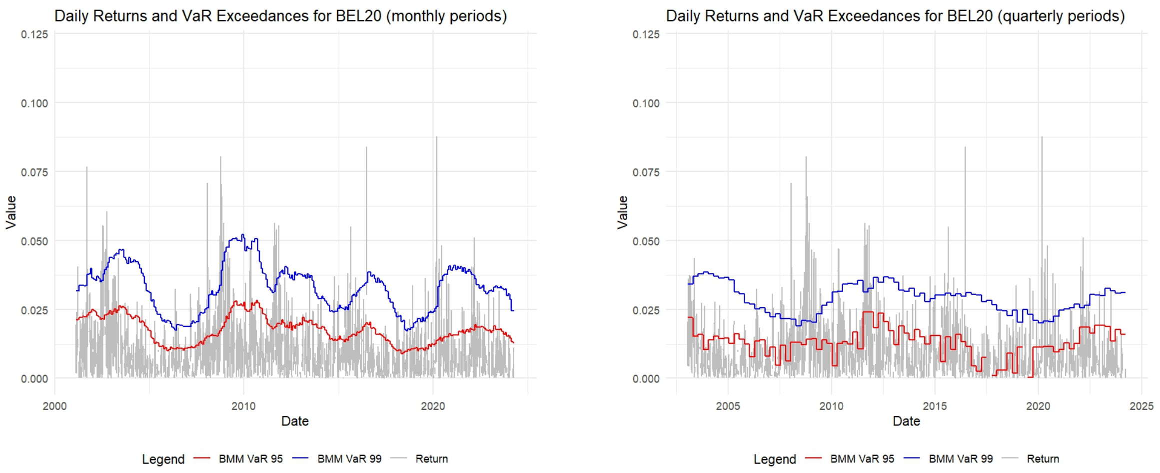
(a) (b)

Figure A.4: VaR Estimates Using Normal Distribution Method and EWMA Method for (a) monthly and (b) quarterly periods for CAC40



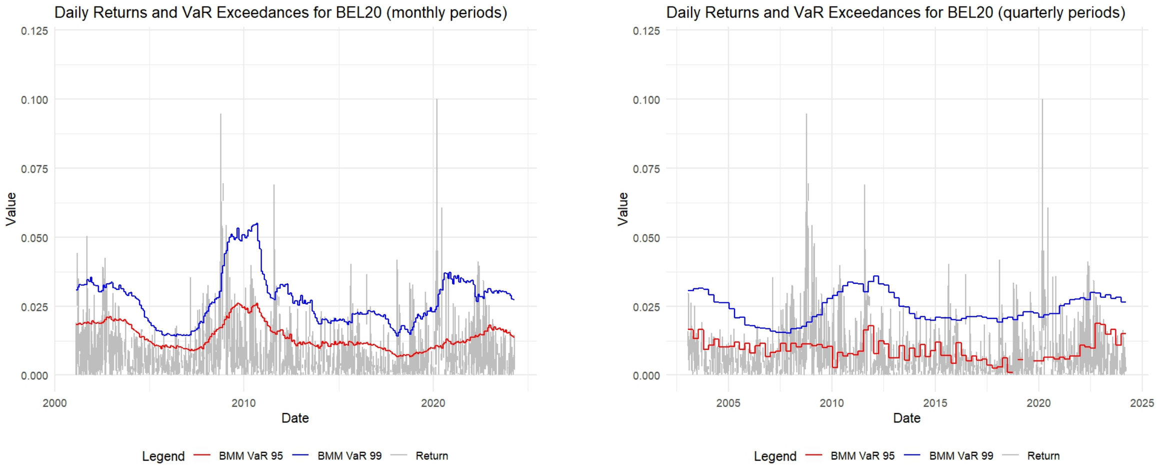
(a) (b)

Figure A.5: VaR Estimates Using Normal Distribution Method and EWMA Method for (a) monthly and (b) quarterly periods for S&P500



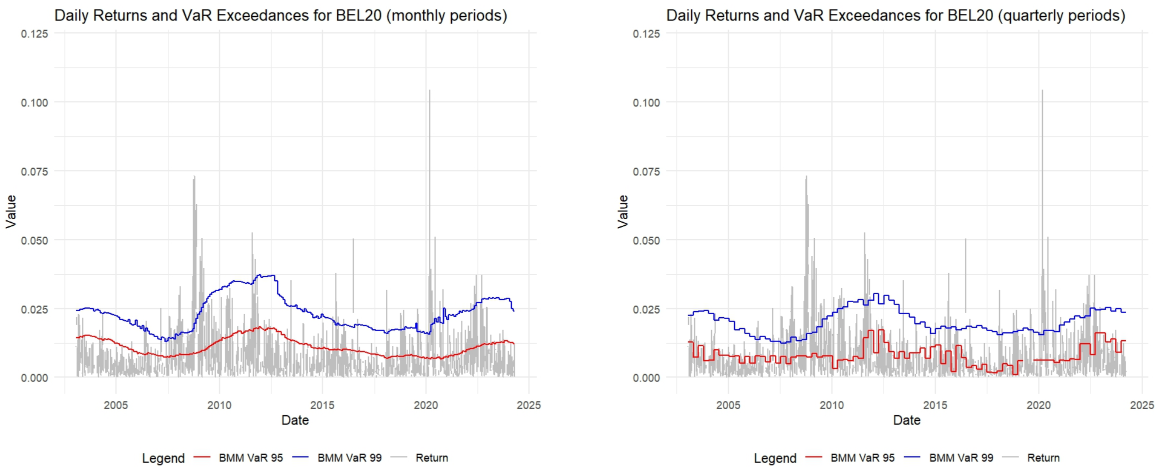
(a) (b)

Figure A.6: Peak-Over Threshold model VaR 95% and 99% estimations for CAC40 on (a) monthly and (b) quarterly periods



(a) (b)

Figure A.7: Block Maxima Model VaR 95% and 99% estimations for S&P500 on (a) monthly and (b) quarterly periods



(a) (b)

Figure A.8: Block Maxima Model VaR 95% and 99% estimations for MSCI World on (a) monthly and (b) quarterly periods

UNIVERSITÉ CATHOLIQUE DE LOUVAIN
Louvain School of Management

Place des Doyens, 1 bte L2.01.01, 1348 Louvain-la-Neuve
Boulevard Emile Devreux 6, 6000 Charleroi, Belgique
Chaussée de Binche 151, 7000 Mons, Belgique

www.uclouvain.be/lsm