Erasmus University Rotterdam

Erasmus School of Economics

Bachelor Thesis Econometrics and Operational Research

Comparative risk analysis of green and conventional stock markets using extreme quantile neural network estimators and bootstrapping

Koen van Woudenberg (531492)

Supervisor:	Prof. Dr. Chen Zhou
Second assessor:	Bram van Os
Date final version:	July 2, 2023

The views stated in this thesis are those of the author and not necessarily those of the supervisor, second assessor, Erasmus School of Economics or Erasmus University Rotterdam.

Abstract

The growing awareness among both private and institutional investors regarding the urgent need to address climate change has led to an increasing demand for green and sustainable investments. This study aims to examine whether investing in green stocks entails a higher level of risk compared to investing in conventional stocks. To explore this, a neural network-based extreme quantile estimator for heavy-tailed distributions is employed to investigate potential disparities in the occurrence of extreme events between the green stock market and the conventional stock market. The analysis encompasses the stock markets of the United States, Europe, and the Asia-Pacific region, and makes use of a bootstrapping comparison method. Concluding from the bootstrap confidence intervals, there are minor indications that the green stock markets in all the regions have a higher risk level compared to their conventional stock markets. However, ultimately, no significant difference in the level of risk is observed.

1 Introduction

The rising trend of investing in green stocks (Gupta and Jham, 2021), increases the importance of whether investments in green stocks are safe and not riskier than other conventional stocks. This study aims to give a clear picture of the effects of switching from a general stock investment strategy to one that focuses on green stocks. The research question is, 'Do investments in the green stock market entail higher risk compared to those in the conventional stock market?'. It is crucial to perform this comparison in the largest stock markets in the world. As a result, this paper examines the stock markets in the United States, Europe, and the Asia-Pacific region. Chakrabarti and Sen (2021) also examines the market risk of investments in green stocks in these regions.

The Intergovernmental Panel on Climate Change's (IPCC) most recent report, which was released in March 2023, stated that it will become harder to slow global warming. Billions of people would suffer as a result of the global warming. Particularly western nations, which are the biggest contributors to global warming, must alter their conduct, see Agarwal and Narain (2016). People must begin leading more environmentally responsible lives. An important way for them to change their carbon footprint is by not supporting big polluting companies. Consequently, through supporting businesses that are green and sustainable. Investing in stocks from green companies is a way to achieve this. Gupta and Jham (2021) address this rising popularity in green investing by showing the rapid growth of green bonds, green funds and green theme indices worldwide. And, not only individual investors, but also larger institutional investors like pension funds and insurance companies follow this trend, which creates even more popularity for the stocks from green companies, see Revelli and Viviani (2015) and Palacios-González and Chamorro-Mera (2018).

In order to address the research inquiry, this paper uses extreme quantile estimations. And to be more precise, it uses a new method introduced by Allouche et al. (2022b), which calculates extreme quantiles from heavy-tailed distributions with neural networks (NNs). Extreme quantiles are the tails of a probability distribution representing the highest and lowest values that can be expected given a certain probability. These quantiles can be used to compare the risk levels between green stocks and standard stocks due to the heavy-tail characteristics of stock time series. The method of Allouche et al. (2022b) establishes a way to estimate and eliminate all B first bias terms, when estimating the extreme quantiles with NNs. As a result, it is possible to approximate the extreme quantiles with a higher order. Thus, a more complex problem can be resolved with a more accurate outcome.

To assess the effectiveness of the methodology, this paper first conducts a simulation test. In this simulation, the estimator of Allouche et al. (2022b) is compared with a number of extreme quantile estimators from heavy-tailed distributions. The test demonstrates whether the new method works better. Following the simulation, the extreme quantiles of the (green) stock markets are estimated and compared to one another. However, simply looking at the estimates to see which is higher or lower cannot be used to compare estimates of NN extreme quantiles from various data sets. The paper uses bootstrapping to compare the green stock markets to the standard stock markets. By employing bootstrapping, this paper is able to construct a confidence interval for the difference between the extreme quantile estimates of two types of stocks.

The use of extreme quantiles is common in the field of risk management. Diebold et al. (2000) provide an overview of the benefits and drawbacks of using extreme value theory and come to the conclusion that it is highly applicable in the field of risk management. Chavez-Demoulin and Embrechts (2010) and Dutta and Biswas (2017) also show the importance of extreme quantiles in financial risk management. Dutta and Biswas (2017) compare multiple quantile estimators with each other and eventually come up with a new estimator. Until now, the method proposed by Allouche et al. (2022b) has not been applied in the area of risk management, thus this study aims to fill in this gap in the literature.

The simulation conducted in this paper demonstrates that the extreme quantile estimation method proposed by Allouche et al. (2022b) is among the most accurate methods currently available for estimating extreme quantiles from heavy-tailed distributions. Across various distributions and parameter combinations, this method consistently yields highly accurate estimates.

Consequently, this estimator is well-suited for the comparative study conducted in this paper. The results of the extreme quantile estimations and the bootstrapping confidence interval estimations, using multiple stock indices such as the DJSI, DJIA, and STOXX 600, reveal some noteworthy finding. Specifically, there are minor indications that investments in all the regions in green stock markets carry a higher level of risk compared to investments in standard stock markets. However, these indications are minimal and lack significance. Consequently, there is no substantial evidence to support a significant difference in risk levels between green stocks and conventional stocks.

The paper is structured as follows: Section 2 presents a comprehensive review of the relevant literature pertaining to this research topic. In Section 3, the methodology is elaborated, with a detailed explanation of the extreme neural network quantile estimator and the comparison methodology. The examination of the data used in the study is presented in Section 4. The results are reported and discussed in Section 5, while Section 6 concludes by answering the research question and providing recommendations for future research.

2 Literature

The existing literature extensively explores the combination of extreme quantiles and risk measurement. Haan and Ferreira (2006) provide a comprehensive introduction to the field of Extreme Value Theory (EVT), where they elucidate the core concept of the EVT. Their work demonstrates the utilization of the distribution's tail behavior as a means to predict extreme quantiles, and more importantly, to establish their connection with risk measurement. A lot of estimators of extreme quantiles for the risk of a time series have been introduced in combination with the EVT, see for example Dutta and Biswas (2017). They review some of these estimators and also introduce a new estimator, which shows encouraging finite sample performances.

Furthermore, Chavez-Demoulin and Embrechts (2010) and Chavez-Demoulin et al. (2014) also delve deeper into the relationship between extreme quantiles and risk. In the aftermath of the 2008 market crash, Chavez-Demoulin and Embrechts (2010) investigate whether existing methods can effectively estimate extreme events, ultimately concluding that statisticians needed to update their estimation techniques in the subsequent years. In a related article, Chavez-Demoulin et al. (2014) propose an extreme-quantile approach to estimating conditional risk measures for stock price time series.

There are dozens of extreme quantile estimators for heavy-tailed distributions. One of the most famous ones is the Weissman estimator, see Weissman (1978). This estimator uses a constant to approximate a slowly-varying function, which could be not very precise in practice.

After the introduction of the Weissman estimator, new and refined estimators followed soon, such as the Corrected Weissman estimator, the Refined Weissman estimator (Allouche et al., 2022a), the Corrected Hill estimator (Gomes et al., 2009), etc. The use of a neural network for the estimation of extreme quantiles is not common in the existing literature. Pasche and Engelke (2022) introduce an extreme quantile regression network to calculate the behavior of extreme events.

In this paper, extreme quantile estimations are employed to assess the disparity in risk levels between investments in green stocks and standard stocks. According to Yousaf et al. (2022), investing in environmentally friendly stocks is not just a nice luxury, but it is also a wise risk-averse action. Chakrabarti and Sen (2021) reveal that whereas green stocks in the United States and Europe exhibit a substantial volatility spillover with the market, they do not in the Asia-Pacific region. They also discover that green stocks are remarkably resilient during stock market crashes. But eventually, they come to the conclusion that the risk in green stocks is due to their chaotic nature.

3 Methodology

In order to assess the risk levels of investments, it is necessary to have a method for estimating extreme quantiles derived from heavy-tailed distributions. In this study, I utilize an independent and identically distributed (i.i.d.) sample, denoted as $X_1, ..., X_n$, drawn from an unknown cumulative distribution function (c.d.f.) represented by F, where n refers to the size of the training dataset. The order statistics of this sample are denoted as $X_{1,n} \leq ... \leq X_{n,n}$. The quantile function is defined as $q(\cdot) := F^{\leftarrow}(\cdot) = \inf x \in \mathbb{R} : F(x) \geq \cdot$, where the extreme level is represented by $1 - \alpha_n$, and it is required that $n\alpha_n \to 0$ as $n \to \infty$.

This paper aims to estimate and compare extreme quantiles across multiple datasets. By comparing these extreme quantiles, a more comprehensive understanding can be obtained regarding the relative risk levels between different investments. To assess the potential presence of similarities or differences, I formulate a null hypothesis and an alternative hypothesis. The hypotheses are as follows:

$$\begin{aligned} H_0 &: q^{(1)} - q^{(2)} = 0, \\ H_a &: q^{(1)} - q^{(2)} > 0, \end{aligned} \tag{3.1}$$

where $q^{(1)}$ and $q^{(2)}$ are the two extreme quantiles from the different heavy tailed-distributions.

3.1 Unconditional extreme quantiles

Allouche et al. (2022b) propose a new method to estimate the extreme quantiles from heavy tailed distributions with a neural network (NN). They suggest the idea to create a connection between the quantile that has to be estimated, $q(1-\alpha_n)$, and an intermediate quantile, $q(1-\delta_n)$, where $k := n\delta_n \to \infty$ as $n \to \infty$. Allouche et al. (2022b) utilize the log-spacing function given as,

$$(x_1, x_2) \in \mathbb{R}_+^2 \longmapsto f(x_1, x_2) = \log U(exp(x_1 + x_2)) - \log U(exp(x_2)) = \gamma x_1 + \psi(x_1 + x_2), \quad (3.2)$$

to help in the estimation process of the NN extreme quantiles. $U(t) := q(1 - \frac{1}{t})$ is the tail quantile function for all $t > 1$, and γ is the tail-index, see for example De Haan and Peng (1998).

Allouche et al. (2022b) construct an alternative representation of the log-spacing function 3.2. They focus on an one-hidden layer feedforward NN, which uses exponential linear units as activation functions (eLU), see for an example of eLU Clevert et al. (2015). By leveraging this NN architecture, I am able to estimate extreme quantiles from heavy-tailed distributions for the B-th order condition. The method effectively estimates and eliminates the first B bias terms, resulting in a precise approximation of the log-spacing function and, consequently, an accurate estimation of the extreme quantile.

I use the NN extreme quantile estimator, as stated in Allouche et al. (2022b),

$$\hat{q}_{\hat{\Psi}}^{NN_B}(1-\alpha_n; 1-\delta_n) := X_{n-k+1,n}\, exp\left(\tilde{f}_{\hat{\Psi}}^{NN_B}(log(\delta_n/\alpha_n), log(1/\delta_n))\right), \quad (3.3)$$

where

$$\tilde{f}_{\hat{\Psi}}^{NN_B}(x_1, x_2) = \tilde{m}_0 x_1 + \tilde{\psi}_{\tilde{\theta}}^{NN_B}(x_1, x_2), \quad (3.4)$$

is the NN construction of the log-spacing function, for all $\tilde{\psi} = (\tilde{m}_0, \tilde{\theta}) \in \Psi := \mathbb{R}_+ \times \Theta$. The NN eLU approximation of ψ is stated as in Allouche et al. (2022b),

$$\tilde{\psi}_{\tilde{\theta}}^{NN_B}(x_1, x_2) := \sum_{i=1}^{B(B-1)/2} m_i^{(1)} \left(\lambda^e(m_i^{(2)} x_1 + m_i^{(3)} x_2) - \lambda^e(m_i^{(4)} x_2)\right), \quad (3.5)$$

with $\theta = \left\{(m_i^{(1)}, m_i^{(2)}, m_i^{(3)}, m_i^{(4)}), i = 1, ..., B(B-1)/2\right\} \in \Theta := (\mathbb{R} \times \mathbb{R}_-^3)^{B(B-1)/2}$. B is the amount of order conditions, the number of bias terms that will be removed when estimating the extreme quantile with NN. The eLU activation function is denoted as $\lambda^e(\cdot)$.

The NN approximation 3.5 is derived by optimizing the construction $\tilde{f}_{\tilde{\psi}}^{NN_B}$ of the log-spacing function to align with the data. This optimization process involves minimizing the distance between two estimations of the $N = (n-1)(n-2)/2$ log-spacings, as demonstrated in the study by Allouche et al. (2022b):

$$\hat{\psi} = \arg\min_{\tilde{\psi}\in\Psi} \frac{1}{N}\sum_{k=2}^{n-1}\sum_{j=1}^{k-1}\left|\hat{D}_{j,k} - \tilde{f}_{\tilde{\psi}}^{NN_B}(log(k/j), log(n/k))\right|^s, \tag{3.6}$$

with $s \in \{1,2\}$ and $\hat{D}_{j,k} := log(X_{n-j+1,n}) - log(X_{n-k+1,n})$. $\hat{D}_{j,k}$ is a data-driven estimation of $log(q(1-j/n)) - log(q(1-k/n))$.

3.2 Simulation

In order to verify whether the estimation strategy using NNs is accurate, I use a simulation. I partially replicate and extend the simulation from Allouche et al. (2022b). In the simulation, extreme quantile estimates from NNs are compared to extreme quantile estimates from other well-known estimators. I study the Weissman estimator (W), the Refined Weissman estimator (RW), the Corrected Weissman estimator (CW), the Corrected-Hill estimator (CH), and the partially reduced bias estimators (CH_p and PRB_p). I use these estimators and the NN estimator on simulated data sets from three heavy tailed distributions, the Burr distribution (Tadikamalla, 1980), the NHW distribution (Allouche et al., 2022b), and the Student distribution (Bening and Korolev, 2005). The parameters γ and ρ_2 have an effect on all three distributions. γ is referred to as the tail-index. A distribution gets a heavier tail when γ is larger. The parameter ρ_2 affects the bias of the extreme quantile estimators. When ρ_2 increases, the asymptotic bias increases, as stated in Allouche et al. (2022b).

The performance of the estimators is analysed using the Relative median-squared error (RMedSE). I use the RMedSE denoted as in Allouche et al. (2022b),

$$RMedSE\left(\hat{q}_{\hat{\Psi}}, \frac{1}{2n}\right) = \underset{r\in\{1,\ldots,R\}}{median}\left[\left(\frac{\hat{q}_{\hat{\Psi}}^{(r)}(1-\frac{1}{2n}; 1-\frac{k^*(r)}{n})}{q(1-\frac{1}{2n})} - 1\right)^2\right], \tag{3.7}$$

with $\hat{q}_{\hat{\Psi}}^{(r)}(1-\frac{1}{2n}; 1-\frac{k^*(r)}{n})$ being an estimator of $q(1-\frac{1}{2n})$ with anchor index $k^*(r)$. R is the amount of replications in the simulation, each replication is a different data sample from a distribution. The estimator with the lowest RMedSE for a given distribution with a given γ and ρ_2, is the best performing estimator.

3.3 Bootstrapping

I propose a method to compare two NN extreme quantile estimates from different time series data sets using bootstrapping. Assume there are two time series data sets, $(X_1^{(1)}, ..., X_n^{(1)})$ and $(X_1^{(2)}, ..., X_n^{(2)})$, from which I want to compare the estimated NN extreme quantiles, with n

being the length of the time series. From each data set I take L bootstrap samples, creating $2 \times L$ groups of bootstrap samples. To not lose possible dependency between the two time series, I create a bootstrap $\{i_1, ..., i_n\}$ with $i \in \{1, ..., n\}$. With this bootstrap, I make the bootstrap samples $(X_{i_1}^{(1)}, ..., X_{i_n}^{(1)})$ and $(X_{i_1}^{(2)}, ..., X_{i_n}^{(2)})$. Now, the extreme quantile from each bootstrap sample can be estimated, creating two samples of L NN extreme quantile estimates, $Q^{(1)} = (\hat{q}_{\hat{\Psi},1}^{NN_B,(1)}, ..., \hat{q}_{\hat{\Psi},L}^{NN_B,(1)})$ and $Q^{(2)} = (\hat{q}_{\hat{\Psi},1}^{NN_B,(2)}, ..., \hat{q}_{\hat{\Psi},L}^{NN_B,(2)})$. Using these samples, I can now compute the bootstrap confidence intervals of the estimates $\hat{q}_{\hat{\Psi}}^{NN_B,(1)}$ and $\hat{q}_{\hat{\Psi}}^{NN_B,(2)}$, denoted as,

$$(Q_{low}, Q_{upper}), \tag{3.8}$$

with Q_{low} and Q_{upper} representing the lower and upper bounds of the $100(1-\alpha)\%$ confidence interval for $Q \in \{Q^{(1)}, Q^{(2)}\}$. I can subsequently compare these intervals.

To counteract the possibility of overlapping confidence intervals, I create a third sample by calculating for each $l \in \{1, ..., L\}$ the difference between the two estimated extreme quantiles,

$$\hat{\tau}_l = \hat{q}_{\hat{\Psi},l}^{NN_B,(1)} - \hat{q}_{\hat{\Psi},l}^{NN_B,(2)}. \tag{3.9}$$

The calculation of the difference creates the sample, $T = (\hat{\tau}_1, ..., \hat{\tau}_l)$. This is the sampling distribution from the difference between the two time series. The characteristics and the confidence interval can now be obtained from the sample. The differential bootstrap interval can be obtained as,

$$(T_{low}, T_{upper}), \tag{3.10}$$

where T_{low} and T_{upper} are the lower and upper bounds of the $100(1-\alpha)\%$ confidence interval.

4 Data

In order to analyze the risk differential between green stocks and standard stocks, I need data from several types of stock indexes. The indexes provide a comprehensive picture of the stock market. Additionally, I examine how risk varies in different parts of the world. I therefore require (green) stock indices from the United States, Europe, and the Asia-Pacific region.

The Dow-Jones-Sustainability-Indexes (DJSI) are a group of indexes that were introduced in 1999 and include the largest companies in the world based on economic, environmental, and social standards. There are indexes for many parts of the world in the DJSI. The DJSI-Europe represent 20% of the 600 sustainable top companies in Europe. The top 40 businesses in the

United States in terms of sustainability are included in the DJSI-US. The DJSI-Asia-Pacific consists of 20% of the 600 sustainable leading companies in the Asia-Pacific region. The official website of Dow-Jones Sustainability Indexes provides access to the historical daily data of these indexes.

I select the STOXX Europe 600 Index, the Dow Jones Industrial Average (DJIA), and the STOXX Asia-Pacific 600 Index as the standard stock market indices for Europe, the United States, and the Asia-Pacific region, respectively. The STOXX Europe 600 Index covers 600 large, mid and small companies from eighteen European countries. The DJIA is a price-weighted average made up of 30 illustrious firms listed on the New York Stock Exchange. The STOXX 600 Asia-Pacific contains the 600 largest companies in the Asia-Pacific region.

I obtain a data sample from each index containing the end-of-the-day trading prices covering the period from January 1, 2020, to December 31, 2022 for the United States and the Asia-Pacific region, and April 1, 2020, to December 31, 2022 for Europe. The Covid-19 peaks are present in the data sets from the United States and the Asia-Pacific region but not in the data set from Europe. The samples size of the data sets are $n = 756$ for the United States indices, $n = 709$ for the indices from Europe, and $n = 773$ for the Asia-Pacific region indices.

I need to confirm that the time series of the stock index prices is stationary before I can estimate the extreme quantiles using a NN. To check for stationarity, I employ the Augmented Dickey Fuller (ADF) test. The test results are displayed in Table 1.

Table 1: The Augmented Dickey Fuller test statistics for the data sets of six stock indices.

	DJIA	STOXX EU	STOXX Asia	DJSI-US	DJSI-EU	DJSI-Asia
Test statistic	-1.602	-2.600	-1.691	-1.433	-2.595	-1.238
p-value	0.481	0.093	0.435	0.567	0.095	0.660

Table 1 demonstrates that there is no statistically significant evidence at a 5% level for any of the six indices that the indices are stationary. Every p-value is above 0.05. I choose to use the log-returns ($log(P_t/P_{t-1})$) of the stock index prices as a result. This creates six stationary time series. Table 2 displays the descriptive statistics of the stationary time series from the six stock indices.

Table 2 reveals that the means of all the indices are very low, meaning that the average percentage change is very little, between the 0 and 0.1%. Another important finding in the descriptive statistics is that the kurtosis is high for all the indices. They all exceed a kurtosis of 4. This demonstrates that the log-returns data samples are heavy-tailed distributed. In addition to this, it is noteworthy that the United States' kurtosis and maximum are larger and

Table 2: Descriptive statistics of the percentile log returns ($log(P_t/P_{t-1}) \times 100$) time series from six stock indices

	Mean	Maximum	Minimum	St. Dev.	Skewness	Kurtosis
DJIA	0.02	10.76	-13.84	1.59	-0.9005	18.9752
DJSI-US	0.02	9.71	-13.12	1.62	-0.7195	15.2390
STOXX Europe 600	0.04	4.58	-4.18	1.10	-0.2135	4.8394
DJSI-Europe	0.03	4.44	-3.79	1.02	-0.1619	4.9964
STOXX Asia-Pacific 600	0.00	6.30	-4.91	1.04	-0.0608	6.4830
DJSI-Asia-Pacific	0.00	7.28	-6.31	1.18	0.0259	6.7559

its minimum is lower than those of the other two regions. This is most likely because the Covid-19 outbreak in March 2020 resulted in significant daily swings in stock prices in the United States. This effect was lower in the Asia-Pacific region, and was not included in the data sample for Europe.

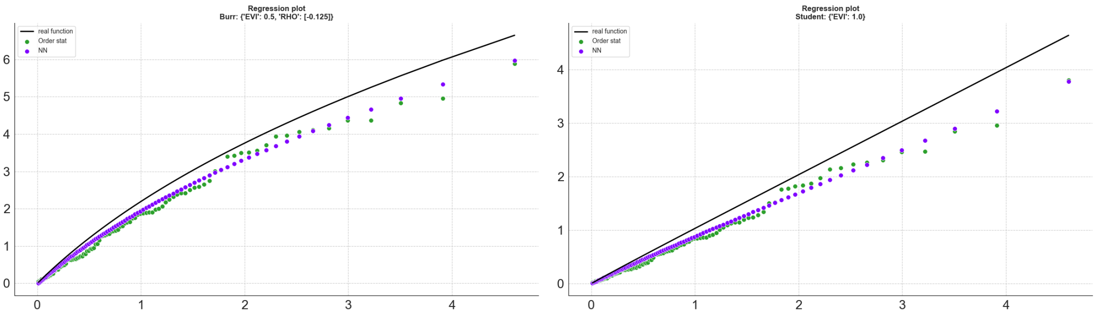
5 Results

In this section, I present the simulation results and the quantile estimates for the three regions, the United States, Europe, and the Asia-Pacific region. I implemented the code to train the neural networks and calculate the extreme quantiles using Python and the PyTorch package. With the parameters $\beta_1 = 0.9$ and $\beta_2 = 0.999$, I used the optimizer Adam Kingma and Ba (2014). For the hyperparameters in the neural network, I applied $B = 4$, giving 12 neurons, and a batch size of 1024. In all the estimations, I used $k_1 = \frac{3n}{100}$ and $k_2 = \frac{3n}{4}$. I used Eviews to obtain the statistical characteristics.

5.1 Simulation

In the simulation, I compare the NN extreme quantile estimator with several well-known quantile estimators. For each distribution, Burr, NHW and Student, multiple combinations of the parameters γ and ρ_2 are used to simulate data sets. The Burr and NHW distributions use a combination of $\gamma \in \{1/4, 1/2\}$ and $\rho_2 \in \{-1/8, -1/4, -1/2\}$, which creates in total 12 configurations. The Student distribution uses a combination of $\gamma \in \{1/4, 1/2, 1\}$ and $\rho_2 = -2\gamma$, thus 3 configurations. 10 replicated data sets of size $n = 500$ are simulated for each configuration, and for each data set are the quantile estimates calculated.

Figure 1 shows a comparison of the NN estimator of the log-spacing function with the theoretical function and an empirical point wise estimator. As shown in Figure 1, the NN extreme quantile estimate outperforms the empirical point-wise estimator for all three distributions. This is the first proof that the NN estimator is preferable.



(a) (b)

Figure 1: A NN estimation of a log-spacing function (purple dots) compared to the theoretical log-spacing function (black line) and an empirical point wise estimation of the log-spacing function (green dots) for a configuration of the Burr and Student distribution.

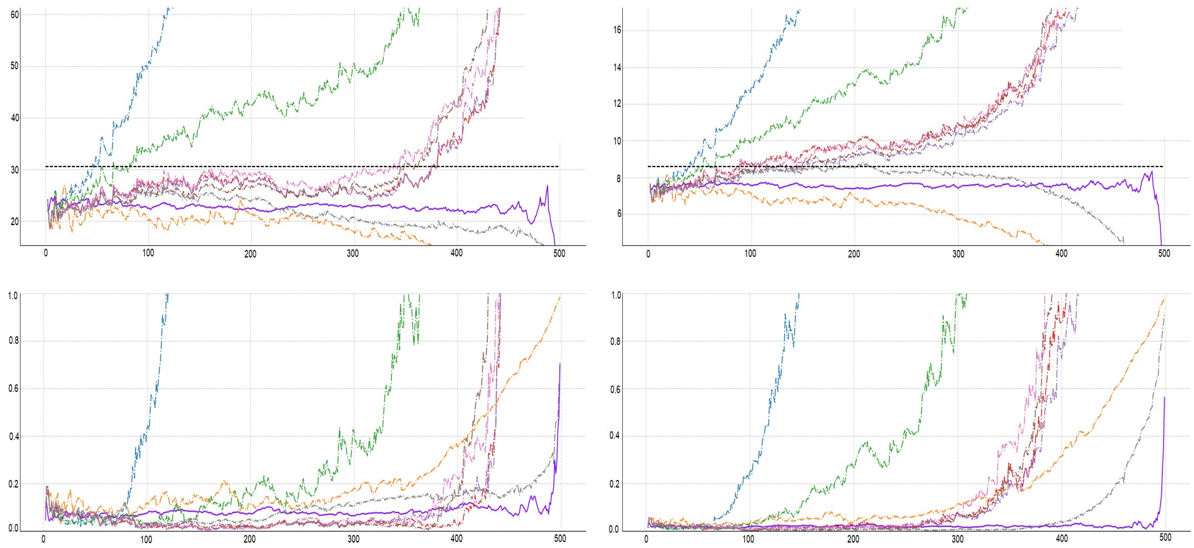
Table 3 presents the results from the simulation of the different distributions.

Table 3: The RMedSE of seven extreme quantile estimators of quantile $q(1-\frac{1}{2n})$ on three distributions with multiple combinations of γ and ρ_2. A RMedSE higher than 1 is not noted, and the best RMedSE of every configuration is emphasized in bold.

			NN	W	RW	CW	CH	CH_p	PRB_p
Burr	$\gamma = 1/4$	$\rho_2 = -1/8$	0.9944	-	**0.8589**	-	-	-	-
		$\rho_2 = -1/2$	0.0169	-	0.0392	0.0318	**0.0070**	0.0112	0.0427
		$\rho_2 = -1$	0.0165	0.1097	0.0287	**0.0163**	0.0711	0.0694	0.0872
	$\gamma = 1/2$	$\rho_2 = -1/8$	**0.2079**	-	0.3467	-	-	-	-
		$\rho_2 = -1/2$	0.0790	-	0.1287	0.1562	0.0260	**0.0257**	0.0375
		$\rho_2 = -1$	0.0585	0.6000	0.0927	**0.0571**	0.2119	0.2122	0.1988
NHW	$\gamma = 1/4$	$\rho_2 = -1/8$	**0.0237**	-	0.1019	0.2541	0.5342	0.5243	0.5181
		$\rho_2 = -1/2$	0.0347	0.4006	0.0694	**0.0067**	0.0709	0.1091	0.1003
		$\rho_2 = -1$	0.0194	**0.0067**	0.0648	0.0641	0.1040	0.1040	0.1453
	$\gamma = 1/2$	$\rho_2 = -1/8$	**0.0813**	-	0.1984	0.6301	0.8481	0.8378	0.8769
		$\rho_2 = -1/2$	0.0979	0.3370	0.1283	**0.0528**	0.2512	0.2350	0.2822
		$\rho_2 = -1$	0.0656	0.162	**0.0518**	0.1242	0.1748	0.1970	0.1665
Student	$\gamma = 1/4$	$\rho_2 = -2\gamma$	0.0194	-	0.0672	0.3224	0.0212	**0.0111**	0.0258
	$\gamma = 1/2$		0.0528	-	0.1538	**0.0142**	0.1466	0.1574	0.1105
	$\gamma = 1$		**0.1576**	0.2695	0.3702	0.4572	0.6913	0.6202	0.6290

Table 3 reinforces this argument. Out of the entire 15 configurations, the NN estimator yields the best estimates for 4 of them. Comparing the results with the results from Allouche et al. (2022b), I can see that the parts that I replicated are very comparable. Although I do not have the exact identical RMedSEs, I do have the same best estimator for the configurations. I was only able to perform 10 replications each configuration, whereas Allouche et al. (2022b) were able to perform 500 replications per configuration, which accounts for the difference in the same exact number. In the configurations that I created which were not in Allouche et al. (2022b), I observe that the NN quantile estimator is not always the best performing estimator. The NN estimator always performs as one of the best, but definitely not always the best. I conclude this from the configurations where $\rho_2 \in \{-1/2, -1\}$.

Figure 2 shows for some of the configurations the median and the RMedSE of the estimators as function of $k \in \{1, ..., n-1\}$.



(a) (b)

Figure 2: An illustration of the median (upper) and RMedSE (lower) for several estimators of quantile $q(1-\frac{1}{2n})$ (black dashed line) on simulated data sets of size $n = 500$ from a Burr distribution (a) with $\gamma = 0.5$ and $\rho_2 = 0.5$, and a Student distribution (b) with $\gamma = 0.25$. The median and RMedSE are computed as functions of $k \in \{2, ..., n-1\}$ with $R = 10$ replications. The estimators are NN (purple), W (blue), RW (orange), CW (green), CH (red), CH_p (violet) and PRB_p (pink).

The figure illustrates that, when considering bias and RMedSE, the NN estimator is by far the most stable estimator in all of the chosen configurations. Even when Table 3 does not show that the NN estimator is the most accurate estimator in certain situation, it might still be the most appreciated one because of its stability over a wide range of k values.

5.2 Risk analysis for the United States

I made a group of $L = 100$ bootstrap samples for the DJIA and the DJSI-US, giving a total of 200 bootstrap samples with a size of $n = 756$. Each sample used a 284,635-piece training data set to develop its neural network. After I estimated all the extreme quantiles of $q(1-\frac{1}{2n})$, I obtained the statistical characteristics of the two distributions from the extreme quantile estimates and of the distribution derived from the differences. Table 4 shows the statistical characteristics.

Table 4: Statistical characteristics of the extreme quantile estimates distributions from the US standard stock index (DJIA), the US green stock index (DJSI-US), and their difference. The estimates are from the quantile $q(1 - \frac{1}{2n})$. The US standard and green distributions are made with each 100 bootstrap samples of size $n = 756$.

	Mean	Maximum	Minimum	St. Dev.	Skewness	Kurtosis
US Standard	0.1367	0.3872	0.0027	0.0921	0.4937	2.5873
US Green	0.1435	0.4003	0.0140	0.0847	1.1185	3.7346
US Green-US Standard	0.0068	0.2627	-0.3213	0.1162	-0.0477	2.8505

In Table 4, I observe that the means of the two quantile estimates distributions are in close proximity. They differ by 0.0068. Additionally, there is not much of a difference between the maximum and minimum, 0.3872 versus 0.4003 and 0.0027 versus 0.0140, respectively. Both the distributions do not have a perfect Skewness and Kurtosis for them to be normally distributed. With the distributions, I create the confidence intervals for the estimated extreme quantiles from the standard stock index and the green stock index. The 90% confidence interval for the standard stock index is as follows,

$$\left(Q_{low}^{US}, Q_{upper}^{US}\right) = (0.0126, 0.3082), \tag{5.1}$$

with Q_{low}^{US} and Q_{upper}^{US} being the lower and upper bound. The 90% confidence interval for the green stock index is,

$$\left(Q_{low}^{US-Green}, Q_{upper}^{US-Green}\right) = (0.0308, 0.3459). \tag{5.2}$$

The confidence intervals are constructed by taking 5th and 95th extreme quantile from the corresponding sorted distribution. The estimated extreme quantiles for the original data sets are $\hat{q}_{US-St}^{NN} = 0.1016$ and $\hat{q}_{US-Gr}^{NN} = 0.0842$.

The confidence intervals for the standard stock index, with 90% confidence, indicate that the true extreme quantile falls between 0.0126 and 0.3082. Similarly, for the green stock index, with the same confidence level, the real extreme quantile falls within the range of 0.0308 and 0.3459. By further analyzing the two 90% confidence intervals, it is evident that the upper and lower bounds of the interval associated with the green stock index are notably higher compared to those of the standard stock index. To further investigate this difference, I analyze the distribution of the differences.

I observe that the mean of the distribution from the differences between the quantiles is slightly above zero. This could indicate that the US Green stock index might be a bit more riskier than the US standard stock index. To conduct further research, I look more closely at the confidence interval of this distribution. This distribution is normally distributed, as evidenced by the Skewness and Kurtosis. The 90% confidence interval is as follows,

$$\left(T_{low}^{US}, T_{upper}^{US}\right) = (-0.1888, 0.2216). \tag{5.3}$$

The 90% confidence interval demonstrates that the real extreme quantile of the difference between the green and standard stock indices falls between -0.1888 and 0.2216 with 90% confidence. Thus, it can be lower than zero, meaning that the standard stock index is riskier, or it can be higher, meaning that the green stock index is riskier. Their is no significant difference between the two.

5.3 Risk analysis for Europe

A total of 200 bootstrap samples were created, comprising $L = 100$ samples for the STOXX Europe 600 and $L = 100$ samples for the DJSI-Europe. Each bootstrap sample was constructed using a training dataset of size 250,278. Through estimating the extreme quantiles at a level of $q(1-\frac{1}{2n})$, I derived statistical characteristics for both distributions, as well as for the distribution resulting from their differences. The statistical characteristics obtained from these estimations are presented in Table 5.

Table 5: Statistical characteristics of the extreme quantile estimates distributions from the Europe standard stock index (STOXX-600), the Europe green stock index (DJSI-Europe), and their difference. The estimates are from the quantile $q(1-\frac{1}{2n})$. The Europe standard and green distributions are made with each 100 bootstrap samples of size $n = 709$.

	Mean	Maximum	Minimum	St. Dev.	Skewness	Kurtosis
EU Standard	0.0374	0.1168	0.0023	0.0221	0.8673	4.0033
EU Green	0.0393	0.1161	0.0052	0.0240	0.8547	3.1915
EU Green-EU Standard	0.0019	0.0660	-0.1011	0.0265	-0.3152	4.6707

The findings presented in Table 5 demonstrate that the means of the extreme quantile bootstrap samples for the standard stock index and the green stock index are relatively close to each other, with a difference of only 0.0019. However, it is noteworthy that the mean of the green index sample is slightly higher, suggesting a potential higher level of risk associated with investing in green stocks. Furthermore, examining the skewness of the two distributions reveals that they are both right-tailed, indicating a tendency towards higher values. Additionally, the kurtosis values being higher than 3 indicate some possible heavy-tailedness. Using the distributions, I construct the confidence interval for the estimated extreme quantiles of the standard and green stock indices. The 90% confidence interval for the standard stock index can be stated as,

$$\left(Q_{low}^{EU}, Q_{upper}^{EU}\right) = (0.0067, 0.0792), \tag{5.4}$$

with Q_{low}^{EU} and Q_{upper}^{EU} being the lower and upper bound. The 90% confidence interval for the green stock index is,

$$\left(Q_{low}^{EU-Green}, Q_{upper}^{EU-Green}\right) = (0.0094, 0.0827). \tag{5.5}$$

The confidence intervals are constructed by taking 5th and 95th extreme quantile from the corresponding sorted distribution. The estimated extreme quantiles for the original data sets are $\hat{q}_{EU-St}^{NN} = 0.0358$ and $\hat{q}_{EU-Gr}^{NN} = 0.0424$.

The standard stock index's confidence intervals show that the actual extreme quantile lies between 0.0067 and 0.0792, with 90% confidence. The actual extreme quantile for the green stock index falls between 0.0094 and 0.0827, with the same level of confidence. The 90% confidence intervals show that both the lower bound and the upper bound are higher for the Europe green stock index compared to the Europe standard stock index. To delve deeper into this discrepancy, an analysis of the distribution of the difference is conducted.

The mean value presented in Table 5 for the distribution derived from the differences between the quantiles strengthens the notion that the Europe green stock index carries a slightly higher level of risk. With a value of 0.0019, which surpasses zero, there is a possibility of an elevated risk associated with the green stock index compared to the standard stock index. To gain further insight into this thought, the confidence interval of this distribution is employed. The 90% confidence interval is as follows,

$$\left(T_{low}^{EU}, T_{upper}^{EU}\right) = (-0.0452, 0.0487). \tag{5.6}$$

The interval shows that, with 90 percent confidence, the true extreme quantile of the difference between the green and standard stock indexes lies between -0.0452 and 0.0487. Thus, there is no significant evidence that investing in Europe in green stocks is riskier than investing in conventional stocks.

5.4 Risk analysis for the Asia-pacific region

I created a set of $L = 100$ bootstrap samples for the STOXX Asia-Pacific 600 and the DJSI-Asia-Pacific, resulting in a total of 200 bootstrap samples with a size of $n = 773$. Each sample utilized a training dataset consisting of 297,606 data points to train its neural network. By estimating the extreme quantiles of $q(1 - \frac{1}{2n})$, I obtained the statistical properties of the two distributions based on the extreme quantile estimates, as well as the distribution derived from their differences. Table 6 presents these statistical characteristics.

Table 6: Statistical characteristics of the extreme quantile estimates distributions from the Asia-Pacific standard stock index (STOXX-600), the Asia-pacific green stock index (DJSI-Asia-Pacific), and their difference. The estimates are from the quantile $q(1 - \frac{1}{2n})$. The Asia-Pacific standard and green distributions are made with each 100 bootstrap samples of size $n = 773$.

	Mean	Maximum	Minimum	St. Dev.	Skewness	Kurtosis
Asia Standard	0.0629	0.1458	0.0009	0.0366	0.3260	2.4723
Asia Green	0.0630	0.1905	0.0036	0.0402	1.1497	4.2390
Asia Green-Asia Standard	0.0002	0.1768	-0.1391	0.0555	0.3778	3.7036

The means of the Asia-Pacific standard stock index and the Asia-Pacific green stock index, as presented in Table 6, demonstrate a negligible difference. The mean of the green stock index is only slightly higher by 0.0002. This minute disparity does not provide sufficient evidence to determine which index carries a higher level of risk. However, it is noteworthy that the skewness of the green stock index surpasses that of the standard stock index. This signifies a more pronounced right-tailed distribution in the green stock index, implying a propensity towards higher values. The 90% confidence intervals for the estimated extreme quantile of the standard stock index is as follows,

$$\left(Q_{low}^{Asia}, Q_{upper}^{Asia}\right) = (0.0074, 0.1348), \tag{5.7}$$

with Q_{low}^{Asia} and Q_{upper}^{Asia} being the lower and upper bound. The 90% confidence interval for the green stock index is,

$$\left(Q_{low}^{Asia-Green}, Q_{upper}^{Asia-Green}\right) = (0.0135, 0.1534). \tag{5.8}$$

By selecting the 5th and 95th extreme quantiles from the associated sorted distribution, the confidence intervals are created. The estimated extreme quantiles for the original data sets are $\hat{q}^{NN}_{Asia-St} = 0.0396$ and $\hat{q}^{NN}_{Asia-Gr} = 0.0496$.

According to the standard stock index's confidence intervals, there is a 90% chance that the real extreme quantile is located between 0.0074 and 0.1348. With the same level of confidence, the green stock index's real extreme quantile ranges from 0.0135 to 0.1534. The 90% confidence interval for the difference distribution is presented as,

$$\left(T^{Asia}_{low}, T^{Asia}_{upper}\right) = (-0.0943, 0.1072). \tag{5.9}$$

By examining the confidence interval, it becomes evident that the actual extreme quantile of the difference between the actual extreme quantiles of the green and standard stock indices lies between -0.0943 and 0.1072, with 90% confidence. It can be lower than zero, and higher than zero. Thus, there is no significant evidence that the extreme quantiles differ. The risk of green stocks do not differ significantly compared to the conventional stocks in the Asia-Pacific region.

5.5 Discussion

There is no clear evidence of a significant difference in the level of risk between investing in green stocks and standard stocks across the United States, Europe, and the Asia-Pacific region. This finding aligns with the research conducted by Chakrabarti and Sen (2021), which suggests that investing in green stocks is a risk-averse choice, with the only risk being their unpredictable nature. Interestingly, the mean values of the green stock indices and standard stock indices, as observed in the extreme quantile bootstrap samples for all three regions, are very close to each other. After evaluating the data used to estimate the extreme quantiles, I noticed a slight overlap between the standard stock index and the green stock index in all three regions. This indicates that there were a few companies that appeared in both indices.

In order to investigate whether the presence of overlapping companies impacts the difference in risk levels between the green and standard stock indices, I conducted a recalculation of the bootstrap samples specifically for the United States. This involved utilizing a new dataset containing the log returns of modified prices for the DJIA.

The modification process entailed adjusting the prices to include only non-green stocks. For each day, I computed the total market capitalization of all the companies within the DJIA, as well as the total market capitalization of the non-green companies. By calculating the ratio between these two market caps and multiplying it by the original daily price, I obtained the new daily price of the DJIA comprising solely non-green companies. The estimations derived from these new bootstrap samples are presented in Table 7.

Table 7: Statistical characteristics and 90% confidence intervals of the extreme quantile estimates distributions from the modified US standard stock index (DJIA), the US green stock index (DJSI-US), and their difference. The estimates are from the quantile $q(1-\frac{1}{2n})$. The US standard and green distributions are made with each 100 bootstrap samples of size $n = 751$.

	Mean	Maximum	Minimum	St. Dev.	Lower	Upper
US Standard	0.1378	0.3726	0.0004	0.0875	0.0222	0.3076
US Green	0.1391	0.5392	0.0039	0.0988	0.0194	0.3316
US Green-US Standard	0.0013	0.4931	-0.2302	0.1163	-0.2052	0.2139

The findings presented in Table 7 reveal that the revised estimations of extreme quantiles do not alter the conclusions drawn from the previous results. Once again, it is clear that the means of the two indices exhibit close proximity to each other. Additionally, the lower bound of the differential confidence interval is below zero, while the upper bound is above zero. These are all signs that there is no evidence of a significant difference between investing in green stocks or investing in conventional stocks.

6 Conclusion

In conclusion, this paper makes a valuable contribution to the existing literature by addressing the research question, 'Do investments in the green stock market entail higher risk compared to those in the standard stock market?'. The paper investigates this possible difference by analyzing three different markets, the United States's stock market, the European stock market, and the Asia-Pacific stock market. The analysis employs an unconditional extreme quantile from heavy-tailed distribution estimation method, complemented by neural networks, to gauge the magnitude of extreme events within the stock markets. To compare the green stock markets with the standard stock markets, bootstrapping is being used to create extreme quantile bootstrap distributions, from which confidence intervals are constructed.

Prior to the comparative analyses, a simulation study demonstrated that the NN extreme quantile estimation method is performing as one of the most accurate and stable extreme quantile estimator methods. This indicates a high level of accuracy for the extreme quantiles estimations of the stock markets. The confidence intervals, created from these estimations, for the United States, Europe, and the Asia-Pacific region, of the green stock markets exhibited small indications of a higher risk level compared to their respective standard stock markets. However, these indications were not significant. No significant difference in risk level was observed between the green and conventional stock markets. These findings highlight the reassurance that investors need not be concerned about encountering a higher risk level when investing in green stocks.

This research employed fixed parameters to estimate the extreme quantiles, including an order condition of $B = 4$ and specific hyperparameters in the neural network. Further exploration of this topic could involve investigating alternative values for these parameters, such as increasing the order condition or modifying the batch size of the neural network. Additionally, another extension could be about increasing the sample size of the bootstrap samples to 1000 or higher. This would improve the accuracy of the constructed confidence intervals. Furthermore, it would be worthwhile to investigate the possibility of constructing a confidence interval for the extreme quantile bootstrap distribution using its standard deviation as a statistical measure.

References

A. Agarwal and S. Narain. Environmental colonialism: The perverse politics of climate change. In *Global Environmental Politics: From Person to Planet*, pages 233–237. Routledge, 2016.

M. Allouche, J. El Methni, and S. Girard. A refined weissman estimator for extreme quantiles. *Extremes*, pages 1–28, 2022a.

M. Allouche, S. Girard, and E. Gobet. Estimation of extreme quantiles from heavy-tailed distributions with neural networks. working paper or preprint, 2022b. URL `https://hal.science/hal-03751980`.

V. E. Bening and V. Y. Korolev. On an application of the student distribution in the theory of probability and mathematical statistics. *Theory of Probability & Its Applications*, 49(3): 377–391, 2005.

G. Chakrabarti and C. Sen. Dynamic market risk of green stocks across regions: Where does the devil lie? *Journal of Cleaner Production*, 303:127028, 2021.

V. Chavez-Demoulin and P. Embrechts. Revisiting the edge, ten years on. *Communications in Statistics—Theory and Methods*, 39(8-9):1674–1688, 2010.

V. Chavez-Demoulin, P. Embrechts, and S. Sardy. Extreme-quantile tracking for financial time series. *Journal of Econometrics*, 181(1):44–52, 2014.

D.-A. Clevert, T. Unterthiner, and S. Hochreiter. Fast and accurate deep network learning by exponential linear units (elus). *preprint arXiv: 1511.07289*, 2015.

L. De Haan and L. Peng. Comparison of tail index estimators. *Statistica Neerlandica*, 52(1): 60–70, 1998.

F. X. Diebold, T. Schuermann, and J. D. Stroughair. Pitfalls and opportunities in the use of extreme value theory in risk management. *The Journal of Risk Finance*, 1(2):30–35, 2000.

S. Dutta and S. Biswas. Extreme quantile estimation based on financial time series. *Communications in Statistics-Simulation and Computation*, 46(6):4226–4243, 2017.

M. I. Gomes, D. Pestana, and F. Caeiro. A note on the asymptotic variance at optimal levels of a bias-corrected hill estimator. *Statistics & Probability Letters*, 79(3):295–303, 2009.

L. Gupta and J. Jham. Green investing: Impact of pro-environmental preferences on stock market valuations during turbulent periods. *Australasian Accounting, Business and Finance Journal*, 15(5):59–81, 2021.

L. Haan and A. Ferreira. *Extreme value theory: an introduction*, volume 3. Springer, 2006.

D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *preprint arXiv: 1412.6980*, 2014.

M. M. Palacios-González and A. Chamorro-Mera. Analysis of the predictive variables of the intention to invest in a socially responsible manner. *Journal of cleaner production*, 196:469–477, 2018.

O. C. Pasche and S. Engelke. Neural networks for extreme quantile regression with an application to forecasting of flood risk. *preprint arXiv: 2208.07590*, 2022.

C. Revelli and J.-L. Viviani. Financial performance of socially responsible investing (sri): what have we learned? a meta-analysis. *Business Ethics: A European Review*, 24(2):158–185, 2015.

P. R. Tadikamalla. A look at the burr and related distributions. *International Statistical Review/Revue Internationale de Statistique*, pages 337–344, 1980.

I. Weissman. Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364):812–815, 1978.

I. Yousaf, M. T. Suleman, and R. Demirer. Green investments: A luxury good or a financial necessity? *Energy Economics*, 105:105745, 2022.

7 Appendix

7.1 Code

There are two approaches to executing my code. The first method involves running the code for simulation purposes. In this scenario, it is crucial to incorporate the distribution that requires simulation into multiple functions. The second option entails running the code using your own dataset. In this case, there is no need to simulate any data samples. To execute the code, you need to run the "run.py" file. Initially, this file reads the dataset that will be utilized and then proceeds to read the "config.yaml" file, which contains all the pertinent information about the neural network and data samples. Afterward, a crucial aspect of the estimation process begins, namely, the training of the neural network. This involves creating a model using the "ExtrapolateNN" class in the "network.py" file. The model is then trained using the data, the order statistics of the data, and the specified hyperparameters provided in the "config.yaml" file. In the training, parameters are being estimated, saved and evaluated each iteration of the training process, via multiple functions.

Once the training is complete, all the relevant information is stored in a file to facilitate the subsequent phase of estimation. In this phase, the "model_evaluation" function is employed to assess the model's performance and identify the best criteria across all iterations. The parameters from the best iteration are utilized to compute the extreme quantile for each k anchor point using the "extrapolate" function. By extrapolating these extreme quantiles, the optimal extreme quantile corresponding to the most suitable k value is determined.