Mean univariate- GARCH VaR portfolio optimization: actual portfolio approach

Article (Published Version)


This version is available from Sussex Research Online: http://sro.sussex.ac.uk/59655/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
Mean-univariate GARCH VaR portfolio optimization: Actual portfolio approach

Vladimir Ranković, Mikica Drenovak, Branko Urosevic, Ranko Jelic

1. Introduction

Value-at-Risk (VaR) is defined as the loss associated with the low (typically first or fifth) percentile of the return distribution. The Basel II Capital Accord codifies VaR as the de-facto industry standard for the banking and insurance industries alike (see, BIS [3–5]). In particular, for the market risk exposure of banks, a bank’s internal VaR estimates corresponding to the actual portfolio, i.e. the portfolio represented by its current holdings, translate directly into the regulatory capital charge (see [6]). Motivated by this regulatory feature, we utilized the actual portfolio framework (APF) to determine a set of portfolios characterized by the optimal trade-off between the expected return and VaR (i.e. Pareto-optimal frontier). We further proposed a mean-univariate Generalized AutoRegressive Conditional Heteroscedasticity (GARCH) VaR portfolio optimization model. We assumed that portfolio returns, standardized by time varying volatility, have a conditional Student’s t distribution, while conditional variance follows a GARCH (1, 1) process. The Student’s t distribution efficiently captures the fat tails of standardized asset returns (see [9,10]) whilst the GARCH model addresses issues related to volatility clustering observed in the data. To the best of our knowledge, this is the first paper that studies mean-VaR portfolio optimization using the actual portfolio approach and also the first paper that uses the univariate GARCH VaR model in this context.

Previous studies on mean-VaR optimization implicitly assumed fixed weights (i.e. fractions of assets) over the observed time period. Since prices change over time, maintaining the fixed portfolio weights (FWA) requires frequent trading and leads to changes in the number of shares of each asset in a portfolio. Regulatory capital charges, however, are determined by the VaR of an actual portfolio where the number of shares of an asset (rather than its weighting) is fixed over the observed period. It is, therefore, APF, rather than FWA, that is more relevant for asset managers facing regulatory VaR limits. To illustrate the effectiveness of our APF approach and univariate GARCH model, we compared our results with two benchmarks. Our first benchmark is the mean-historical VaR approach developed in Rockafellar and Uryasev [20,21] and Krokhmal et al. [22]. These authors mapped conditional VaR (CVaR) optimization into a linear programming problem and argued that the mean-CVaR efficient frontier provided near-

---

[1] For classification and comparison of risk measures see [1,2], among others.
[2] GARCH was introduced in Engle [7] and GARCH (1, 1) in Bollerslev [8].
[3] For more on superiority of GARCH VaR compared to historical VaR see [11–15], among others.
[4] For example, [16–21].
optimal solutions in the context of the mean-historical VaR trade-off.\[^{5}\] We referred to this benchmark as the Linear Programming (LP) model. The second point of comparison for our approach is the mean-multivariate GARCH VaR optimization that can be mapped into the Quadratic Programming (QP) problem (see [26]).

Use of the univariate GARCH approach for VaR modeling, however, makes the mean-VaR optimization problem rather complex. Previous literature documented that Multi-Objective Evolutionary Algorithms (MOEA) could reliably and efficiently be applied in complex portfolio optimization problems.\[^{3}]](27) For example, Anagnostopoulos and Mamanis [31] studied the effectiveness of different MOEA (e.g. Nondominated Sorting Genetic Algorithm (NSGA-II), Strength Pareto Evolutionary Algorithm (SPEA), Pareto Envelope-based Selection Algorithm (PESA), etc.) in solving various complex mean-variance optimization problems. They reported the best NSGA-II and SPEA2 average performance in terms of hyper-volume indicators, while PESA performed best in terms of the proximity to the Pareto-optimal frontier. The same authors (see [32]) also examined the mean-variance, mean-ES, and mean-VaR optimization problem with quantity, cardinality and class constraints. They showed that NSGA-II, SPEA2 and PESA performed efficiently and their performance was independent of the risk measure used.\[^{7}\] Deb et al. [34] reported NSGA-II’s advantages over the Pareto Archived Evolutionary Strategy (PAES) and SPEA. Deb et al. [35] developed a hybrid NSGA-II procedure for handling a mean-variance portfolio optimization problem with the cardinality constraint and lower and upper bounds as investment criteria. The authors provided evidence of NSGA-II’s superiority over classical quadratic programming approaches. Branke et al. [36] considered the mean-variance problem with the maximum exposure constraint. The authors generated mean-variance Pareto-optimal frontiers by using a hybrid algorithm that combined NSGA-II with the critical line algorithm.

In this study we used the NSGA-II algorithm. Our choice was motivated by the above studies, whose results highlighted the advantages of NSGA-II in tackling complex portfolio optimization problems.\[^{8}\] NSGA-II was first introduced by Deb et al. [34] and subsequently developed in Deb et al. [37]. Recently, a new version of NSGA (NSGA-III) was developed by Deb and Jain [38] and also Jain and Deb [39]. NSGA-III is better suited to many-objective optimization problems (three and more objectives) and offers the ability to define the desired part of a solution space defined by reference points.\[^{7}\] Here, we prefer NSGA-II to NSGA-III, since we consider only two objectives and are interested in the entire Pareto-optimal frontier. Knowledge of the entire frontier of the risk-return trade-offs is particularly useful for asset managers in banks and other financial institutions in order to comply with different in-house and regulatory requirements.

For our numerical tests we selected 40 of the largest US stocks in the Standard and Poor’s stock market index (S&P 100) for which sufficient data was available. First, we determined the mean-historical VaR Pareto-optimal frontier, in the APF framework, using the NSGA-II algorithm and compared it to the LP Pareto optimal frontier. The comparison showed that NSGA-II produced better mean-historical VaR trade-offs compared to the LP optimization approach, in both Low and High volatility samples. Second, we compared our mean-univariate GARCH VaR Pareto-optimal frontier with the frontiers obtained by the two benchmarks (mean-historical VaR and mean-multivariate GARCH VaR). In comparison with the two benchmarks, the proposed univariate GARCH VaR procedure provided actual portfolios with better mean-univariate GARCH VaR trade-offs, in both Low and High volatility samples.

Previous portfolio optimization studies typically neglect the differences between using portfolios with fixed weights and portfolios with fixed holdings of assets. We contribute to portfolio optimization literature by addressing recent real world regulatory changes which impose VaR based on actual portfolio holdings. The rare MOEA portfolio optimization studies that measured risk by using VaR tend to use historical rather than analytical VaR (see [33, 41–43]).\[^{10}\] We therefore also contribute to MOEA literature by examining the mean-VaR optimization problem with analytical univariate GARCH VaR instead of historical VAR.

The remainder of this paper is organized as follows. In Section 2 we introduce historical and GARCH VaR models. The optimization problem is introduced in Section 3. MOEA and implementation details are discussed in Section 4. Section 5 contains sample descriptive statistics. Results of mean-historical VaR optimization are presented in Section 6. Results of our univariate and multi-variate mean-GARCH VaR optimization are presented and discussed in Section 7. We conclude in Section 8. All acronyms and notations are listed in Appendix.

2. VaR models

2.1. Traditional (historical) VaR

For a given portfolio, significance level $\alpha$ and time horizon $h$, portfolio VaR is a loss that will be exceeded, on average, only $\alpha \times 100$ percent of the time. If expressed in terms of portfolio value, $\text{VaR}$ is the $\alpha$-quantile of profit and loss distribution (cash VaR), while if expressed in terms of portfolio return $r$, it is the $\alpha$-quantile of the return distribution (relative VaR or, simply, VaR). We focused on the $\alpha$-quantile of the return distribution. Consider time horizon of $h = 1$ day and return $r_{\alpha}$ such that probability $p(r < r_{\alpha}) = \alpha$. In that case, 1-day VaR with significance level $\alpha$ is

$$\text{VaR}^\alpha = - r_{\alpha}. \tag{1}$$

The minus sign is needed since $\text{VaR}$ is defined as a loss. Given the cumulative distribution function of returns $F(\alpha)$, the $\alpha$-quantile is calculated as $r_{\alpha} = F^{-1}(\alpha)$. The main reason for the popularity of this method is ease of its implementation and the fact that it makes no assumptions about the parametric form of return distributions. On the other hand, the historical VaR often slowly reacts to abrupt changes in market conditions (see [11]).

2.2. Analytical VaR

A popular alternative to historical simulation is provided by analytical (or parametric) VaR models. These models take a stand on the shape of the return distribution and capture the following stylized facts about asset returns (see [5]). First, asset returns are difficult to predict based on their past realizations. Second, volatility of daily

[^5]: CVaR (or Expected Shortfall, ES) is the expected loss, conditional that loss is higher than VaR. Thus, CVaR and VaR are closely related risk measures. The LP model is widely used in CVaR and VaR optimization literature (see [23–25], among others).

[^6]: For comprehensive surveys of MOEA applications in portfolio optimization, see [27–30].

[^7]: Anagnostopoulos and Mamanis also showed that NSGA-II, SPEA2, and PESA provide a good approximation of risk-return trade-offs in a 3 objectives optimization problem (see [33]).

[^8]: NSGA-II algorithm is also the most widely used MOEA in portfolio optimization literature. According to the number of studies reviewed in [30], NSGA-II was used in twice as many studies compared to second most popular method (SPEA2).

[^9]: The number of the Pareto optimal solutions could also be reduced by combination of NSGA-II with Data Envelopment Analysis (see [40]).

[^10]: Only 4.27% of MOEA studies use VaR as one of the objectives in the portfolio optimization problem (see [27]).
returns dominates their mean. Third, unlike returns, the variance of daily volatile returns tends to exhibit clustering over time. Namely, days with highly volatile returns are likely followed by days with highly volatile returns and vice versa. Fourth, after periods of high or low asset volatility, volatility tends to return toward a stable long run level. Fifth, even though asset returns are often modeled using a normal distribution, an unconditional return distribution usually has much heavier tails than predicted by a normal distribution. Using the first two conditional moments of distribution, the return can be presented in the form:

$$r_t = \mu_t + \sigma_t z_t$$  \hspace{1cm} (2)

where \(r_t\) is the return at time \(t\), \(\mu_t\) is the conditional mean, \(\sigma_t\) is conditional volatility of the return, and \(z_t\) is the residual (innovation term) of the process. It is assumed that \(z_t\) are independently and identically distributed and follow a known theoretical distribution \(D\) with zero mean and unit variance \(D(0,1)\).

To specify the analytical VaR model we specified the volatility updating model, as well as the shape of the conditional return distribution \(D(0,1)\). A commonly applied class of conditional volatility models that captures all of the above stylized facts of returns is the GARCH model. In estimating GARCH parameters on daily data, we have taken into account that the conditional mean is dominated by the standard deviation of returns (see [9,15,44]). This implies:

$$Vt = \sigma_t z_t$$  \hspace{1cm} (3)

We took into account non-normality of standardized returns by assuming they follow a standardized Student’s \(t\) distribution (with zero mean and unit variance) with \(d\) degrees of freedom \(t(d)\). The degree of freedom parameter is an additional parameter estimated jointly with the other model parameters using Maximum Likelihood Estimation (MLE) method. The 1-day ahead analytical VaR estimate with significance level \(\alpha\), calculated at time \(t\), is obtained as follows:

$$VaR^t_w = -\sigma_t + \alpha^{-1}(d)$$  \hspace{1cm} (4)

where \(\sigma_{t+1}\) is the 1-day ahead conditional volatility estimate at time \(t\) obtained by applying the proposed model, \(d\) are degrees of freedom of the estimated Student’s \(t\) distribution of standardized portfolio returns, and \(\alpha^{-1}(d)\) is \(\alpha\)-quantile of the standardized Student’s \(t\) distribution with \(d\) degrees of freedom.

Conditional portfolio volatility can be modeled directly using the time series of portfolio returns (we referred to this method as the Univariate GARCH approach). Alternatively, one can estimate conditional portfolio volatility using the conditional variance–covariance matrix estimated via the multivariate GARCH model. We referred to this method as the multivariate GARCH approach.

**2.2.2. Multivariate GARCH VaR**

In a multivariate GARCH modeling framework, the 1-day ahead portfolio conditional volatility can be estimated as a function of the returns on individual assets (from the opportunity set) as follows:

$$\sigma_{t+1}^2 = w_H w_t$$  \hspace{1cm} (7)

where \(H_{t+1}\) is \(N \times N\) conditional variance–covariance matrix of returns of individual assets, and \(w_t\) is the vector of portfolio weights at time \(t\). Importantly, the multivariate GARCH implicitly assumes a fixed portfolio weights approach (FWA). The portfolio weights are thus fixed at \(w_{t+1} = w_t\) for all, \(t \leq T\), where \(T\) denotes optimization date. Thus, the multivariate GARCH VaR is

$$VaR^T_w = -\left(w'H_t w\right)^{1/2} \alpha^{-1}(d)$$  \hspace{1cm} (8)

3. Optimization problem

3.1. Mean-univariate GARCH VaR optimization (APF approach)

The problem we are trying to solve has the following general form:

$$\min_{w} VaR^T_w = -\left(\omega + \theta r_t^2 + \beta \sigma_t^2\right)^{1/2} \alpha^{-1}(d)$$  \hspace{1cm} (9)

$$\max E(r)$$  \hspace{1cm} (10)

subject to \(\sum_{i=1}^{N} w_i = 1\)

$$0 \leq w_i \leq 1, \quad i = 1, ..., N$$  \hspace{1cm} (12)

Here, \(w\) denotes the vector of portfolio weights at optimization date \(t=T\), its components are \(w_i\), \(r\) is the return on the portfolio, \(VaR^T_w\) denotes the portfolio risk measure that we try to minimize, and \(E(r)\) is the expected return on the portfolio. \(\alpha^{-1}(d)\) here refers to a quantile of a univariate Student’s \(t\) distribution with \(d\) degrees of freedom. Eq. (11) describes the standard budget constraint which requires that weights sum up to 1. Eq. (12) states that no short sales are allowed.

To calculate actual portfolio VaR and the corresponding expected return, a time series of portfolio returns with fixed asset holdings is needed. We determined asset holdings at the optimization day by adopting an arbitrary dollar portfolio value (set, without loss of generality, to 1). Thus, we assumed that, at optimization date \(t=T\), dollar portfolio value \(V_{p, t=T}\) was 1. Portfolio holdings \(n_i\) at time \(T\) were determined based on weights at \(t=T\):

$$n_i = \frac{w_i V_{p, t=T}}{P_{i, t=T}} = \frac{w_i}{P_{i, t=T}}$$  \hspace{1cm} (13)

Here \(P_{i, t}\) denotes the price of shares in company \(i\) at time \(t\). In the actual portfolio approach, holdings are held fixed over time. The actual portfolio return at time \(t \leq T\) is, therefore, determined

---

11. Although this is partially rectified when returns are standardized by time-varying volatility, some residual non-normality still remains.
12. GARCH process was first applied to VaR modeling in [44].
13. Alternatively, one can apply the same model to mean-adjusted returns.
14. The second condition assures stationarity of the conditional volatility process.
15. See, for example, [26,45].
16. Boldface denotes matrices and vectors. Matrices are in upper case, whereas vectors are in lower case.
17. As discussed in [46], the fixed weights assumption further implies continuous portfolio rebalancing.
18. Short sale prohibition is a common constraint imposed on large institutional investors such as mutual or pension funds.
as follows:

\[ r_t = \frac{V_{p,t}}{V_{p,t-1}} - 1 = \frac{\sum_{i=1}^{N} n_i P_{i,t}}{\sum_{i=1}^{N} n_i P_{i,t-1}} - 1 \quad (14) \]

\( V_{p,t} \) denotes the dollar portfolio value at time \( t \). The actual portfolio mean-\( \text{VaR} \) optimization problem simply means that we used Eq. (14) when determining input returns for the above optimization problem.

In the actual portfolio approach, portfolio weights at time \( t \leq T \), are given by the expression:

\[ w_{i,t} = \frac{n_i P_{i,t}}{\sum_{i=1}^{N} n_i P_{i,t}} \quad (15) \]

Eqs. (14) and (15) imply that

\[ r_t = \sum_{i=1}^{N} w_{i,t} r_{i,t} \quad (16) \]

where \( r_{i,t} \) is simple return on asset \( i \) at time \( t \).

3.2. Mean-multivariate GARCH \( \text{VaR} \) optimization (FWA approach)

The portfolio optimization problem (Eqs. (9–12)) can be specified as follows:

\[ \min_{\mathbf{w}} \text{VaR}_{t+1}^D = - (\mathbf{w}^T \mathbf{H}_{t+1} \mathbf{w})^{1/2} t_{a}^{-1}(d) \quad (17) \]

subject to \( \mathbf{w}^T \mathbf{\mu} = \tau \quad (18) \)

subject to \( \sum_{i=1}^{N} w_i = 1 \quad (19) \)

\( 0 \leq w_i \leq 1, \quad i = 1, ..., N \quad (20) \)

where \( \mathbf{H}_{t+1} \) denotes the conditional variance–covariance matrix of individual returns estimated at the time of optimization \( t = T \), \( \mathbf{\mu} \) denotes vector of expected returns of individual assets, and \( \tau \) denotes the expected portfolio return as a weighted average of constituents’ expected returns.

A portfolio return for an arbitrary, \( t \leq T \), is given by the expression:

\[ r_t = \sum_{i=1}^{N} w_{i,t} r_{i,t} \quad (21) \]

Quantile \( t_{a}^{-1}(d) \) refers to a quantile of a multivariate Student’s \( t \) distribution with \( d \) degrees of freedom. Here, \( d \) does not depend on a portfolio composition. Namely, \( d \) and, thus, quantile \( t_{a}^{-1}(d) \) depend solely on the choice of constituent return series and not on the way in which they are combined into a particular portfolio. \( \mathbf{H}_{t+1} \) and \( d \) are obtained using the Dynamic Conditional Correlation (DCC) model of Engle [47].

DCC decomposes conditional variance–covariance matrix \( \mathbf{H}_{t+1} \) into conditional standard deviations and correlations, that is:

\[ \mathbf{H}_{t+1} = \mathbf{D}_{t+1} \mathbf{R}_{t+1} \mathbf{D}_{t+1} \quad (22) \]

where \( \mathbf{D}_{t+1} = \text{diag}(\sigma_{1,t+1}, ..., \sigma_{N,t+1}) \). Here diag () is the operator transforming a \( N \times 1 \) vector into a \( N \times N \) diagonal matrix. We assume that conditional variances \( \sigma_{i,t+1}^2, i = 1, ..., N \), follow a standard univariate GARCH (1,1) model.

Matrix \( \mathbf{R}_{t+1} \) is a symmetric positive definite conditional correlation matrix defined as

\[ \mathbf{R}_{t+1} = \text{diag}(\mathbf{Q}_{t+1})^{-1/2} \mathbf{Q}_{t+1} \text{diag}(\mathbf{Q}_{t+1})^{-1/2} \quad (23) \]

where \( \mathbf{Q}_{t+1} \) is a proxy process which is assumed to follow GARCH-type dynamics:

\[ \mathbf{Q}_{t+1} = (1 - \theta - \beta) \mathbf{Q} + \theta \mathbf{z}_t \mathbf{z}_t^T + \beta \mathbf{Q} \quad (24) \]

Vector \( \mathbf{z}_t = (z_{1,t}, ..., z_{N,t})^T \) has elements \( z_{i,t} = r_{i,t}\sigma_{i,t} \) (standardized unexpected returns or innovations), \( \mathbf{Q} \) is the \( N \times N \) covariance matrix of \( \mathbf{z}_t \) and \( \theta \) and \( \beta \) are non-negative scalar parameters satisfying \( \theta + \beta < 1 \).

Since \( t_{a}^{-1}(d) \) is constant, the optimization problem (Eq. (17)) is equivalent to a problem represented by quadratic programming formulation (referred to as \( QP \) model):

\[ \min_{\mathbf{w}} \mathbf{w}^T \mathbf{H}_{t+1} \mathbf{w} \quad (25) \]

It should be emphasized that quantile \( t_{a}^{-1}(d) \) in this approach is constant and therefore not portfolio specific. Thus, the multivariate GARCH approach is more restrictive compared to the univariate GARCH approach where the degrees of freedom of the estimated standardized Student’s \( t \) distribution (and therefore quantile \( t_{a}^{-1}(d) \)) were portfolio specific.

3.3. Summary of assumptions

3.3.1. General

i) Short sales are not allowed (Eqs. (12) and (21));

ii) The standard budget constraint which requires that weights must sum up to 1 (Eqs. (11) and (20)).

3.3.2. Mean-univariate GARCH \( \text{VaR} \) optimization

i) Portfolio returns, standardized by time varying volatility, follow a conditional standardized Student’s \( t \) distribution (with zero mean and unit variance);

ii) Conditional variance follows GARCH (1, 1) process, \( \sigma_i^2 = \omega + \theta \sigma_{i-1}^2 + \beta \sigma_{i-1}^2 \) (Eq. (5));

iii) In estimating GARCH parameters on daily data, the mean value of daily returns is dominated by the standard deviation of returns and that \( r_t \approx \sigma_z \) (Eq. (3));

iv) At optimization date \( t = T \), without loss of generality, dollar portfolio value \( V_{p,T} = 1 \).

3.3.3. Mean-multivariate GARCH \( \text{VaR} \) optimization

i) Portfolio weights are fixed over the observed period;

ii) Returns of assets from the opportunity set jointly follow a multivariate Student’s \( t \) distribution;

iii) \( \mathbf{Q}_{t+1} \) (Eq. (23)) follows GARCH-type dynamics: \( \mathbf{Q}_{t+1} = (1 - \theta - \beta) \mathbf{Q} + \theta \mathbf{z}_t \mathbf{z}_t^T + \beta \mathbf{Q} \);

iv) In matrix \( \mathbf{D} \) \( t+1 \) conditional variance \( \sigma_i^2 \) follows a univariate GARCH (1, 1) process.

4. Portfolio optimization methodology

4.1. Evolutionary algorithms (EA)

EA are efficient stochastic search techniques for solving complex optimization and search problems (e.g., optimization

---

19 We used ‘rmgarach’ package [48] within software R [49]. For fitting GARCH model we used ‘dccfit’ method. For the 1-day ahead estimation of conditional covariance matrix we used ‘dccforecast’ method.

20 This problem is of the standard Markowitz type (see [50]) with added short sales constraints.
problems with non-differentiable objective functions, large and non-convex solution spaces, complex constraints, etc.) through an emulation of natural selection i.e. survival of the fittest (see [51]).

EA begin with a set of randomly generated candidate solutions, referred to as a population. In each of the iterations (i.e. generations) the following processes are performed: (i) Good solutions from the current population are selected and transferred into a set of potential parents (mating pool); (ii) Randomly selected parent solutions are combined (crossover) producing new solutions (offspring solutions); (iii) Some randomly selected offspring solutions are slightly modified (mutation); (iv) Generated offspring solutions constitute the population of the next generation. By performing these processes in each generation (until the termination condition is satisfied) the solutions “evolve” and become even better in fulfilling the stated objectives.

Over the past two decades, significant advances have been made on EA methods to solve multi-objective optimization problems.\(^{21}\) This has led to the creation of multi-objective evolutionary algorithms (MOEA). In general, MOEA address two important issues: (i) How to generate a Pareto-optimal frontier; and (ii) How to provide diversity in alternative solutions (i.e. how to avoid convergence to a single point on the Pareto-optimal frontier). To achieve the first goal, most MOEA implementations use ranking based on the concept of dominance, while different diversity-preserving techniques are employed to achieve the second goal.

4.2. MOEA implementation

We directly borrowed the NSGA-II algorithm from Deb et al. [37].\(^{22}\) Implementation of NSGA-II involves adopting settings for the solution representation, the population size, the crossover and mutation probabilities and the termination condition. Solution representation depends upon the specific optimization problem. For the model at hand, the solution was defined as a non-negative real-valued vector of portfolio weights at time \(t=T\), that is, the vector of fractions of the total budget invested in financial securities (see Section 4). Population size (the number of candidate portfolios in each generation), was set to 100. The Pareto-optimal frontier was, therefore, approximated with 100 points. The step by step flow chart which presents a schematic view of the proposed portfolio optimization method is shown in Fig. 1.

In the preparatory phase our software generated and printed a data file (data.csv) with a time series of occurrences of assets under consideration. Next, it generated and printed an R script file (Script.R) with commands for estimation of the univariate GARCH VaR. It then executed the NSGA-II algorithm in the execution phase.\(^{23}\)

For the breeding of an offspring population we used the uniform crossover operator. Two solutions (portfolios) from the current population were randomly selected and were recombined with a predefined crossover probability. If the two solutions undergo recombination, every allele (individual asset weight) is exchanged between the pair of randomly selected solutions with a certain probability, known as the swapping probability, otherwise, the two offspring are simply copies of their parents (see [51]).\(^{24}\) In accordance with previous literature, we set the swapping probability to be 0.5.\(^{25}\)

\(^{21}\) For a detailed introduction to MOEA, see [52].
\(^{22}\) The algorithm was coded in C and run on a personal computer with Intel i5 processor and 4 GB of RAM.
\(^{23}\) For more details on the algorithm steps, see [37].
\(^{24}\) Candidate solutions are also referred to as chromosomes, decision variables are referred to as genes and the values of decision variables are called alleles.
\(^{25}\) See [31–33,51].

We applied a uniform mutation operator (uniform replacement). When applying uniform mutation, each allele is selected with a predefined mutation probability and replaced with a realization of a random variable, uniformly distributed in the range defined by the lower and upper domain bounds. The selected crossover and mutation operator ensured that the constraint defined by Eq. (12) was satisfied for each offspring. However, these operators did not ensure satisfaction of the budget constraint (Eq. (11)). Hence, we had to normalize each of the offspring solutions.\(^{26}\)

In order to find the appropriate parameter values for the crossover and mutation probabilities, we performed a series of experiments. The performance was assessed using the \(\varepsilon\)-indicator, and the hypervolume metric (see [53]). The \(\varepsilon\)-indicator is a binary performance metric which is used to measure how close the approximation set is to the reference set. As a reference set, the true or the best known Pareto-optimal frontier was used. The \(\varepsilon\)-indicator determines a minimum value the reference set must be multiplied by in order for every solution in the reference set to become weakly dominated by at least one solution in the approximation set. If the approximation set matches the reference set exactly then the \(\varepsilon\)-indicator takes the value of one. For this metric, values close to one indicate that the approximation set very closely matches the reference set. The hypervolume indicator was used to measure the diversity of the approximation set. The hypervolume measure quantifies the volume of the objective space dominated by an approximation set. For optimization problems with two objectives, it quantifies the area of the objective space dominated by the approximation set, bounded by a predefined reference point. Thus, for this metric higher values are preferable.

To determine the best performing mutation and crossover probability sets, we tested four crossover probabilities (0.7, 0.8, 0.9, 1.0) and five uniform mutation probabilities (0.001, 0.005, 0.01, 0.05, 0.1). The above tests resulted in 20 different configurations for each set of objectives (mean-historical VaR, mean-GARCH VaR) and for each data sample that we used. For each configuration, the algorithm was left to run until 100,000 solutions were generated.

\(^{26}\) We have done this by dividing each weight by sum of all weights.
The interactions between the NSGA-II algorithm and statistical computing.

Fig. 2. Interactions between the NSGA-II algorithm and statistical computing.

5. Sample descriptive statistics

For the purposes of this study, we selected 40 constituents of the S&P 100 with the highest market capitalization as of September 6, 2013. We observed the constituents’ time series of 1421 returns in the period January 15, 2008–September 6, 2013. We used a rolling window of 1000 returns for VaR estimation. For each day within the rolling estimation period of 421 days (January 4, 2012–September 6, 2013) we estimated 1-day ahead daily volatility of the S&P 100 time series. The volatilities were estimated using the univariate GARCH (1,1) model. Standardized returns were assumed to have a standardized t-distribution.

GARCH volatility estimations of the S&P 100 index were based on a rolling window of 1000 daily returns. Maximum volatility of 0.01294 (20.5% in annual terms) was determined on June 29, 2012. Minimum volatility of 0.00583 (9.3% in annual terms) was determined on July 31, 2013. From the chosen sample of 40 stocks we created two samples with 1001 daily prices: (i) First, ending on the maximum volatility date (referred to as the High volatility sample); and (ii) Second, ending on the minimum volatility date (referred to as the Low volatility sample). All VaR values in this paper correspond to a significance level of 0.01, with a time horizon of one day. Table 1 presents a summary statistics of the time series comprising the Low and High volatility samples.

6. Mean-historical VaR optimization

Rockafellar and Uryasev [20,21] and Krokhmal et al. [22] argue that the mean-CVaR efficient frontier generated by the LP model also provides near-optimal solutions in the context of mean-historical VaR optimization. We tested this assertion in an actual portfolio framework. Namely, for each portfolio belonging to the mean-CVaR Pareto-optimal frontier generated by a LP (FWA approach), we created a time series of actual portfolio returns (applying Eqs. (13) and (14)), and calculated the corresponding 1% historical VaR and expected return, and represented the portfolios in the actual portfolio mean-VaR plane. In this way we generated an approximate Pareto-optimal frontier corresponding to the LP solution. Next, we determined the mean-historical VaR Pareto-optimal frontier (in the actual portfolio framework directly) using the NSGA-II optimization approach (see Fig. 3). Fig. 3 shows that the NSGA-II frontiers dominate the LP frontiers, especially in the High volatility sample.

Best performing NSGA-II mean-historical VaR approximate of the Pareto-optimal set was selected based on the hypervolume metric. To compare the NSGA-II mean-historical VaR Pareto-optimal frontier to the LP Pareto-optimal set, we used the ε-indicator. As a reference set, we adopted the NSGA-II mean-historical VaR Pareto-optimal frontier. The comparison is presented in Table 2. The values of the hypervolume parameter and reference points, together with the ε-indicator, were for the best performing NSGA-II approximates of the mean-VaR Pareto-optimal frontier.

The values for the ε-indicator confirm that the NSGA-II optimization approach provides better mean-historical VaR trade-offs of actual portfolios compared to the LP solutions in both the High and Low volatility samples. The relatively higher value of the ε-indicator (1.1526) suggests that NSGA-II performed particularly well during the High volatility period.

Table 1

<table>
<thead>
<tr>
<th>Sample</th>
<th>Average expected return (p.a.)</th>
<th>Average volatility (p.a.)</th>
<th>Average historical VaR</th>
<th>Average CVaR VaR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low volatility</td>
<td>16.61</td>
<td>23.47</td>
<td>3.96</td>
<td>3.12</td>
</tr>
<tr>
<td>High volatility</td>
<td>13.47</td>
<td>37.94</td>
<td>6.95</td>
<td>4.12</td>
</tr>
</tbody>
</table>

GARCH VaR volatility estimations of the S&P 100 index were based on a rolling window of 1000 daily returns. Maximum volatility of 0.01294 (20.5% in annual terms) was determined on June 29, 2012. Minimum volatility of 0.00583 (9.3% in annual terms) was determined on July 31, 2013. From the chosen sample of 40 stocks we created two samples with 1001 daily prices: (i) First, ending on the maximum volatility date (referred to as the High volatility sample); and (ii) Second, ending on the minimum volatility date (referred to as the Low volatility sample). All VaR values in this paper correspond to a significance level of 0.01, with a time horizon of one day. Table 1 presents a summary statistics of the time series comprising the Low and High volatility samples.

27 We used ‘rugarch’ package [54] within software R [49]. For fitting GARCH model we used ‘ugarchfit’ method; For the 1-day ahead estimation of conditional volatility we used ‘ugarchforecast’ method.

28 We obtained these solutions and included them into the initial population of each MOEA execution. Maximum return solution was obtained analytically, while minimum risk solution was obtained using single-objective genetic algorithm.

29 Initially, 43 stocks were considered of which 3 were discarded (tickers: PM, V, KFT) due to incomplete data during the sample period.

30 We annualize volatility using square root of time assuming 252 trading days per annum.
7. Results for mean-GARCH VaR optimization

7.1. Pareto-optimal frontiers in the Low and High volatility samples

In this section we compared the Pareto-optimal frontiers for:

(i) the mean-univariate GARCH VaR (or univariate GARCH for short); (ii) the mean-multivariate GARCH VaR (benchmark 1); and (iii) the mean-historical VaR (benchmark 2). In order to obtain solutions in the actual portfolio framework for both benchmark portfolio solutions (represented by the corresponding vector of weights \( w \)), we generated a time series of portfolio returns. The time series of portfolio returns was generated by employing Eqs. (13) and (14); and (Eqs. (4) and (5)).

The results are shown in Fig. 4. The Pareto-optimal frontiers, for actual portfolios, are for the two volatility regimes when 1% VaR was estimated using the univariate GARCH model. Triangle markers represent the univariate GARCH frontier (obtained via NSGA-II), while filled dots represent benchmark 1 solutions and empty dots represent benchmark 2 solutions.

The differences are particularly prominent in the area of low returns (and low risk). In this segment of the Pareto-optimal set, thus, the opportunities for improving the mean-VaR trade-off through optimization are the greatest. It is worth noting that this segment of the Pareto-optimal set is associated with most diversified portfolios. In contrast, with higher expected return values, the cardinality of efficient portfolios is reduced (the highest return portfolio, by construction, corresponds to a single asset).

To further compare the univariate GARCH Pareto-optimal frontier to the respective benchmarks, we used the \( \varepsilon \)-indicator. As a reference set, we adopted the univariate GARCH frontier. Table 3 presents the values of the hypervolume parameter and reference points for the best performing NSGA-II mean-univariate GARCH VaR approximate Pareto-optimal sets. The table also shows \( \varepsilon \)-indicators corresponding to the respective benchmarks in the Low and High volatility samples.

The above presented \( \varepsilon \)-indicators confirm superiority of the univariate GARCH VaR optimization approach. The advantages of the univariate GARCH VaR optimization approach are particularly pronounced in the Low volatility sample. Measured by the \( \varepsilon \)-indicators, the mean-historical VaR approach provides the worst approximations out of the three approaches. The above result is particularly important given that approximately 75% of banks tend to use the historical VaR models for portfolio optimization (see [55]).

7.2. Process time (CPU time)

As expected, our CPU time is longer compared to the CPU time of the traditional methods (e.g. QP). For example, the execution of one generation of NSGA-II, when solving the mean-univariate GARCH VaR problem, lasted 41 s. Hence, to execute 1000 generations of NSGA-II we needed 41,000 s. The total CPU time for mean-univariate GARCH VaR optimization with the QP solver was

---

31 As presented in Fig. 3, the NSGA-II method provided superior actual portfolio mean-historical VaR trade-offs compared to the LP model, in both samples. Consequently, we adopted the NSGA-II mean-historical VaR portfolios as benchmark 2.

32 During this process some of benchmark 1 and benchmark 2 portfolio solutions became dominated and thus have been discarded from the approximation set.

33 In comparison, it took only 1 s for the LP model to generate mean-historical VaR Pareto-optimal frontier consisting of 100 solutions.
Our additional analysis, however, revealed that the implied capital charge are calculated using 8. Conclusions

less pronounced in the Low volatility sample. As expected, the differences between the results

tion models used in our study (see

compared out-of-sample performance of the different optimiza-

tion revealed that 99% of final hypervolume was achieved after 50 generations (12 s) in the Low volatility sample, and after 52 generations (12.5 s) in the High volatility sample. The same analysis applied on the mean-univariate GARCH VaR optimization revealed that 99% of final hypervolume was achieved after 40 generations (1640 s) in the Low volatility sample. The corresponding values in the High volatility sample were 20 generations and 820 s. Given its longer CPU time, the use of NSGA-II could be justified by its greater flexibility and ability to deal with more complex real-life portfolio optimization problems compared to the QP method (see [35]).

7.3. Out-of-sample estimates

Bank managers and regulators are interested in the out-of-sample performance of different optimization models. Thus, we compared out-of-sample performance of the different optimization models used in our study (see Fig. 5). The optimization dates for our High and Low volatility samples were dates with the highest (29 June 2012) and lowest (31 July 2013) volatilities. Based on the estimated portfolios on the optimization dates, we calculated respective mean and VaR values for dates which fall exactly 1 month later (30 July 2012 and 3 September 2013 respectively).

The results of the out-of-sample analysis were consistent with the results reported in Fig. 4, thus confirming the superiority of estimates based on the mean-univariate GARCH model in both volatility samples. As expected, the differences between the results for the mean-univariate and mean-multivariate GARCH models are less pronounced in the Low volatility sample.

8. Conclusions

According to Basel regulation, the riskiness of a portfolio and the implied capital charge are calculated using VaR estimations of an 424 s. Our additional analysis, however, revealed that the NSGA-II CPU time could be shorter since the execution of 1000 generations was not always necessary. For example, for the mean-historical VaR optimization, 99% of final hypervolume was achieved after 50 generations (12 s) in the Low volatility sample, and after 52 generations (12.5 s) in the High volatility sample. The same analysis applied on the mean-univariate GARCH VaR optimization revealed that 99% of final hypervolume was achieved after 40 generations (1640 s) in the Low volatility sample. The corresponding values in the High volatility sample were 20 generations and 820 s. Given its longer CPU time, the use of NSGA-II could be justified by its greater flexibility and ability to deal with more complex real-life portfolio optimization problems compared to the QP method (see [35]).

Table 2
Comparison of NSGA-II mean-historical VaR and LP models.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Mutation probability</th>
<th>Crossover rate</th>
<th>Hypervolume</th>
<th>Reference point</th>
<th>$\epsilon$-indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low volatility</td>
<td>0.05</td>
<td>0.8</td>
<td>4.2366E–05</td>
<td>(0.05210)</td>
<td>1.1300</td>
</tr>
<tr>
<td>High volatility</td>
<td>0.05</td>
<td>1</td>
<td>6.9160E–05</td>
<td>(0.08230)</td>
<td>1.1526</td>
</tr>
</tbody>
</table>

Table 3
Parameters of the best performing mean-univariate GARCH VaR.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Mutation probability</th>
<th>Crossover rate</th>
<th>Hypervolume</th>
<th>Reference point</th>
<th>$\epsilon$-indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low volatility</td>
<td>0.05</td>
<td>0.7</td>
<td>5.0973E–05</td>
<td>(0.0504, 0)</td>
<td>1.2608</td>
</tr>
<tr>
<td>High volatility</td>
<td>0.01</td>
<td>0.9</td>
<td>5.3451E–05</td>
<td>(0.0581, 0)</td>
<td>1.0866</td>
</tr>
</tbody>
</table>

Fig. 5. (a) Out-of-sample test in the High volatility sample. (b) Out-of-sample test in the Low volatility sample.
set consisting of 40 large US stocks belonging to the S&P 100 index in two volatility regimes. We found that the NSGA-II method resulted in better actual portfolio mean-historical VaR trade-offs in both Low and High volatility regimes.

Next, we compared the mean-univariate GARCH VaR Pareto-optimal frontier to the mean–multivariate GARCH VaR and to the mean–historical VaR Pareto-optimal frontiers. In comparison to the two benchmarks, the proposed univariate GARCH VaR procedure, again, provided actual portfolios with a superior mean–univariate VaR frontier for both volatility samples. The results suggest that the multivariate GARCH modeling framework lacks flexibility to conform to the actual portfolio framework which is inherent in regulation. At the same time, the mean–historical VaR frontier provided the worst mean–univariate GARCH VaR trade-offs. Overall, our results bear two important implications for financial institutions and their regulators. First, the results highlighted differences between the actual portfolio approach and the approach based on the fixed weights. Second, the results show the importance of carefully selecting amongst different VaR methodologies used in portfolio optimization.

Recently, some multiple criteria decision making (MCDM) tools were combined with NSGA-III to either rank or reduce the Pareto optimal frontier (see [56]). Further enhancement of our optimization approach with some of the MCDM reduction techniques could be an interesting area for future research.

Acknowledgments

We would like to thank two anonymous referees for their helpful suggestions. Research presented in this paper was supported by Serbian Ministry of Education, Science and Technology, Grants OH 179005 and III-44010.

Appendix A

See Tables A1 and A2.

References


Table A1

<table>
<thead>
<tr>
<th>Acronyms</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>APF</td>
<td>actual portfolio framework</td>
</tr>
<tr>
<td>CPU Time</td>
<td>process time</td>
</tr>
<tr>
<td>CR</td>
<td>capital requirements</td>
</tr>
<tr>
<td>CVaR</td>
<td>conditional value-at-risk</td>
</tr>
<tr>
<td>DCC</td>
<td>dynamic conditional correlation model</td>
</tr>
<tr>
<td>EA</td>
<td>evolutionary algorithm</td>
</tr>
<tr>
<td>ES</td>
<td>expected shortfall</td>
</tr>
<tr>
<td>FWA</td>
<td>fixed weights approach</td>
</tr>
<tr>
<td>GARCH</td>
<td>generalized autoregressive conditional heteroscedasticity</td>
</tr>
<tr>
<td>LP</td>
<td>linear programming</td>
</tr>
<tr>
<td>MCDM</td>
<td>multiple criteria decision making</td>
</tr>
<tr>
<td>MLE</td>
<td>maximum likelihood estimation</td>
</tr>
<tr>
<td>MOEA</td>
<td>multi-objective evolutionary algorithm</td>
</tr>
<tr>
<td>NSGA-II (NSGA-III)</td>
<td>nondominated sorting genetic algorithm</td>
</tr>
<tr>
<td>PESA</td>
<td>Pareto envelope-based selection algorithm</td>
</tr>
<tr>
<td>PAES</td>
<td>Pareto archived evolutionary strategy</td>
</tr>
<tr>
<td>QP</td>
<td>quadratic programming</td>
</tr>
<tr>
<td>S&amp;P 100</td>
<td>Standard and Poor’s 100 stock market index</td>
</tr>
<tr>
<td>SPEA</td>
<td>strength Pareto evolutionary algorithm</td>
</tr>
<tr>
<td>VaR</td>
<td>Value-at-Risk</td>
</tr>
</tbody>
</table>

Table A2

<table>
<thead>
<tr>
<th>Notations</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$r_i$</td>
<td>portfolio return</td>
</tr>
<tr>
<td>$\mu_r$</td>
<td>conditional expected return</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>conditional variance</td>
</tr>
<tr>
<td>$\Sigma$</td>
<td>residual (innovation) term</td>
</tr>
<tr>
<td>$p$</td>
<td>portfolio probability</td>
</tr>
<tr>
<td>$\rho$</td>
<td>VaR</td>
</tr>
<tr>
<td>$V_{j, t}$</td>
<td>dollar portfolio value</td>
</tr>
<tr>
<td>$\sigma_{t, t}$</td>
<td>1-day ahead conditional volatility</td>
</tr>
<tr>
<td>$\tilde{f}(\alpha)$</td>
<td>cumulative distribution function of returns</td>
</tr>
<tr>
<td>$a$</td>
<td>significance level</td>
</tr>
<tr>
<td>$h$</td>
<td>time horizon</td>
</tr>
<tr>
<td>$d$</td>
<td>degrees of freedom</td>
</tr>
<tr>
<td>$H_{i, t}$</td>
<td>1-day ahead conditional variance–covariance matrix of returns</td>
</tr>
<tr>
<td>$\mathbf{w}_i$</td>
<td>vector of portfolio weights</td>
</tr>
<tr>
<td>$\text{VaR}_{i, t}$</td>
<td>GARCH VaR of a portfolio</td>
</tr>
<tr>
<td>$R(t)$</td>
<td>portfolio expected return</td>
</tr>
<tr>
<td>$P_{i, t}$</td>
<td>company $i$ share price</td>
</tr>
<tr>
<td>$n_i$</td>
<td>portfolio holdings</td>
</tr>
</tbody>
</table>


