The impact of covariance misspecification in risk-based portfolios

Essay

Jean-Philippe Gagnon Fleury

Master degree in financial engineering
Master of Science (M.Sc.)

Québec, Canada

Contents

1 Introduction

Foundations of Modern Portfolio Theory are based on the work by Markowitz [1952]. Given a mean-variance optimization (MVO) problem, an investor can obtain an optimal portfolio for a universe of investable assets. The portfolio's expected return is maximized under a risk constraint. The set of portfolios obtained by letting varying the risk constraint is called the efficient frontier.

Although very appealing, the MVO suffers from many drawbacks. MVO's inputs are statistical estimates and hence estimated with error. Mean-variance optimization amplifies the effects of these errors which leads to extreme allocations. Moreover, MVO is very sensitive to changes in the inputs. This is especially true for expected returns, where small changes tend to produce a completely different portfolio allocation (Michaud [1989]). The instability is challenging for portfolio managers who face transaction costs.

Given the lack of diversification and instability of the MVO, one alternative is to ignore all available information and invest equally in each asset. However, this portfolio is expensive to manage because it needs to be rebalanced constantly. Another alternative is to invest in risk-based portfolios. Risk-based investments gained in popularity among practitioners and many academic papers were written on the subject. Those portfolios are heuristic-based and focus on diversification. They do not use expected performance measure as input. They consist in "allocating" in risk instead of capital.

Empirically, risk-based portfolios achieve higher Sharpe ratios than the MVO and are more robust to inputs parameters, especially in turmoil periods. Also, dynamic risk-based strategies generate less turnover (Bruder and Roncalli [2012]). In addition, risk-based portfolios tend to be more stable compared to mean-variance portfolio when mean and volatility are varying (Roncalli [2013]). From a practical perspective, they are therefore very attractive strategies.

Nevertheless, performance of risk-based strategies still depends on the accuracy of the input parameter since the covariance matrix is subject to estimation risk. The choice of the covariance model is therefore important. Zakamulin [2015] studies the robustness of the minimum-volatility with respect to several covariance matrix estimators. Results show that practitioners would gain significant benefits from adopting a multivariate GARCH model instead of the standard sample covariance in a minimum-volatility portfolio context. Also, the use of multivariate GARCH improves the tracking error and results in a better performance in terms of risk optimization.

The aim of our research is to study the impact of covariance misspecification for a wide set of risk-based portfolios. This set is composed of the equal-risk-contribution portfolio (Maillard, Roncalli, and Teïletche [2010]), the maximum-diversification portfolio (Choueifaty and Coignard [2008]), the risk-efficient portfolio (Amenc, Goltz, Martellini, and Retkowsky [2011]) and the minimum-variance portfolio. Because of its simplicity, we also included the inverse-volatility portfolio.

For the covariance models, the sample covariance is considered. However, the sample covariance matrix becomes inefficient as the number of assets increases compared to the number of time-series observations. For this reason, we also consider the shrinkage estimator of Ledoit and Wolf [2003]. That estimator displays a larger bias but is very efficient for short time-series. The gain in efficiency compensates for the increase in bias so that the overall error is reduced. In short, it consists of a weighted average of the sample covariance matrix and the covariance matrix given by the Sharpe's single factor model.

The two previous covariance models are only valid under the assumption that the covariance matrix is constant over time. However, it is well known that the second moments of stock returns present dynamic behaviors. Hence, the Exponential-Weighted-Smoothing-Average (EWMA) covariance matrix, popularized by RiskMetrics [1996] and widely used among practitioners, is also tested. The last covariance model used in this study is developed by Engle [2002]. It is a more complex and sophisticated model called Dynamic-Conditional-Correlation (DCC).

In the first part of our study, a Monte Carlo experiment is performed to measure the robustness of risk-based portfolios to the choice the covariance estimator. We find that allocations given by the inverse-volatility portfolio and the equal-risk-contribution portfolio are less sensitive to the error in the covariance matrix estimation. On the other hand, misspecification of the covariance matrix has a huge impact on the minimum-volatility. For each portfolio, the dynamic-conditional-correlation model improves significantly the asset allocation. The exponential-weighted-moving-average has an overall performance not far behind the DCC which can be interesting for practitioners who want to keep things simple.

In the second part, we study the impact of the covariance on the portfolio turnover for risk-based investments. Hence, historical simulations are performed on returns from 30 stocks included in the Dow Jones universe for a period ranging from March 2008 to September 2014. A brief performance analysis is also provided. We find that using either the sample or the Ledoit-Wolf covariance matrix results in much more stable portfolios. Using the dynamic-conditional-correlation or the exponential-weighted-smoothing-average increases significantly the turnover. Superiority in returns resulting from a dynamic approach for the covariance estimators might not be enough to encounter the cost incurred by the high portfolio's turnover.

From a practical standpoint, this study gives a general idea of how an optimized risk-based portfolio reacts to the covariance input. It also presents the trade-off between performance and turnover and explicits some issues associated with risk-based approach.

The rest of this essay is organized as follows. Section 3 describes the risk-based portfolios. Section 4 describes the covariance matrices. Section 5 presents the Monte Carlo study. Section 6 presents the backtest. Section 7 concludes.

2 R package

The computer code developed to perform the various analyses is available in the R package `RiskPortfolios` which can be downloaded from https://github.com/ArdiaD/RiskPortfolios/. The functions in the package allow us to:

1. Estimate expected returns according to various methodologies;
2. Estimate covariance matrices according to various methodologies;
3. Estimate implied expected returns from the market;
4. Find optimal allocation of various risk-based portfolios.

In addition to our package, we also rely on the package `rmgarch`[1] created by Alexio Ghalanos for the Dynamic Conditional Correlation calculations. We slightly modified the code to perform the composite likelihood estimation.

[1] http://cran.r-project.org/web/packages/rmgarch/index.html

3 Risk-based portfolios

This section describes the alternatives to the market capitalization portfolio that are used in this study. Most of those alternative portfolios are heuristics but are mean-variance efficient under different assumptions. For this section, we use the notation $\mathbf{\Sigma}$ as the $(N \times N)$ covariance matrix of arithmetic returns, $\mathbf{w}$ as the $(N \times 1)$ vector of weights and $\boldsymbol{\iota}$ as a $(N \times 1)$ vector of ones.

3.1 Minimum-variance portfolio

The minimum-variance portfolio is the allocation that minimize the risk of the portfolio regardless of its expected return:

$$\mathbf{w}_{\text{min}} \equiv \underset{\mathbf{w} \in \mathcal{C}_{\text{FI}}}{\text{argmin}} \left\{ \mathbf{w}' \mathbf{\Sigma} \mathbf{w} \right\} , \tag{1}$$

where $\mathcal{C}_{\text{FI}} \equiv \{\mathbf{w} \in \mathbb{R}^+ | \mathbf{w}'\boldsymbol{\iota} = 1\}$ is a constraint where short sales are not allowed and where wealth is fully invested. A potential problem with the minimum-variance portfolio is that it puts too much emphasis on stocks with low volatility and is therefore very concentrated.

3.2 Maximum diversification portfolio

The maximum-diversification portfolio aims at maximizing the benefits from diversification and is an alternative to the minimum-variance portfolio. Let $\boldsymbol{\sigma} \equiv \sqrt{Diag\{\mathbf{\Sigma}\}}$. In order to obtain effective diversification, the diversification ratio:

$$DR(\mathbf{w}) \equiv \frac{\mathbf{w}'\boldsymbol{\sigma}}{\sqrt{\mathbf{w}'\mathbf{\Sigma}\mathbf{w}}} , \tag{2}$$

must be greater than one. Thus, weights are obtained from the following optimization problem:

$$\mathbf{w}_{\text{md}} \equiv \underset{\mathbf{w} \in \mathcal{C}_{\text{FI}}}{\text{argmax}} \left\{ \frac{\mathbf{w}'\boldsymbol{\sigma}}{\sqrt{\mathbf{w}'\mathbf{\Sigma}\mathbf{w}}} \right\} .$$

The maximum-diversification portfolio has a maximal Sharpe ratio if the returns are proportional to their volatility (Choueifaty and Coignard [2008]). Furthermore, this portfolio tend to be less concentrated than the mean-variance portfolio (Clarke, De Silva, and Thorley [2013]).

3.3 Inverse-volatility portfolio

The inverse-volatility portfolio is an easy-to-implement portfolio extensively used in practice. It can be seen as a risk-based approach. This portfolio is built such that asset weights are inversely proportional to their volatility ($\boldsymbol{\sigma}$). More precisely, the weights are computed as:

$$\mathbf{w}_{\text{iv}} \equiv \frac{\boldsymbol{\iota}./\boldsymbol{\sigma}}{\boldsymbol{\iota}'(\boldsymbol{\iota}./\boldsymbol{\sigma})} , \tag{3}$$

where $./$ is the element-by-element division. A weakness of this approach is that correlations between assets are not taken into account. Contrarily to the equal-risk-contribution portfolio which is discussed below, each asset in the portfolio does not contribute to the same risk in the portfolio. Indeed, portfolio volatility is not an additive function of the individual asset volatilities. If for each stock the Sharpe ratio is the same and if all pairwise correlations are equal, then the inverse-volatility portfolio is mean-variance efficient and has the highest Sharpe-ratio (Leote de Carvalho, Lu, and Moulin [2012]).

3.4 Equal-risk-contribution portfolio

It is possible to allocate portfolio's assets according to their risk. Here, for instance, the notion of risk could be defined as the volatility, the Value-at-Risk (VaR) or the Expected Shortfall (ES). In this study, we use the volatility is the risk measure. The equal-risk-contribution portfolio relies on the risk-contribution of asset i to the portfolio:

$$\mathcal{RC}_i \equiv \mathbf{w}_i \frac{\partial \mathcal{R}(\mathbf{w})}{\partial \mathbf{w}_i} \,,$$

where $\mathcal{R}(\mathbf{w})$ is the total portfolio risk. According to the Euler decomposition, the sum of the risk-contribution of each asset is equal to the total portfolio risk:

$$\mathcal{R}(\mathbf{w}) = \sum_{i=1}^{N} \mathbf{w}_i \frac{\partial \mathcal{R}(\mathbf{w})}{\partial \mathbf{w}_i} = \sum_{i=1}^{N} \mathcal{RC}_i \,.$$

The equal-risk-contribution portfolio consists of building a portfolio such as each asset has the same risk-contribution in the portfolio. Numerically, optimal weights are computed as:

$$\mathbf{w}_{\text{erc}} \equiv \underset{\mathbf{w} \in \mathcal{C}_{\text{FI}}}{\operatorname{argmin}} \left\{ \sum_{i=1}^{N} (\%\mathcal{RC}_i - \tfrac{1}{N})^2 \right\} \,,$$

where $\%\mathcal{RC}_i \equiv \frac{w_i [\mathbf{\Sigma w}]_i}{\mathbf{w}'\mathbf{\Sigma w}}$. As suggested by Roncalli [2013], the Sequential Quadratic Programming (SQP) algorithm is used for the optimization. The volatility of the equal-risk-contribution portfolio lie between those of the minimum-volatility and the equally-weighted-portfolio (Maillard et al. [2010], Roncalli [2013]). The equal-risk-contribution is mean-variance efficient under the condition that all assets contribute equally to the portfolio excess return (Bruder and Roncalli [2012]).

3.5 Risk-efficient portfolio

This risk-efficient portfolio (Amenc et al. [2011]), weighs more on stocks that has a larger contribution to the portfolio's Sharpe ratio than for the stocks with a smaller contribution. Notice that this portfolio requires some information about expected returns. Instead of using usual expected returns in the optimization problem, a methodology similar to Fama and French [1993] is used. Decile portfolios are created according to the stocks' semi-deviation. Then, the median semi-deviation of each decile portfolio, $\boldsymbol{\xi} \equiv (\xi_1, \ldots, \xi_{10})'$, is computed and is attributed to each stock in this decile portfolio as an estimate of its expected return. The risk-efficient portfolio is defined as:

$$\mathbf{w}_{\text{ref}} \equiv \underset{\mathbf{w} \in \mathcal{C}_{\text{FI}^*}}{\operatorname{argmax}} \left\{ \frac{\mathbf{w}'\mathbf{J}\boldsymbol{\xi}}{\sqrt{\mathbf{w}'\mathbf{\Sigma w}}} \right\} \,,$$

with $\mathbf{J}$ a $(N \times 10)$ matrix of zeros with the element (i, j) equal to 1 if the semi-variance of stock i belongs to decile j, and $\mathcal{C}_{\text{FI}^*} \equiv \{\mathbf{w} \in \mathbb{R}^N \,|\, \mathbf{w}'\boldsymbol{\iota} = 1, (1/2N)\boldsymbol{\iota} \leq \mathbf{w} \leq (2/N)\boldsymbol{\iota}\}$ is the investment constraint on weights.

4 Covariance matrix specifications

We study covariance misspecification impact by considering four estimators for the covariance matrix. The first one is the sample covariance matrix. The second, developed by Ledoit and Wolf [2003], is a combination of the sample covariance matrix and a structured covariance matrix based a (market) single factor mode. The last two are dynamic models: the exponential weighted moving average and the dynamic conditional correlation models. For ease of presentation, expected returns are assumed to be zero.

4.1 Sample covariance

The sample covariance matrix is a standard approach used in statistics and is easy to implement. However, its simplicity has a cost. Indeed, the sample covariance matrix can create a lot of problems when used as input in an portfolio optimization setting. That is because when the number of assets is large, relative to the number of observations, the covariance matrix is estimated with a lot of error. For a window of T observations, the $(N \times N)$ sample covariance matrix is defined as:

$$\hat{\mathbf{\Sigma}} \equiv \sum_{t=1}^{T} w_t \mathbf{r}_{T-(t-1)} \mathbf{r}'_{T-(t-1)} \,, \tag{4}$$

where $\mathbf{r}_t \equiv (r_{1,t}, \dots, r_{N,t})'$ and $w_t \equiv 1/(T-1)$.

4.2 Ledoit-Wolf

Ledoit and Wolf [2003] develop a covariance matrix based on a statistical technique called *shrinkage* which dates back to Stein et al. [1956]. They propose to take a weighted average of the sample covariance matrix with the Sharpe single-index model covariance estimator. The weight or optimal shrinkage intensity assigned to the Sharpe ratio model lies between zero and one and controls how much structure is imposed to the estimator. Hence, the single-index model is the *shrinkage target*.

The shrinkage intensity depends on the correlation between estimation error. Indeed, when estimation error between the shrinkage target and the sample covariance matrix are positively (negatively) correlated, benefit from shrinkage is smaller (larger).

In order to compute the Ledoit-Wolf covariance matrix, let first define $\mathbf{F}$ as the shrinkage target matrix based on a single index market factor model:

$$r_{i,t} = \alpha_i + \beta_i r_{m,t} + \epsilon_{i,t} \,, \tag{5}$$

where $r_{i,t}$ is the asset i return at time t, $r_{m,t}$ is the market return at time t and $\epsilon_{i,t}$ is an error term. From (5), the shrinkage target $\mathbf{F}$ can be written as:

$$\mathbf{F} = s_m^2 \mathbf{b}\mathbf{b}' + \mathbf{D},$$

where s_m^2 is the sample variance of the market returns $r_{m,t}$, $\mathbf{b} \equiv (\beta_1, \dots, \beta_N)'$ and $\mathbf{D}$ is a diagonal matrix of the residuals variance. The Ledoit-Wolf covariance matrix estimator, $\hat{\mathbf{\Sigma}}_{Shrink}$, is a weighted average of the single factor model based estimator $\hat{\mathbf{F}}$ and the sample covariance matrix estimator $\hat{\mathbf{S}}$:

$$\hat{\boldsymbol{\Sigma}}_{Shrink} \equiv \delta^* \hat{\mathbf{F}} + (1 - \delta^*) \hat{\mathbf{S}} \,. \tag{6}$$

In (6), δ^* is referred to as the optimal shrinkage intensity. The intuition behind (6) is that a compromise between two extreme estimators should perform better than either extreme. As mentioned by Ledoit-Wolf, *"most people would prefer the 'compromise' of one bottle of Bordeaux and one steak to either 'extreme' of two bottles of Bordeaux (and no steak) or two steaks (and no Bordeaux)"*.

The optimal shrinkage intensity is based on the minimization the expected quadratic loss function. When the number of observations is large, the optimal shrinkage intensity is close to zero: the Ledoit-Wolf covariance matrix converges to the sample covariance matrix.

4.3 Exponential weighted moving average

The Exponential weighted moving average (EWMA) covariance matrix is a dynamic model popularized by RiskMetrics [1996]. Instead of setting the weights of (4) equals to $1/(T-1)$, weights are proportional to a decay factor, λ, which lies between zero and one:

$$w_k \equiv (1 - \lambda)\lambda^{k-1},\, k \in \{1, 2, \ldots, \infty\} \,,$$

where $w_k \geq 0$ and $\lim_{k \to \infty} \sum_{s=1}^{k} w_s = 1$. Hence, more recent observations have more weight in the estimator. Using recursion, the EWMA covariance matrix can be expressed as:

$$\hat{\boldsymbol{\Sigma}}_t \equiv \lambda \hat{\boldsymbol{\Sigma}}_{t-1} + (1 - \lambda) \mathbf{r}_{t-1} \mathbf{r}_{t-1} \,,$$

where $\hat{\boldsymbol{\Sigma}}_0$ is typically initialized to the sample covariance matrix. We use $\lambda = 0.94$ which is selected by RiskMetrics [1996] as the optimal decay factor for daily data.

4.4 The dynamic conditional correlation model

We denote $\mathcal{F}_{t-1}$ as the set of information available at time $t-1$ and $\mathbf{r}_t \equiv (r_{1t}, r_{2t}, ..., r_{Nt})'$ as a vector of demeaned return at time t with $\mathbb{E}[\mathbf{r}_t | \mathcal{F}_{t-1}] = 0$ and $\mathbb{V}\text{ar}(\mathbf{r}_t | \mathcal{F}_{t-1}) = \mathbf{H}_t$. In a dynamic framework, returns satisfy the following relation:

$$\mathbf{r}_t \,|\, \mathcal{F}_{t-1} = \mathbf{H}_t^{1/2} \mathbf{z}_t \,,$$

where $\mathbf{z}_t$ is a $(N \times 1)$ i.i.d. white noise vector. The constant conditional correlation (CCC) model by Bollerslev [1990] defines the conditional covariance matrix as:

$$\mathbf{H}_t \equiv \mathbf{D}_t \mathbf{R} \mathbf{D}_t \,,$$

where $\mathbf{D}_t$ is a $(N \times N)$ diagonal matrix of conditional volatilities and $\mathbf{R}$ is a $(N \times N)$ matrix of constant correlations. The conditional variances are estimated using the GARCH(1,1) model. Engle [2002] relaxes the constant correlations hypothesis with the dynamic conditional correlation (DCC) model:

$$\mathbf{H}_t \equiv \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t \,.$$

As for the CCC, with the assumption that the variance processes are GARCH(1,1), the matrix $\mathbf{D}_t^2$ is defined as:

$$\mathbf{D}_t^2 \equiv Diag\{\boldsymbol{\omega}\} + Diag\{\boldsymbol{\kappa}\}\mathbf{r}_{t-1}\mathbf{r}'_{t-1} + Diag\{\boldsymbol{\lambda}\}\mathbf{D}_{t-1}^2 \,.$$

where $\boldsymbol{\omega}$, $\boldsymbol{\kappa}$ and $\boldsymbol{\lambda}$ are $(N \times 1)$ vectors of parameters.

The estimation of the conditional correlation matrix is not straightforward since a proxy process $\mathbf{Q}_t$ must be estimated and rescaled thereafter. This proxy process, defined as:

$$\mathbf{Q}_t \equiv (1-\alpha-\beta)\bar{\mathbf{Q}} + \alpha(\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}'_{t-1}) + \beta\mathbf{Q}_{t-1} \,,$$

needs non-negative values for α and β with the restriction $\alpha + \beta < 1$ to ensure both stationarity and definitiveness of $\mathbf{Q}_t$. Here, $\boldsymbol{\epsilon}_t = \mathbf{D}^{-1}\mathbf{r}_t$ and $\bar{\mathbf{Q}}$ is the unconditional correlation matrix of $\boldsymbol{\epsilon}$. Finally, the conditional correlation matrix $\mathbf{R}_t$ is obtained by rescaling $\mathbf{Q}_t$:

$$\mathbf{R}_t \equiv Diag\{\mathbf{Q}_t\}^{-1/2}\mathbf{Q}_t Diag\{\mathbf{Q}_t\}^{-1/2} \,.$$

4.4.1 Covariance forecast

The forecast of elements on the diagonal of $\mathbf{D}_{t+k}$ are easily obtained from the GARCH model. The conditional variance forecast k-step ahead for asset i is obtained as:

$$\mathbb{E}[h_{i,t+k}] = \bar{\sigma}_i^2 + (\kappa_i + \lambda_i)^k(h_{i,t} - \bar{\sigma}_i^2) \,.$$

where $\bar{\sigma}$ is the unconditional volatility of the asset i. Because of the nonlinearity of the correlation process, some assumptions are made for the computation of the conditional correlation forecast which results in an approximation of the true forecast matrix. Those assumptions are that $\bar{\mathbf{Q}} = \bar{\mathbf{R}}$ and $\mathbb{E}_t[\mathbf{Q}_{t+1}] = \mathbb{E}_t[\mathbf{R}_{t+1}]$, leading to:

$$\mathbf{Q}_{t+1} = (1-\alpha-\beta)\bar{\mathbf{Q}} + \alpha\mathbb{E}[\mathbf{z}_t\mathbf{z}'_t] + \beta\mathbf{Q}_t \,,$$

and $\mathbf{Q}_{t+1}$ is standardized according to:

$$\mathbf{R}_{t+1} = Diag\{\mathbf{Q}_{t+1}\}^{-1/2}\mathbf{Q}_{t+1} Diag\{\mathbf{Q}_{t+1}\}^{-1/2} \,.$$

Finally, the forecast k-step ahead of the conditional correlation matrix is obtained as:

$$\mathbb{E}_t[\mathbf{R}_{t+k}] = \sum_{i=0}^{k-2}(1-\alpha-\beta)\bar{\mathbf{R}}(\alpha+\beta)^i + (\alpha+\beta)^{k-1}\mathbf{R}_{k+1} \,. \qquad (7)$$

Engle and Sheppard [2001] find that this approximation is the one that leads to the smallest bias.

4.4.2 Parameter estimation

Define θ as the parameters related to $\mathbf{D}$ and ϕ the parameters related to $\mathbf{R}$. Given that the random vector $\mathbf{z}_t = \mathbf{H}_t^{-1/2}\boldsymbol{\epsilon}_t$ follows a multivariate normal distribution where $\mathbb{E}[\mathbf{z}_t] = 0$ and $\mathbb{E}[\mathbf{z}_t\mathbf{z}'_t] = \mathbf{I}$, the likelihood function is equal to:

$$L(\theta,\phi) \equiv \prod_{i=1}^{T} \frac{1}{2\pi^{n/2}|\mathbf{H}_t|^{1/2}} \exp\left\{-\frac{1}{2}\boldsymbol{\epsilon}'_t\mathbf{H}_t^{-1}\boldsymbol{\epsilon}_t\right\} \,. \qquad (8)$$

Taking the logarithm of (8) combined with some manipulations, results in the following log-likelihood form:

$$l(\theta,\phi) \equiv l_V(\theta) + l_C(\theta,\phi)\,,$$

where:

$$l_V(\theta) \equiv -\frac{1}{2}\sum_{t=1}^{T}\left(n\log(2\pi) + \log|\mathbf{D}_t|^2 + \boldsymbol{\epsilon}_t'\mathbf{D}_t^{-2}\boldsymbol{\epsilon}_t\right)\,,$$

$$l_C(\theta,\phi) \equiv -\frac{1}{2}\sum_{t=1}^{T}\left(\log|\mathbf{R}_t| + \mathbf{z}_t'\mathbf{R}_t^{-1}\mathbf{z}_t - \mathbf{z}_t'\mathbf{z}_t\right)\,.$$

Maximizing each individual term is equivalent to maximizing (8). In order to get the DCC parameters, a two-steps approach is used. The first step consists in:

$$\hat{\theta} \equiv \arg\max_{\theta}\{L_V(\theta)\}\,,$$

and takes the estimated set of parameters as input in the second step:

$$\hat{\phi} \equiv \arg\max_{\phi}\{L_C(\hat{\theta},\phi)\}\,.$$

However, Engle, Shephard, and Sheppard [2008] show that in large-scale estimation settings, parameters α and β are estimated with bias when using the conventional Maximum Likelihood. Hence, the solution is the Composite Likelihood (CL) defined as:

$$l_C(\theta,\phi) \equiv \sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{j>1}\log f(\theta,\phi,r_{i,t},r_{j,t})\right)\,,$$

where $f(\theta,\phi,r_{i,t},r_{j,t})$ is the bivariate normal distribution of asset pair i and j and where covariance targeting is imposed. The composite log-likelihood averages the log-likelihoods of pairs of assets. Each pair yields a valid (but inefficient) likelihood for α and β, but averaging over all pairs produces an estimator which is relatively efficient, numerically fast, and free of bias even in large-scale problems. In our context, $l_C(\theta,\phi)$ is equal to:

$$l_C(\theta,\phi) \equiv -\frac{1}{2}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N}\sum_{j>1}\left(\log(1-\rho_{ij,t}^2) + \frac{z_{i,t}^2 + z_{j,t}^2 - 2\rho_{ij,t}z_{i,t}z_{i,t}}{1-\rho_{ij,t}^2}\right)\right)\,.$$

We use a dataset of dimension $N = 30$ in our study, which can be considered as a large data set. The composite likelihood technique is therefore more robust.

5 Monte Carlo study

In portfolio construction, required inputs are subject to estimation risk which may result in excessive portfolio turnover as well as poor out-of-sample performance. Consider a portfolio manager who rebalances its portfolio actively every day according to the minimum-variance portfolio (1). On one hand, the portfolio manager can use the sample covariance matrix as input. On the other hand, he can use a forecast of the covariance matrix over the investment horizon. In the latter case, (1) is now be defined as:

$$\mathbf{w}_t^* \equiv \underset{\mathbf{w}\in\mathcal{C}_{\text{FI}}}{\operatorname{argmin}} \left\{ \mathbf{w}'\boldsymbol{\Sigma}_{t+1|t}\mathbf{w} \right\} , \tag{9}$$

where $\boldsymbol{\Sigma}_{t+1|t}$ is the covariance matrix forecast one-day ahead. Whatever is the chosen covariance matrix, the portfolio weights are subject to estimation error since the correct specification of the covariance matrix is unknown. This estimation risk becomes important when rebalancing is frequent and trading costs are taken into account.

5.1 Setup

The aim of this section is to measure portfolio misallocation arising from the covariance matrix estimation for the previously defined risk-based portfolios. To that aim, we rely on a Monte Carlo study. The data generating process (DGP) used for the simulations is:

$$\mathbf{r}_t \,|\, \mathcal{F}_{t-1} \sim \mathcal{N}(0, \mathbf{H}_t) \,,$$

where $\mathbf{H}_t$ is computed according to a dynamic conditional correlation model (DCC).

DCC parameters are calibrated with $N = 30$ stocks belonging to the Dow Jones Composite index in date of September 2014 with daily adjusted returns. The estimated parameters are presented in Table 2. In general, shocks on volatilities tend to persist over time since the sum of the coefficients $\alpha + \beta$ are close to one. Persistence in correlations is also observed.

Setting k as the number of days between two rebalancing dates, $1736 + k$ observations are generated for each stock. Estimated covariances matrix are computed using the first 1736 observations while the remaining k observations are used to compute the true covariance matrix at the rebalancing date. The covariance estimation error on weights is then computed as,

$$d \equiv \sum_{i=1}^{30} \left| w_{i,\text{true}} - w_{i,\text{estimated}} \right| . \tag{10}$$

where $w_{i,\text{true}}$ are the portfolio weights obtained using the true covariance matrix and $w_{i,\text{estimated}}$ are the portfolio weights obtained using the misspecified covariance matrix. More precisely, we follow the steps below:

1. Use the fitted DCC model on daily-adjusted returns of the 30 stock belonging the Dow Jones index as the DGP;

2. Simulate $1736 + k$ observations for each stock according to the DGP;

3. Compute:

- the true covariance;
- the covariance forecast according to a DCC model fitted to the simulated data;
- the sample covariance matrix with the simulated data;
- the EWMA covariance matrix with the simulated data;
- the Ledoit-Wolf covariance matrix with the simulated data;

4. Compute:

   - the equal-risk-contribution portfolio for each computed covariance matrix;
   - the maximum-diversification portfolio for each computed covariance matrix;
   - the risk-efficient portfolio for each computed covariance matrix;
   - the minimum-variance portfolio for each computed covariance matrix;

5. Compute the distance following (10);

6. Repeat step one to step five 1000 times.

5.2 Results

First, we set the forecast horizon equal to one day and proceed to one thousand simulations. Figure 1 presents the box plots of the distance (see 10) between portfolio computed with the estimated covariance matrix and the true covariance matrix.

[Insert Figure 1 about here.]

At first, we see that there is a common pattern between each portfolio. In general, the allocation can be significantly improved with a dynamic approach for the covariance estimator. Overall, the covariance forecast that provides the most accurate allocation among portfolios is the dynamic-conditional-correlation model. That is followed by the exponential-weighted-smoothing average model. Finally the sample and Ledoit-Wolf covariances matrices result in the worst allocation.

The asset weights given by the minimum-volatility portfolio are extremely sensitive to covariance misspecification in terms of both average distance and uncertainty around the distance. In fact, for the tested portfolio set, the minimum-volatility portfolio is the most sensitive and the less robust to covariance misspecification. That is followed by the maximum-diversification and the risk-efficient portfolio.

It is clear that the portfolios for which the covariance misspecification has the less impact are the inverse-volatility and the equal-risk-contribution respectively. For both, the average distance and the uncertainty around the distance are similar except for the exponential-weighted-moving-average. In the case of the inverse volatility, the allocation with the exponential-weighted-moving-average is not that far from the allocation with the dynamic-conditional-correlation and the variance of the portfolio's allocation is also relatively small. The risk-efficient portfolio as well as the maximum-diversification portfolio are more sensitive than the equal-risk-contribution but still less than the minimum-volatility portfolio.

[Insert Figure 2 about here.]

For the weekly rebalancing results presented in Figure 2, the characteristics of the distance distribution are almost identical to the ones from the daily rebalancing. The main difference is in the portfolio allocations given the dynamic-conditional-correlation model where the distance distribution becomes more volatile and the average distance from the real optimal allocation increased compared.

6 Backtest

In this section, we proceed as in Maillard et al. [2010]. The aim is to study the out-of-sample performance of the risk-based portfolios and assess their sensitivity with respect to the covariance matrix estimator.

6.1 Setup

Every combination portfolio/covariance is considered as an investment strategy for which backtest is performed using returns of the 30 stocks belonging to the Dow Jones Composite index from March 2008 to November 2014. A three-year rolling window is used for the covariance estimation. Hence, the first rebalancing date is on April 2011. When necessary, log-return covariance matrices are transformed into arithmetic return covariance matrices following Meucci [2001].

We present results only for daily rebalancing since weekly and monthly rebalancing results are similar. For each strategy, the portfolio's annual performance, annual volatility as well as the information ratio are computed (return divided by the volatility). Risk measures such as the daily historical Value-at-Risk (VaR) 95% and daily historical Expected Shortfall (ES) 95% are also computed. The average turnover ($\bar{T}$) is included in order to have a better idea of which covariance estimator leads to more stable weights over time. The turnover of each period is measured as:

$$\text{Turn} = \sum_{i=1}^{N} \left| w_{i,t+1} - w_{i,t^+} \right| ,$$

where $w_{n,t+1}$ are the weights after the rebalancing and w_{n,t^+} are the weights just before rebalancing such as:

$$\hat{w}_{i,t^+} = \frac{w_{i,t}(1 + r_{i,t+1})}{\sum_{n=1}^{N} w_{i,t}(1 + r_{i,t+1})} . \tag{11}$$

Here, the turnover measure can be interpreted as the average percentage of wealth traded on each period. As measure of diversification, the average value of the Gini coefficient ($\bar{G}_w$) of the portfolio weights is computed. We use the average value of the Gini coefficient associated with the asset risk-contribution ($\bar{G}_{\mathcal{RC}}$) as a measure of risk concentration. Recall that the Gini coefficient is a measure of inequality and lies between zero and one. A lower Gini coefficient involves better repartition between individuals. Results are presented in Table 1.

[Insert Table 1 about here.]

6.2 Results

In terms of information-ratio, each risk-based strategy dominates the equally-weighted portfolio with the maximum-diversification as the best performer. As expected, the minimum-volatility strategy has the smallest volatility. Notice that for maximum-diversification and the minimum-volatility portfolios, both Gini coefficients associated to the stocks' weights and stocks' risk-contribution are high ($\sim 70 - 85\%$). It means that those portfolios are concentrated in few assets and that the portfolios' risk is driven by few stocks. Recall that by definition of the minimum-volatility portfolio, assets' risk-contribution are equal to their respective weights.

Better diversification properties are achieved by the equal-risk-contribution and the inverse-volatility. However, their information-ratios are lower. A good alternative might be the risk-efficient portfolio which has a higher information-ratio without erasing too much diversification. Indeed, the Gini coefficients for weights are around 30%, about twice the one of the equal-risk-contribution. The risk-contribution is also more concentrated, but still much less than in the minimum-volatility.

The choice of a dynamic approach (DCC and EWMA) instead of a static approach (sample and Ledoit-Wolf) for the covariance matrix estimator has a significant impact on portfolios' turnover. Indeed, the use of the dynamic-conditional-correlation or the exponential-weighted-smoothing-average results in turnover between 2% and 27%. For the sample and Ledoit-Wolf estimators, the turnover lies between 0.5% and 1.5%.

As an extreme case, the minimum-volatility strategy using a dynamic approach results in a turnover over 15 times higher than if a static approach were used. The impact on the maximum-diversification strategy is also important, even more for the exponential-weighted-smoothing-average which is twice the turnover given by using the dynamic-conditional-correlation model. The portfolio's turnover for the risk-efficient strategy is about half the one of the minimum-volatility portfolio for the DCC and EWMA specification.

Yet, the inverse-volatility and the equal-risk-contribution portfolios are the ones for which the choice of the covariance matrix estimator has the least impact and are the most stable. As an example, the inverse-volatility portfolio using the dynamic-conditional-correlation covariance matrix gives a turnover about 7 times smaller than when the dynamic-conditional-correlation is used for the minimum-volatility.

Overall, the resulting turnover from dynamic approach for the covariance estimators does not seem justifiable when looking at return and risk. Indeed, except for the maximum diversification, it is not obvious which covariance results in the best risk-adjusted returns. Though, it seems that the EWMA tend to provide slightly better returns. In the end, the biggest impact of the covariance matrix estimator is definitely on portfolios' turnover.

From a practical point of view, a portfolio manager should be aware of the impact that his choice of covariance matrix estimator might have on the stability of his portfolio.

7 Conclusions

An increasing amount of investors has been forsaking traditional investment methodologies for risk-based methodologies which consist in allocating risk instead of capital and does not require the use expected returns. Based on this approach, portfolios which have been the most extensively used are the equal-risk-contribution portfolio, the maximum-diversification portfolio, the minimum-variance portfolio, the risk-efficient portfolio and the inverse-volatility portfolio. Although these portfolios present interesting properties, they are still subject to some concerns.

On one hand, we show in a Monte Carlo framework that the portfolios that are the less sensitive portfolios to misspecification of the covariance matrix are the inverse-volatility and the equal-risk-contribution. On the opposite, the minimum-volatility portfolio is extremely sensitive to the covariance matrix estimator. In general, the dynamic-conditional-correlation and the exponential-weighted-smoothing-average covariances can improve significantly the asset allocation when compared to the sample or Ledoit-Wolf covariance matrix estimators.

On the other hand, in a historical simulation framework, using the exponential-weighted-smoothing-average or the dynamic-conditional-correlation dynamic as a covariance estimator increases significantly portfolios turnover without resulting in significant higher returns. Hence, a portfolio manager subject to transaction cost might be better off using the sample or Ledoit-Wolf covariance estimator.

Finally, portfolios managers should be aware that, as for the MVO, concerns persist for risk-based portfolios. This research could be extend to a set of different investment universes as well as a widen set of covariance matrix estimators.

Bibliography

Noël Amenc, Felix Goltz, Lionel Martellini, and Patrice Retkowsky. Efficient indexation: An alternative to cap-weighted indices. *Journal Of Investment Management, Fourth Quarter*, 2011.

David Ardia and Kris Boudt. Implied expected returns and the choice of a mean-variance efficient portfolio proxy. *Journal of Portfolio Management*, 41(4):68–81, 2015.

Michael J Best and Robert R Grauer. Capital asset pricing compatible with observed market value weights. *The Journal of Finance*, 40(1):85–103, 1985.

Fischer Black and Robert Litterman. Global portfolio optimization. *Financial Analysts Journal*, 48(5):28–43, 1992.

Tim Bollerslev. Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *The Review of Economics and Statistics*, 72(3):498–505, 1990.

Benjamin Bruder and Thierry Roncalli. Managing risk exposures using the risk budgeting approach. Available at SSRN 2009778, 2012.

Yves Choueifaty and Yves Coignard. Toward maximum diversification. *Journal of Portfolio Management*, 35(1):40, 2008.

Roger Clarke, Harindra De Silva, and Stephen Thorley. Minimum variance, maximum diversification, and risk parity: an analytic perspective. *Journal of Portfolio Management*, 39(3):39–53, 2013.

Victor DeMiguel, Lorenzo Garlappi, and Raman Uppal. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, 22(5):1915–1953, 2009.

Robert Engle. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20 (3):339–350, 2002.

Robert F Engle and Kevin Sheppard. Theoretical and empirical properties of dynamic conditional correlation multivariate garch. Technical report, National Bureau of Economic Research, 2001.

Robert F Engle, Neil Shephard, and Kevin Sheppard. Fitting and testing vast dimensional time-varying covariance models. NYU Working Paper No. FIN-07-046, 2008.

Eugene F Fama and Kenneth R French. Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56, 1993.

Alexios Ghalanos. *rmgarch: Multivariate GARCH models*. https://cran.r-project.org/web/packages/rmgarch/index.html, 2012. R package version 0.98.

Olivier Ledoit and Michael Wolf. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of empirical finance*, 10(5):603–621, 2003.

Olivier Ledoit and Michael Wolf. Honey, i shrunk the sample covariance matrix. *The Journal of Portfolio Management*, 30(4):110–119, 2004.

Raul Leote de Carvalho, Xiao Lu, and Pierre Moulin. Demystifying equity risk–based strategies: A simple alpha plus beta description. *The Journal of Portfolio Management*, 38(3):56–70, 2012.

Moshe Levy and Richard Roll. The market portfolio may be mean/variance efficient after all. *Review of Financial Studies*, 23(6):2464–2491, 2010.

John Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The review of economics and statistics*, 47(43):13–37, 1965.

Sébastien Maillard, Thierry Roncalli, and Jérôme Teïletche. The properties of equally weighted risk contribution portfolios. *The Journal of Portfolio Management*, 36(4):60–70, 2010.

Harry Markowitz. Portfolio selection*. *The journal of finance*, 7(1):77–91, 1952.

Robert C. Merton. On estimating the expected return on the market. *Journal of Financial Economics*, 8(4):323–361, 1980.

Attilio Meucci. Common pitfalls in mean-variance asset allocation. *Wilmott Tech*, 2001.

Richard O Michaud. The markowitz optimization enigma: is' optimized'optimal? *Financial Analysts Journal*, 45(1):31–42, 1989.

RiskMetrics. *RiskMetrics Technical Document*. J. P. Morgan/Reuters,, 1996.

Thierry Roncalli. *Introduction to Risk Parity and Budgeting*. CRC Press, 2013.

William F Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk*. *The journal of finance*, 19(3):425–442, 1964.

Charles Stein et al. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. *Proceedings of the Third Berkeley symposium on mathematical statistics and probability*, 1(399):197–206, 1956.

Jack L Treynor. Toward a theory of market value of risky assets. Unpublished manuscript, 1961.

Valeriy Zakamulin. A test of covariance-matrix forecasting methods. *The Journal of Portfolio Management*, 41(3):97–108, 2015.

A Tables

Table 1: **Backtest results**: This table presents the performance statistics for the 6 different strategies when the portfolios are rebalanced on a daily basis. Here, r is the annualized daily average return and σ is the annualized volatility, both in percent. Hence, the Information Ratio is also on an annual basis. VaR_{95} is the daily historical Value-at-Risk 95%. ES_{95} is the daily historical Expected Shortfall 95%. $\bar{T}$ is the daily average turnover of the strategy expressed in percent. $\bar{Q}_w$ is the average Gini coefficient associated the the weights and $\bar{Q}_{\mathcal{RC}}$ is the average Gini coefficient associated to the risk-contribution. Both are expressed in percent.

Strategy	Cov.	r	σ	I-R	VaR_{95}	ES_{95}	$\bar{T}$	$\bar{Q}_w$	$\bar{Q}_{\mathcal{RC}}$
$\mathbf{w}_{\text{ew}}$	dcc	16.74	14.24	1.18	-1.37	-2.18	0.62	0.00	15.92
	ewma	16.74	14.24	1.18	-1.37	-2.18	0.62	0.00	18.45
	lw	16.74	14.24	1.18	-1.37	-2.18	0.62	0.00	16.49
	sample	16.74	14.24	1.18	-1.37	-2.18	0.62	0.00	16.50
$\mathbf{w}_{\text{iv}}$	dcc	16.32	13.39	1.22	-1.35	-2.06	3.58	14.34	7.43
	ewma	16.29	13.39	1.22	-1.35	-2.04	2.58	14.49	11.76
	lw	16.40	13.38	1.23	-1.34	-2.05	0.62	14.36	6.06
	sample	16.40	13.38	1.23	-1.34	-2.05	0.62	14.36	6.11
$\mathbf{w}_{\text{erc}}$	dcc	16.85	13.22	1.28	-1.36	-2.01	3.76	15.59	0.00
	ewma	17.40	13.22	1.32	-1.29	-2.01	3.86	17.93	0.00
	lw	16.54	13.24	1.25	-1.31	-2.03	0.64	16.41	0.00
	sample	16.54	13.24	1.25	-1.31	-2.03	0.64	16.31	0.00
$\mathbf{w}_{\text{ref}}$	dcc	18.22	13.62	1.34	-1.28	-2.03	12.43	32.83	33.42
	ewma	18.66	13.73	1.36	-1.29	-2.04	8.78	33.00	31.39
	lw	17.42	13.43	1.30	-1.28	-2.04	2.48	32.92	32.84
	sample	17.43	13.46	1.29	-1.29	-2.04	2.50	32.91	32.98
$\mathbf{w}_{\text{md}}$	dcc	21.92	12.60	1.74	-1.19	-1.83	13.61	69.59	69.42
	ewma	24.04	13.00	1.85	-1.24	-1.81	24.41	77.94	77.90
	lw	18.85	12.66	1.49	-1.19	-1.90	1.46	71.19	69.84
	sample	18.92	12.76	1.48	-1.22	-1.90	1.51	71.42	70.23
$\mathbf{w}_{\text{min}}$	dcc	15.01	10.68	1.41	-1.00	-1.52	26.87	84.92	84.92
	ewma	17.26	10.98	1.57	-1.05	-1.53	21.57	83.45	83.45
	lw	15.20	10.61	1.43	-1.00	-1.54	1.09	84.48	84.48
	sample	15.31	10.63	1.44	-0.98	-1.54	1.17	84.62	84.62

B Figures

Figure 1: **Daily rebalancing:** This figure presents the distribution of the allocation error for the four choices of covariance matrices (DCC, EWMA, LW, Sample) and the five risk-based portfolios ($\mathbf{w}_{\text{erc}}$,$\mathbf{w}_{\text{iv}}$, $\mathbf{w}_{\text{md}}$,$\mathbf{w}_{\text{min}}$, $\mathbf{w}_{\text{ref}}$). The errors are computed according to (10) and assuming that the strategy is rebalanced on a daily basis.

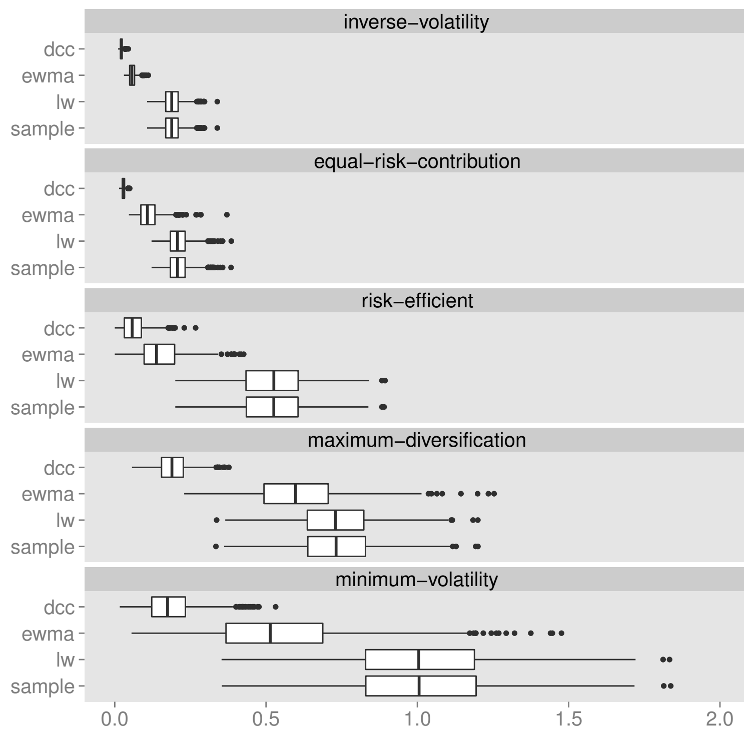
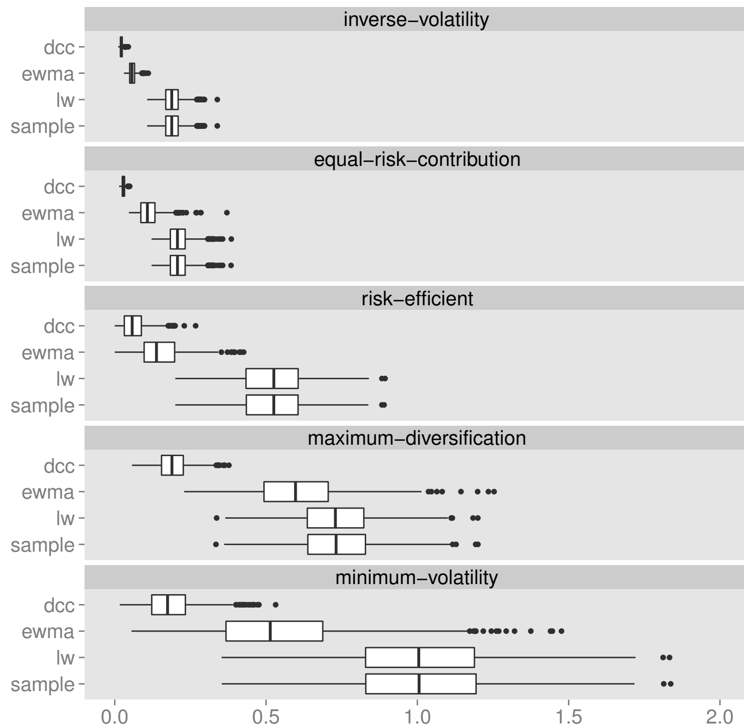


Figure 2: **Weekly rebalancing:** This figure presents the distribution of the allocation error for the four choices of covariance matrices (DCC, EWMA, LW, Sample) and the five risk-based portfolios ($\mathbf{w}_{\text{erc}}$,$\mathbf{w}_{\text{iv}}$, $\mathbf{w}_{\text{md}}$,$\mathbf{w}_{\text{min}}$, $\mathbf{w}_{\text{ref}}$). The errors are computed according to (10) and assuming that the strategy is rebalanced on a weekly basis.



C DGP parameters

Table 2: DGP parameters. The parameters are estimated with the past three years of daily returns for the 30 stock belonging to the Dow Jones index.

Asset	ω	α	β	$\alpha + \beta$
		GARCH		
1	<0.000	0.07291	0.9138	0.9867
2	<0.000	0.08979	0.8890	0.9788
3	<0.000	0.08923	0.9037	0.9929
4	<0.000	0.07432	0.9151	0.9895
5	<0.000	0.09601	0.8917	0.9878
6	<0.000	0.09845	0.8896	0.9881
7	<0.000	0.02129	0.9648	0.9861
8	<0.000	0.09252	0.8794	0.9719
9	<0.000	0.10817	0.8781	0.9863
10	<0.000	0.09193	0.8925	0.9845
11	<0.000	0.05631	0.9377	0.9940
12	<0.000	0.06886	0.9268	0.9957
13	<0.000	0.10323	0.8754	0.9786
14	<0.000	0.04905	0.9342	0.9833
15	<0.000	0.13462	0.8185	0.9531
16	<0.000	0.07956	0.9145	0.9941
17	<0.000	0.14918	0.8115	0.9607
18	<0.000	0.04233	0.9470	0.9893
19	<0.000	0.07422	0.9048	0.9790
20	<0.000	0.05460	0.9315	0.9861
21	<0.000	0.07954	0.8763	0.9559
22	<0.000	0.05553	0.9359	0.9915
23	<0.000	0.13296	0.8354	0.9684
24	<0.000	0.07643	0.9144	0.9908
25	<0.000	0.07657	0.9114	0.9879
26	<0.000	0.04309	0.9526	0.9957
27	<0.000	0.09182	0.8843	0.9761
28	<0.000	0.15431	0.8235	0.9778
29	<0.000	0.05364	0.9329	0.9865
30	<0.000	0.08574	0.8973	0.9830
		DCC		
-	-	0.0222	0.9566	0.9788