Recursive MEWMA Projections of Conditional Covolatilities in Large Portfolios

Radek Hendrych[1] | *Charles University, Prague, Czechia*
Tomáš Cipra[2] | *Charles University, Prague, Czechia*

Received 26.11.2024 (revision received 28.2.2025), **Accepted** (reviewed) 8.4.2025, **Published** 12.9.2025

Abstract

Dynamic predictions of large dimensional conditional covariance matrices are considered in the context of large financial portfolios. Since numerically simple prediction methods are usually recommended for multivariate conditional covariances (covolatilities), one prefers in this paper the multivariate EWMA (exponentially weighted moving average) processes extending the recursive estimation of EWMA processes to the multivariate case. Moreover, various modifications of recursive MEWMA projections are suggested to improve the quality of covolatility projections. An extensive numerical study for real stock indices portfolios compares types of covolatility projections employing various criteria and tests.

Keywords	***DOI***	***JEL code***
Covolatilities projections, large covariance matrices, multivariate EWMA, multivariate GARCH models, recursive estimation	*https://doi.org/10.54694/stat.2024.80*	*C51, C53, C58, G11*

INTRODUCTION

A frequent problem in multivariate dynamic systems is estimating and predicting the corresponding conditional covariance matrices. Typically, one forecasts conditional volatilities and covolatilities (i.e., multivariate volatilities) in a portfolio of financial assets. However, technical applications for systems of signals are also possible.

The MEWMA model (Multivariate Exponential Weighted Moving Average) is the multivariate generalization of the univariate model EWMA, which has been primarily developed as a numerically

[1] Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovska 83, 186 75 Prague, Czechia. E-mail: hendrych@karlin.mff.cuni.cz. ORCID: <https://orcid.org/0000-0003-3222-3148>.
[2] Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovska 83, 186 75 Prague, Czechia. E-mail: cipra@karlin.mff.cuni.cz. ORCID: <https://orcid.org/0000-0003-0604-324X>.

simple alternative to the GARCH model for anticipating future volatility of data records, usually financial returns (this univariate case is supported by the commercial software RiskMetrics (1996) popular in risk management). The name of this concept originates from the conditional variance being an exponentially weighted sum of historical squared (financial) returns with the geometrically declining weights going backward in time (similarly in the multivariate case for conditional covariances). Therefore, the MEWMA model can easily track conditional volatilities and covolatilities changes.

The single unknown parameter of the classic MEWMA model determines the geometrically declining weights. This parameter (denoted usually as lambda) is conventionally prescribed by experts or users. Alternatively, it can be estimated using standard (online or batch) statistical identification procedures (e.g., the conditional maximum likelihood method).

Various models for estimating and predicting conditional variances (volatilities in the univariate case) or conditional covariance matrices (covolatilities in the multivariate case) have been suggested in this context. However, surprisingly, such models (including MEWMA models) are rarely calibrated recursively even though recursive estimation performed by recursive algorithms is undoubtedly convenient: to evaluate the parameter estimates at a time step, recursive methods operate only with the actual measurements and the quantities estimated in previous steps. It contrasts sharply with the non-recursive (batch) estimation, where all data are collected first, and then the given model is fitted. The recursive estimation techniques are effective in terms of memory storage and computational complexity, and this efficiency can be employed only in the framework of high-frequency time series data. Alternatively, it is possible to apply these methods to monitor or forecast (co)volatilities in real time, evaluate risk measures, detect faults, check model stability, identify structural changes, etc.

The recursive estimation is even more desirable in the case of multivariate GARCH (MGARCH) models, e.g., in various portfolio applications where the dynamic covolatilities play an essential role when the risk is controlled and regulated. On the other hand, one should respect that the multivariate GARCH models with higher numbers of parameters fulfilling strict constraints are complex since one must identify and calibrate large dynamic systems employing data that are (co)volatile in time. Therefore, we confine ourselves in this paper to the simplest models of this type, which are just MEWMA models (in the dynamic MEWMA model, only a single parameter suffices to be estimated recursively in time, and its modifications suggested in this paper are parameterized by a reasonable number of parameters).

The paper extends not only the recursive approach applied for the univariate EWMA model by Hendrych and Cipra (2019) to the multivariate MEWMA model and its derived variants but also presents an extensive numerical study for large stock portfolios in this context including statistical analysis of results (compared with smaller currency portfolios in Cipra and Hendrych, 2019).

The paper is organized as follows. A literature review is performed in Section 1. Section 2 presents a general framework of recursive estimation of multivariate volatility models and applies this approach to the multivariate MEWMA process. Section 3 deals with various modifications of the MEWMA process to better model its dynamic character. An extensive empirical data study in Section 4 numerically demonstrates the recursive approach to the multivariate volatility models by comparing various techniques. Finally, the last section contains concluding remarks.

1 LITERATURE REVIEW

As forecasts of conditional volatilities and covolatilities in portfolios of financial assets are concerned, various authors have dealt with this problem frequently in financial practice, see, e.g., Caldeira et al. (2017), Chiriac and Voev (2011), Cipra and Hendrych (2019), Clements et al. (2012), Engle and Sheppard (2007), Engle et al. (2008). However, technical applications for systems of signals are also frequent, see, e.g., Ledoit and Wolf (2004), Ljung and Söderström (1983). Various models for estimating and predicting conditional covariance matrices have been suggested in this context, see, e.g., Caporin and McAleer (2013), Engle and Colacito (2006), Hafner and Franses (2009), Laurent et al. (2012), and others.

Finally, one must mention in this context the generalized autoregressive score models denoted as GAS models which are observation-driven time series models that update the parameters over time using the scaled score of the likelihood function (see Creal et al., 2003, and others). The GAS framework encompasses a large number of models including the models described in this paper where the estimation of parameters of such models runs recursively in time.

In the introductory section the convenience of recursive estimation in dynamic systems has been stressed. Recently, several recursive estimation schemes suitable for the stochastic models of type GARCH have been introduced (see Hendrych and Cipra, 2016; 2018; 2019), which one can advantageously apply to the MEWMA model in this contribution. They can be characterized as numerically effective techniques that can estimate and control the model parameter (and consequently the model behavior) in real time.

2 RECURSIVE ESTIMATION OF MULTIVARIATE VOLATILITY PROCESS

A general framework for recursive estimation of multivariate volatility can be formulated employing a practical scheme that enables modeling conditional covolatilities (i.e., the conditional covariance matrices) of an m dimensional stochastic processes $\{\boldsymbol{r}_t\}$. If one ignores a possible nonzero conditional mean vector, then the corresponding model in this context usually has the form:

$$\boldsymbol{r}_t = \mathbf{H}_t^{1/2} \cdot \boldsymbol{\varepsilon}_t, \tag{1}$$

where $\{\boldsymbol{\varepsilon}_t\}$ is a sequence of *iid* random vectors with zero mean and identity covariance matrix and $\mathbf{H}_t$ is an $m \times m$ (symmetric) positive definite Ω_{t-1}-measurable matrix with square root matrix $\mathbf{H}_t^{1/2}$ such that $\mathbf{H}_t = \mathbf{H}_t^{1/2}(\mathbf{H}_t^{1/2})^{\mathrm{T}}$. It represents the covariance matrix conditioned by the information observed till and including time t - 1. Hence, the conditional moments of $\{\boldsymbol{r}_t\}$ are obviously:

$$\mathrm{E}(\boldsymbol{r}_t | \Omega_{t-1}) = \mathbf{0}, \quad \mathrm{cov}(\boldsymbol{r}_t | \Omega_{t-1}) = \mathbf{H}_t, \tag{2}$$

where Ω_t is the smallest σ-algebra such that $\{\boldsymbol{r}_s\}$ is measurable for all $s \leq t$. The additional part of this model is the covolatility equation for the matrices $\mathbf{H}_t$ which specifies particular models (e.g., model MEWMA in Section 3) and contains unknown parameters ordered in a column vector $\boldsymbol{\theta}$ so that one should write $\mathbf{H}_t(\boldsymbol{\theta})$ or even $\mathbf{H}_{t|t-1}(\boldsymbol{\theta})$ (the latter symbol underlines the fact that the conditional elements of the given matrix can be looked upon as one-step-ahead covolatility predictions for time t constructed at time $t - 1$). Similarly, $\mathbf{H}_{t|t}(\boldsymbol{\theta})$ should be a proper symbol of the covolatility estimations for time t constructed at time t.

Assuming the normality of $\boldsymbol{\varepsilon}_t$, it is possible to apply the conditional probability density:

$$f(\boldsymbol{r}_t | \Omega_{t-1}) = |2\pi\, \mathbf{H}_t(\boldsymbol{\theta})|^{-1/2} \exp\left\{-\frac{1}{2}\boldsymbol{r}_t^{\mathrm{T}} \mathbf{H}_t(\boldsymbol{\theta})^{-1} \boldsymbol{r}_t\right\}, \tag{3}$$

for construction of ML estimate of parameters $\boldsymbol{\theta}$ by minimizing the loss function (which is based on the negative conditional log-likelihood criterion):

$$\min_{\boldsymbol{\theta}} \sum_{t=1}^{T} \left[\ln|\mathbf{H}_t(\boldsymbol{\theta})| + \boldsymbol{r}_t^{\mathrm{T}} \mathbf{H}_t(\boldsymbol{\theta})^{-1} \boldsymbol{r}_t \right]. \tag{4}$$

Since the objective is to estimate the multivariate volatility processes recursively, we shall generalize the recursive prediction error method, which has been successfully applied for univariate conditional heteroscedasticity processes of the type GARCH (see Cipra and Hendrych, 2019; Hendrych and Cipra, 2018) to the multivariate case. It can be described algorithmically by the system of the following recursive formulas, which are well-known from the literature on the identification of dynamic systems, see, e.g., Ljung and Söderström (1983) and others:

$$\hat{\boldsymbol{\theta}}_t = \hat{\boldsymbol{\theta}}_{t-1} - \eta_t \mathbf{R}_t^{-1} \mathbf{F}_t'(\hat{\boldsymbol{\theta}}_{t-1})^{\mathrm{T}}, \tag{5}$$

$$\mathbf{R}_t = \mathbf{R}_{t-1} + \eta_t \left[\widetilde{\mathbf{F}}_t''(\hat{\boldsymbol{\theta}}_{t-1}) - \mathbf{R}_{t-1} \right], \tag{6}$$

$$\eta_t = \frac{1}{1 + \alpha_t / \eta_{t-1}} \text{ for } \textit{forgetting factor } \alpha_t = \widetilde{\alpha} \cdot \alpha_{t-1} + (1 - \widetilde{\alpha}), \ \ \alpha_0, \widetilde{\alpha} \in (0, 1), \eta_0 = 1, \tag{7}$$

where:

$$F_t(\boldsymbol{\theta}) = \ln \left| \mathbf{H}_t(\boldsymbol{\theta}) \right| + \boldsymbol{r}_t^{\mathrm{T}} \mathbf{H}_t(\boldsymbol{\theta})^{-1} \boldsymbol{r}_t, \tag{8}$$

$\mathbf{F}_t'(\boldsymbol{\theta})$ denotes the gradient of $F_t(\boldsymbol{\theta})$, $\mathbf{F}_t''(\boldsymbol{\theta})$ is the Hessian matrix of $F_t(\boldsymbol{\theta})$, and $\widetilde{\mathbf{F}}_t''(\boldsymbol{\theta})$ is the approximation of $\mathbf{F}_t''(\boldsymbol{\theta})$ such that:

$$\mathrm{E}\left(\mathbf{F}_t''(\hat{\boldsymbol{\theta}}_{t-1}) - \widetilde{\mathbf{F}}_t''(\hat{\boldsymbol{\theta}}_{t-1}) \middle| \Omega_{t-1} \right) = \mathbf{0}, \tag{9}$$

this approximation simplifies the calculation of the corresponding Hessian matrix for particular types of models; see, e.g., the model MEWMA in Section 3. The application of forgetting factor α_t is typical in the literature on the identification of dynamic systems since it improves the convergence properties, including the statistical consistency of the corresponding recursive estimators of the type (5)-(7).

Various technicalities are involved when applying these procedures, namely the choice of forgetting factor and the initialization of the recursive calculations. For the recursive algorithm of the type (5)-(7), one recommends choosing both values $\widetilde{\alpha}$ and α_0 close to one (e.g., $\widetilde{\alpha}$ = 0.95 and α_0 = 1). Regarding the initialization of these recursive algorithms, the general approach by Hendrych and Cipra (2018) can be simply adapted for the models proposed in this paper. Finally, one should introduce a simple (general) projection scheme, which completes the algorithm (5)-(7) and ensures that the algorithm will not degenerate:

$$\left[\hat{\boldsymbol{\theta}}_t \right]_{\boldsymbol{D}} = \begin{cases} \hat{\boldsymbol{\theta}}_t & \text{if } \hat{\boldsymbol{\theta}}_t \in \boldsymbol{D}, \\ \hat{\boldsymbol{\theta}}_{t-1} & \text{if } \hat{\boldsymbol{\theta}}_t \notin \boldsymbol{D}, \end{cases} \tag{10}$$

where $\boldsymbol{D}$ is a compact subset of $R^{dim(\boldsymbol{\theta})}$ reflecting all constraints on parameters (e.g., their positivity) that should be fulfilled to guarantee the symmetry and positive definiteness of $\mathbf{H}_t$.

The theoretical properties of the suggested recursive estimation algorithm coincide with the conventional non-recursive case (as t goes to infinity) when the corresponding negative conditional log-likelihood criterion is minimized. Namely, convergence and asymptotic distributional properties are identical for a sufficiently large number of observations (refer to Ljung and Söderström (1983) for the theoretical features of the prediction error method). The derivation employs instruments from the ordinary differential equation theory. Moreover, an extensive simulation study performed by the authors demonstrated the convergence behavior of the proposed recursive estimation technique.

3 MEWMA MODEL AND ITS MODIFICATIONS

In this Section, we will present explicit forms of the recursive estimation algorithm (5)-(7) for some simple multivariate models of conditional covariance matrices (covolatilities), which are based on the MEWMA principle.

3.1 Recursive MEWMA Model

In the classic MEWMA model, the covolatility equation for the matrix $\mathbf{H}_t$ has the form:

$$\mathbf{H}_t = (1-\lambda)\boldsymbol{r}_{t-1}\boldsymbol{r}_{t-1}^{\mathrm{T}} + \lambda \mathbf{H}_{t-1}, \;\; \lambda \in (0,1), \tag{11}$$

where λ is the single parameter to be estimated ($0 < \lambda < 1$). This means that the model based on the relation (11) is very parsimonious compared to the classic MGARCH models, and the constraints on this parameter are straightforward. In the context of MGARCH models, the MEWMA model could be regarded as a particular case of the so-called scalar BEKK(1,1) model (see, e.g., Bauwens et al., 2006; Cipra, 2020). Notice that the relation (11) produces conditional covariance matrices that are (symmetric) positive definite. The classic estimation approach to the model (11) consists in the (non-recursive or batch) ML estimation of the parameter λ.

Alternatively, the general algorithm (5)-(7) can be applied to estimate MEWMA recursively. The corresponding recursive MEWMA model has the following form (consult Appendix):

$$\hat{\lambda}_t = \hat{\lambda}_{t-1} - \eta_t R_t^{-1} \left[\mathrm{tr}\left(\mathbf{H}_t^{-1}(\hat{\lambda}_{t-1}) \frac{\partial \mathbf{H}_t(\hat{\lambda}_{t-1})}{\partial \lambda} \right) - \boldsymbol{r}_t^{\mathrm{T}} \mathbf{H}_t^{-1}(\hat{\lambda}_{t-1}) \frac{\partial \mathbf{H}_t(\hat{\lambda}_{t-1})}{\partial \lambda} \mathbf{H}_t^{-1}(\hat{\lambda}_{t-1}) \boldsymbol{r}_t \right], \tag{12}$$

$$R_t = R_{t-1} + \eta_t \left[\mathrm{tr}\left(\mathbf{H}_t^{-1}(\hat{\lambda}_{t-1}) \frac{\partial \mathbf{H}_t(\hat{\lambda}_{t-1})}{\partial \lambda} \mathbf{H}_t^{-1}(\hat{\lambda}_{t-1}) \frac{\partial \mathbf{H}_t(\hat{\lambda}_{t-1})}{\partial \lambda} \right) - R_{t-1} \right], \tag{13}$$

$$\mathbf{H}_{t+1}(\hat{\lambda}_t) = (1-\hat{\lambda}_t)\boldsymbol{r}_t\boldsymbol{r}_t^{\mathrm{T}} + \hat{\lambda}_t \mathbf{H}_t(\hat{\lambda}_{t-1}), \tag{14}$$

$$\frac{\partial \mathbf{H}_{t+1}(\hat{\lambda}_t)}{\partial \lambda} = -\boldsymbol{r}_t\boldsymbol{r}_t^{\mathrm{T}} + \mathbf{H}_t(\hat{\lambda}_{t-1}) + \hat{\lambda}_t \frac{\partial \mathbf{H}_t(\hat{\lambda}_{t-1})}{\partial \lambda}, \tag{15}$$

$$\eta_t = \frac{1}{1+\alpha_t/\eta_{t-1}} \;\; \text{for } \alpha_t = \tilde{\alpha}\cdot\alpha_{t-1} + (1-\tilde{\alpha}), \;\; \alpha_0, \tilde{\alpha} \in (0,1), \eta_0 = 1. \tag{16}$$

Notice the recursive calculation of the matrix derivatives in Formula (15).

To simplify the denotation, the one-step-ahead covolatility predictions (i.e., the covariance matrix prediction) for time t constructed at time $t - 1$ utilizing the MEWMA model will be written as:

$$\hat{\mathbf{H}}_t = (1-\hat{\lambda}_{t-1})\boldsymbol{r}_{t-1}\boldsymbol{r}_{t-1}^{\mathrm{T}} + \hat{\lambda}_{t-1}\hat{\mathbf{H}}_{t-1}, \tag{17}$$

where $\hat{\lambda}_t$ is the recursive MEWMA estimate (12) of the parameter λ in the model (11) at time t. Point out that one might put $\boldsymbol{D} = [\delta, 1 - \delta]$ with $\delta \in (0, 0.5)$ in (10).

3.2 Recursive Diagonal BEKK MEWMA Model

The multivariate BEKK model is one of the most popular MGARCH models (see Bauwens et al., 2006; Cipra, 2020). In particular, the diagonal form of this model, due to a substantial reduction of parameters, provides acceptable predictions of conditional covariance matrices even for high dimensions. Such a model denoted as diagonal BEKK(1,1) can be looked upon as a valuable modification of the recursive MEWMA model, mainly when its volatility equation is written in the following form:

$$\begin{aligned}\mathbf{H}_t = {} & \mathrm{diag}\left\{\sqrt{1-\lambda^1},\ldots,\sqrt{1-\lambda^m}\right\}\ \boldsymbol{r}_{t-1}\boldsymbol{r}_{t-1}^{\mathrm{T}}\ \mathrm{diag}\left\{\sqrt{1-\lambda^1},\ldots,\sqrt{1-\lambda^m}\right\}\ + \\ & + \mathrm{diag}\left\{\sqrt{\lambda^1},\ldots,\sqrt{\lambda^m}\right\}\ \mathbf{H}_{t-1}\ \mathrm{diag}\left\{\sqrt{\lambda^1},\ldots,\sqrt{\lambda^m}\right\},\end{aligned} \tag{18}$$

where the symbol diag{...} denotes a diagonal matrix with given diagonal elements $\lambda^1, \ldots, \lambda^m$ which are m parameters to be estimated ($0 < \lambda^1, \ldots, \lambda^m < 1$). The diagonal BEKK recursive MEWMA estimates of parameters λ^k are obtained separately for each $k = 1, \ldots, m$ according to the algorithm (12)-(16) applied to diagonal elements (k, k) in (19), i.e.

$$\left(\hat{\mathbf{H}}_t\right)_{kk} = (1-\hat{\lambda}_{t-1}^k)\left(\boldsymbol{r}_{t-1}\boldsymbol{r}_{t-1}^{\mathrm{T}}\right)_{kk} + \hat{\lambda}_{t-1}^k\left(\hat{\mathbf{H}}_{t-1}\right)_{kk}, \tag{19}$$

and then the estimates of remaining non-diagonal elements (k, l), $k \neq l$, are completed as:

$$\left(\hat{\mathbf{H}}_t\right)_{kl} = \sqrt{1-\hat{\lambda}_{t-1}^k}\ \sqrt{1-\hat{\lambda}_{t-1}^l}\ \left(\boldsymbol{r}_{t-1}\boldsymbol{r}_{t-1}^{\mathrm{T}}\right)_{kl} + \hat{\lambda}_{t-1}^k\hat{\lambda}_{t-1}^l\left(\hat{\mathbf{H}}_{t-1}\right)_{kl}. \tag{20}$$

The matrices $\hat{\mathbf{H}}_t$ obtained according to (19)-(20) are due to their construction (symmetric) positive definite ones.

3.3 Recursive DCC MEWMA Model

The DCC model (Dynamic Conditional Correlations) by Engle is another helpful model proposed for modeling conditional covariances (see Bauwens et al., 2006; Cipra, 2020). Instead of a formal description of this model, we present for simplicity only the algorithm of the corresponding DCC recursive MEWMA procedure providing the predictions $\hat{\mathbf{H}}_t$ of the conditional covariance matrices exploiting the recursive EWMA approach. The algorithm is applied at time t in two steps:

(1) Standardization $\hat{\boldsymbol{r}}_t^{std}$ of vector $\boldsymbol{r}_t$:

$$\hat{\boldsymbol{r}}_t^{std} = \hat{\mathbf{D}}_t^{-1/2} \boldsymbol{r}_t \ \text{ for } \ \hat{\mathbf{D}}_t = \text{diag}\left\{ \hat{\sigma}_{t1}^2, ..., \hat{\sigma}_{tm}^2 \right\}, \tag{21}$$

where $\hat{\sigma}_{tk}^2$ is the conditional variance provided by the classical (univariate) recursive EWMA model for $\boldsymbol{r}_{tk}$, $k = 1, ..., m$.

(2) Estimation $\hat{\mathbf{H}}_t$ of matrix $\mathbf{H}_t$:

$$\hat{\mathbf{H}}_t = \hat{\mathbf{D}}_t^{1/2} \left(\text{diag}\left\{ \hat{\mathbf{Q}}_t \right\}^{-1/2} \hat{\mathbf{Q}}_t \ \text{diag}\left\{ \hat{\mathbf{Q}}_t \right\}^{-1/2} \right) \hat{\mathbf{D}}_t^{1/2}, \tag{22}$$

where $\hat{\mathbf{Q}}_t$ are the conditional covariance matrices of the standardized process $\hat{\boldsymbol{r}}_t^{std}$ estimated by the recursive MEWMA method as:

$$\hat{\mathbf{Q}}_t = (1 - \hat{\lambda}_{t-1}^{Q}) \hat{\boldsymbol{r}}_t^{std} \left(\hat{\boldsymbol{r}}_t^{std} \right)^{\mathrm{T}} + \hat{\lambda}_{t-1}^{Q} \hat{\mathbf{Q}}_{t-1}. \tag{23}$$

Again, the matrices $\hat{\mathbf{H}}_t$ obtained according to (21)-(23) are due to their construction (symmetric) positive definite ones.

4 NUMERICAL STUDY

4.1 Data

The suggested recursive MEWMA approaches for predicting conditional covariance matrices have been investigated using two data sets. We have used daily log-returns $\boldsymbol{r}_t$ of stocks from two portfolios constituting two well-known stock indices, namely:

– S&P500 (daily log-returns of adjusted close prices from Jan 3, 2018 to Oct 17, 2023);

– DJI30 (daily log-returns of adjusted close prices from Jan 3, 2018 to Oct 17, 2023).

For each stock index and fixed dimension m (namely m = 10, 30, 50, and 70 for S&P500 and 10 and 20 for DJI30), one has randomly chosen 100 sub-portfolios of dimension m to compare the success rate of particular MEWMA procedures in estimating the corresponding covolatilities over the given period.

Finally, to compare the criteria in Section 4.2, one uses available daily values of realized covariance matrices $\mathbf{RC}_t$ of both data sets S&P500 and DJI30 at selected time points $t_1, t_2, ..., t_{32}$ (in particular, the last time point corresponds to Oct 17, 2023). The realized covariance matrices $\mathbf{RC}_t$ represent the proxy of unobservable covolatilities and are estimated as classical sample covariances using 2-minute interval (intraday) data during the given trading date.

Since intraday data necessary for estimating realized covariance matrices are available limitedly, we were able to obtain 32 realized covariance matrices for both indices only. This limitation reflects the constraints associated with acquiring high-frequency financial data, which are typically proprietary and require costly subscriptions. Despite this restriction, we consider the obtained data sufficient for a meaningful evaluation of the suggested procedures within the described framework.

4.2 Comparing criteria and tests

Three criteria, which are based either on numerical metrics or statistical tests, have been used to compare the covolatility prediction results:

(a) *Criterion based on averages of Frobenius norms of deviations from realized covolatilities*:

The metric is constructed as the average Frobenius norm of differences (the average is calculated both over time points $t = t_1, t_2, ..., t_{32}$ (see Section 4.1) and over chosen portfolios $p = 1, \ldots, 100$ of dimension m)

between the conditional matrices $\hat{\mathbf{H}}_t^{(p)}$ predicted by the corresponding MEWMA procedure and the realized covariance matrices $\mathbf{RC}_t$:

$$M^{Frob} = \frac{1}{100}\sum_{p=1}^{100}\left\{\frac{1}{32}\sum_{i=1}^{32}\left\|\hat{\mathbf{H}}_{t_i}^{(p)} - \mathbf{RC}_{t_i}\right\|_F\right\}. \tag{24}$$

The Frobenius norm of a matrix is the square root of the sum of its squared elements.

In fact, this criterion evaluates the one-step prediction's out-of-sample performance compared to the realized covariance matrices $\mathbf{RC}_t$ in summary over portfolios of a given dimension.

(b) *Criterion based on averages of deviations from realized minimum portfolio variances*:

The criterion is constructed as the average difference (the average is again calculated both over time points $t = t_1, t_2, ..., t_{32}$ (see Section 4.1) and over chosen portfolios $p = 1, \ldots, 100$ of dimension m) between the minimum portfolio variance with covariance matrix $\hat{\mathbf{H}}_t^{(p)}$ predicted by the corresponding MEWMA procedure and the minimum portfolio variance with the realized covariance matrix $\mathbf{RC}_t$ estimated similarly as in the previous criterion:

$$M^{minvar} = \frac{1}{100}\sum_{p=1}^{100}\left\{\frac{1}{32}\sum_{i=1}^{32}\left|\frac{1}{\mathbf{1}^{\mathrm{T}}(\hat{\mathbf{H}}_{t_i}^{(p)})^{-1}\mathbf{1}} - \frac{1}{\mathbf{1}^{\mathrm{T}}(\mathbf{RC}_{t_i})^{-1}\mathbf{1}}\right|\right\}, \tag{25}$$

where, e.g., the minimal portfolio variance corresponding to the portfolio construction problem:

$$\min \mathbf{w}_t^{\mathrm{T}}\hat{\mathbf{H}}_t^{(p)}\mathbf{w}_t \quad \text{s.t.} \;\; \mathbf{w}_t^{\mathrm{T}}\mathbf{1} = 1. \tag{26}$$

This criterion evaluates the out-of-sample performance of the minimal portfolio variance calculated by the one-step prediction $\hat{\mathbf{H}}_t^{(p)}$ compared to the minimal portfolio variance calculated by the realized covariance matrices $\mathbf{RC}_t$ in summary over portfolios of a given dimension. The concept of MVP (Minimum Variance Portfolio) is usually used in a similar context (see, e.g., Caldeira et al., 2017).

(c) *Test based on MCS approach*:

A substantial part of the existing literature on covolatility predictions focuses on purely statistical measures of prediction accuracy, the model confidence set (MCS) approach by Hansen et al. (2011). The MCS is used to evaluate the significance of any differences in performance among models. One starts with a complete set of all candidate models and sequentially discards some elements to achieve a smaller set of models. This resulting *model confidence set* contains the best model with a given level of confidence $(1 - \alpha)$ (see, e.g., Becker et al., 2015; Čech and Baruník, 2017).

4.3 Results

For all randomly chosen sub-portfolios from data sets described in Section 4.1, we predicted the conditional covariance matrices $\hat{\mathbf{H}}_t$ over the given period Jan 3, 2018–Oct 17, 2023 using the following MEWMA methods:

1. rec MEWMA: recursive MEWMA; see Formula (17);

2. rec dBEKK-MEWMA: recursive diagonal BEKK MEWMA; see Formulas (19)-(20);

3. rec DCC-MEWMA: recursive DCC MEWMA; see Formulas (21)-(23).

Moreover, several variants of the behavior of forgetting factors in the recursive MEWMA are considered:

– $\alpha_0 = 0.950,\ \tilde{\alpha} = 0.990$;

– $\alpha_0 = 0.950,\ \tilde{\alpha} = 1.000$;

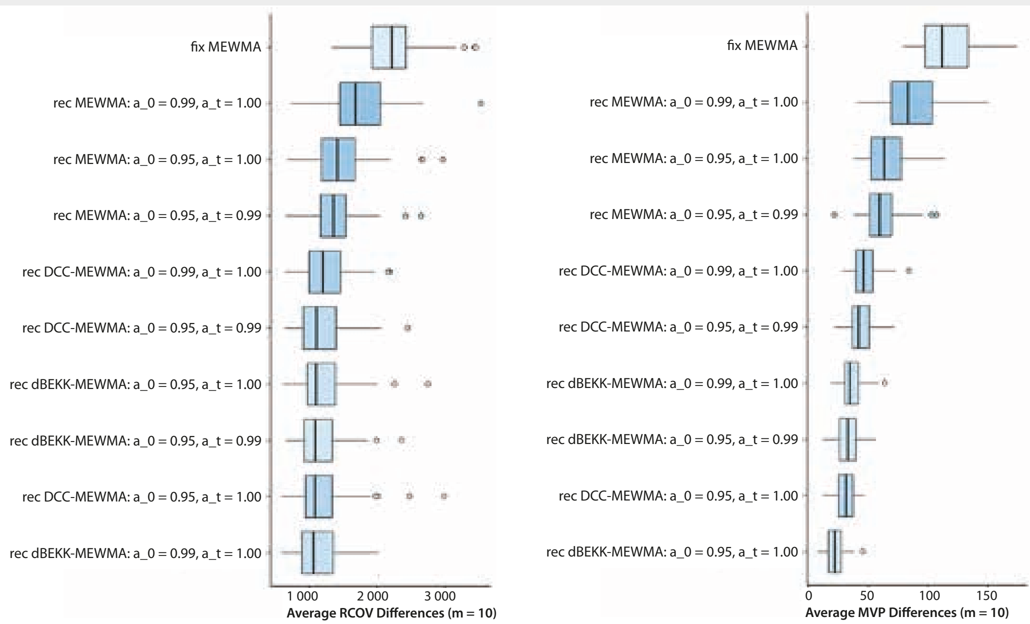
– $\alpha_0 = 0.990,\ \tilde{\alpha} = 1.000$.

Figures 1–4 (for S&P500 data and 100 sub-portfolios of dimensions 10, 30, 50, and 70) and Figures 5–6 (for DJI30 data and 100 sub-portfolios of dimensions 10 and 20) display (a) boxplots of 100 M^{Frob} according to (24) (i.e., averages of Frobenius norms of deviations from realized covolatilities) and (b) boxplots of 100 M^{minvar} according to (25) (i.e., averages of deviations from realized minimum portfolio variances) corresponding to particular modifications of MEWMA.

The graphical results corresponding to the application of metrics M^{Frob} and M^{minvar} can be summarized as follows:

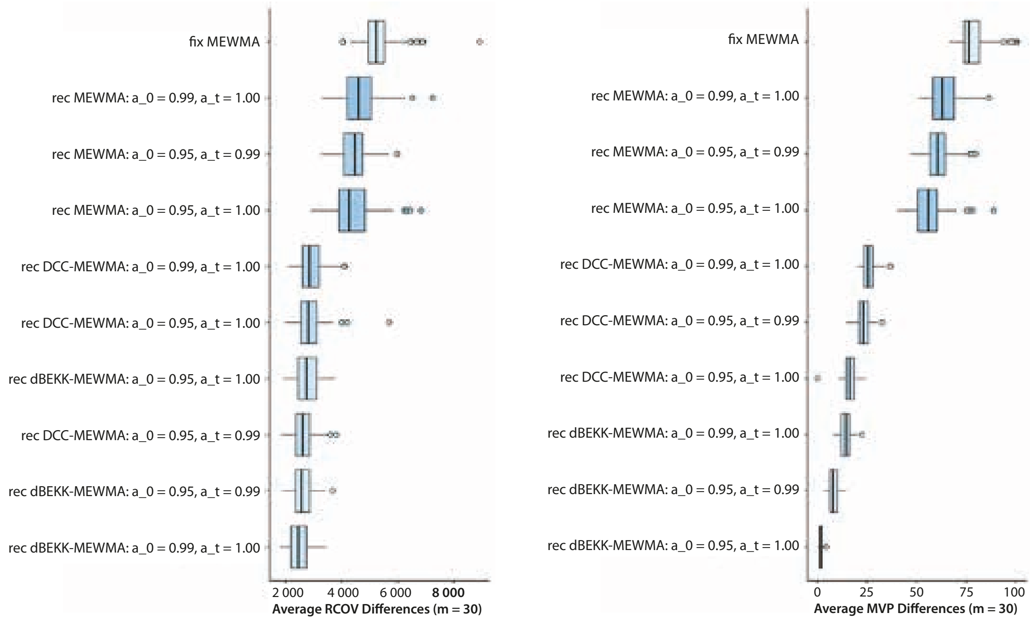
– Suggested modifications of MEWMA with recursive estimation of dynamic coefficients lambda outperform MEWMA with fixed (non-recursively estimated) coefficients. This holds without exception for the data set DJI30. For the dataset S&P500, only the recursive MEWMA model applied for sub-portfolios of dimension $m = 70$ provides comparable results, according to M^{Frob} and M^{minvar}.
– By evaluating M^{Frob} and M^{minvar} simultaneously, the recursive diagonal BEKK MEWMA model dominates over other considered (single) modifications.
– As to the choice of coefficients that influence the behavior of the forgetting factor, the variant $\alpha_0 = 0.990,\ \tilde{\alpha} = 1.000$ and $\alpha_0 = 0.950,\ \tilde{\alpha} = 1.000$ can be preferred by M^{Frob} and M^{minvar}, respectively.

Figure 1 Boxplots of 100 averages of (a) Frobenius norms of deviations from realized covolatilities (left panel) and (b) deviations from realized minimum portfolio variances (right panel) for S&P500 data and $m = 10$ (original values multiplied by 10^6)



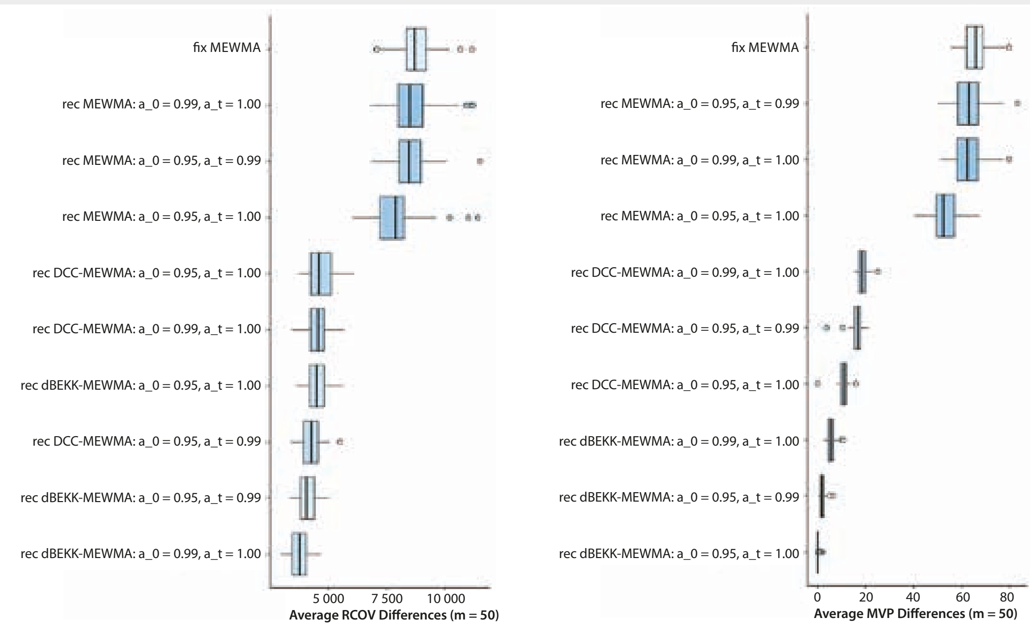
Source: Authors' calculation

Figure 2 Boxplots of 100 averages of (a) Frobenius norms of deviations from realized covolatilities (left panel) and (b) deviations from realized minimum portfolio variances (right panel) for S&P500 data and $m = 30$ (original values multiplied by 10^6)



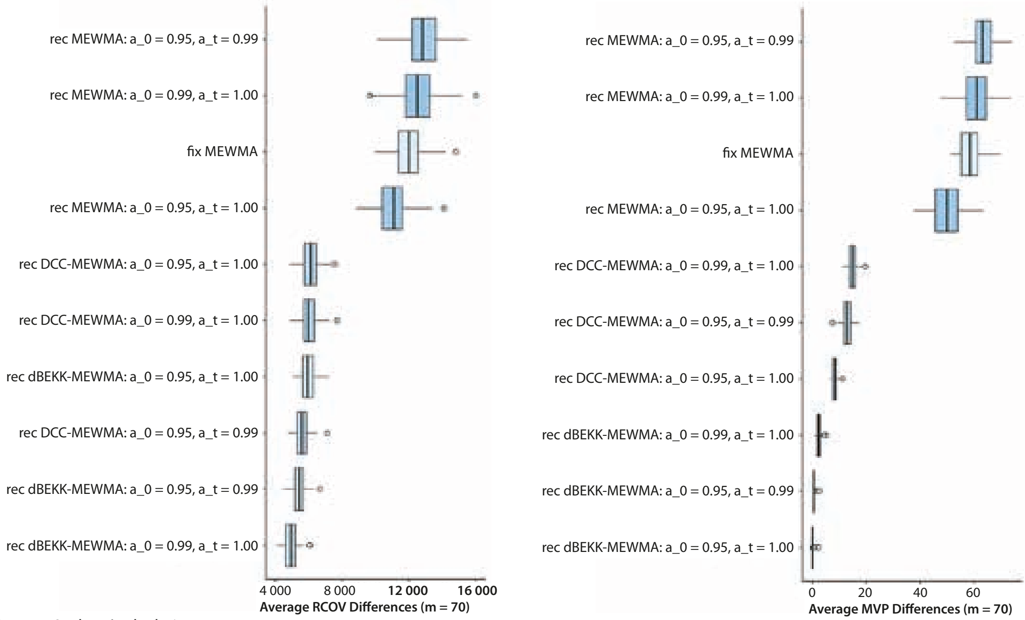
Source: Authors' calculation

Figure 3 Boxplots of 100 averages of (a) Frobenius norms of deviations from realized covolatilities (left panel) and (b) deviations from realized minimum portfolio variances (right panel) for S&P500 data and $m = 50$ (original values multiplied by 10^6)



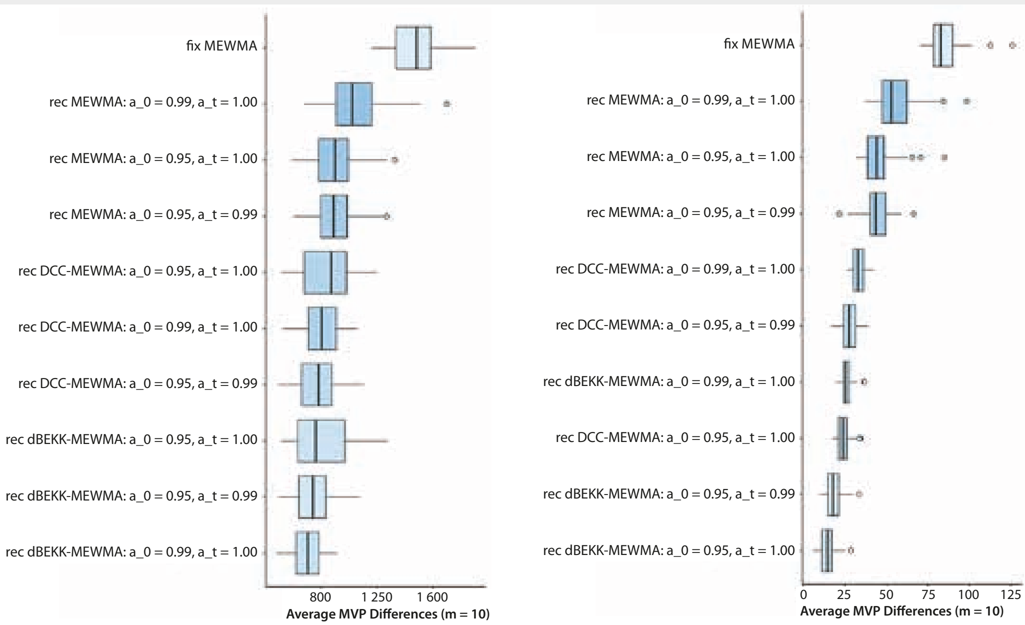
Source: Authors' calculation

Figure 4 Boxplots of 100 averages of (a) Frobenius norms of deviations from realized covolatilities (left panel) and (b) deviations from realized minimum portfolio variances (right panel) for S&P500 data and $m = 70$ (original values multiplied by 10^6)



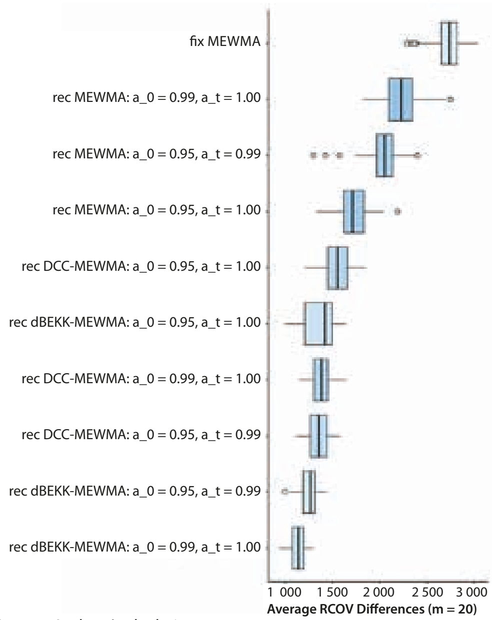
Source: Authors' calculation

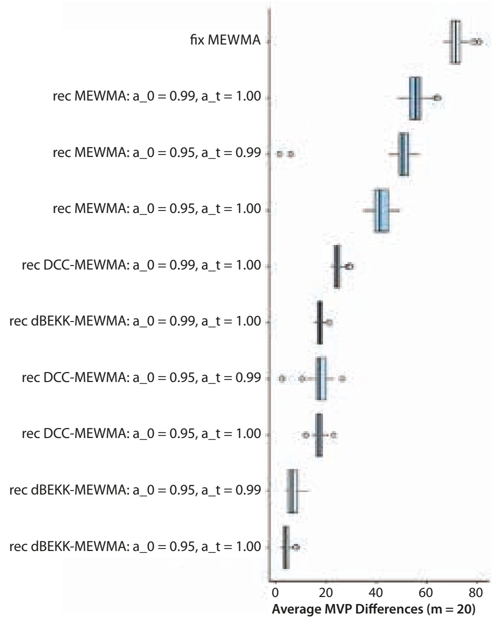
Figure 5 Boxplots of 100 averages of (a) Frobenius norms of deviations from realized covolatilities (left panel) and (b) deviations from realized minimum portfolio variances (right panel) for DJI30 data and $m = 10$ (original values multiplied by 10^6)



Source: Authors' calculation

Figure 6 Boxplots of 100 averages of (a) Frobenius norms of deviations from realized covolatilities (left panel) and (b) deviations from realized minimum portfolio variances (right panel) for DJI30 data and m = 20 (original values multiplied by 10^6)





Source: Authors' calculation

As far as comparison through statistical tests is concerned, Table 1 (for S&P500 data and 100 sub-portfolios of dimensions 10, 30, 50, and 70) and Table 2 (for DJI30 data and 100 sub-portfolios of dimensions 10 and 20) present the results of the MCS test with a confidence level of 95%. Namely, these Tables for each considered modification of MEWMA deliver the counts of achievement of the model confidence set corresponding to modifications of MEWMA.

Table 1 Counts of achievement of the model confidence set MCS with the level of confidence 95% corresponding to modifications of MEWMA over 100 sub-portfolios of S&P500 data

MEWMA	$\alpha_0 = 0.950$ $\tilde{\alpha} = 0.990$	$\alpha_0 = 0.950$ $\tilde{\alpha} = 1.000$	$\alpha_0 = 0.990$ $\tilde{\alpha} = 1.000$
m = 10:			
fix MEWMA	1	1	1
rec MEWMA	→24	19	4
rec dBEKK-MEWMA	↑71	→↑81	↑79
rec DCC-MEWMA	66	→77	41
m = 30:			
fix MEWMA	0	0	0
rec MEWMA	0	0	0
rec dBEKK-MEWMA	↑70	↑64	→↑86
rec DCC-MEWMA	→60	39	26

Table 1 (continuation)

MEWMA	$\alpha_0 = 0.950$ $\tilde{\alpha} = 0.990$	$\alpha_0 = 0.950$ $\tilde{\alpha} = 1.000$	$\alpha_0 = 0.990$ $\tilde{\alpha} = 1.000$
$m = 50$:			
fix MEWMA	0	0	0
rec MEWMA	0	0	0
rec dBEKK-MEWMA	↑64	↑45	→↑90
rec DCC-MEWMA	→38	9	23
$m = 70$:			
fix MEWMA	0	0	0
rec MEWMA	0	0	0
rec dBEKK-MEWMA	↑66	↑42	→↑91
rec DCC-MEWMA	→36	10	13

Note: Vertical arrow ↑ denotes the column maximum and horizontal arrow → indicates the row maximum for particular *m*.
Source: Authors' calculation

Table 2 Counts of achievement of the model confidence set MCS with the level of confidence of 95% corresponding to modifications of MEWMA over 100 sub-portfolios of DJI30 data

MEWMA	$\alpha_0 = 0.950$ $\tilde{\alpha} = 0.990$	$\alpha_0 = 0.950$ $\tilde{\alpha} = 1.000$	$\alpha_0 = 0.990$ $\tilde{\alpha} = 1.000$
$m = 10$:			
fix MEWMA	0	0	0
rec MEWMA	28	→33	5
rec dBEKK-MEWMA	↑90	→↑98	↑85
rec DCC-MEWMA	79	77	65
$m = 20$:			
fix MEWMA	0	0	0
rec MEWMA	1	1	0
rec dBEKK-MEWMA	→↑ 100	→↑ 100	→↑ 100
rec DCC-MEWMA	54	34	→78

Note: Vertical arrow ↑ denotes the column maximum and horizontal arrow → indicates the row maximum for particular *m*.
Source: Authors' calculation

The results corresponding to the application of the MCS set can be summarized as follows:

- All suggested modifications of MEWMA with recursive estimation of dynamic coefficients lambda outperform MEWMA with fixed (non-recursive) coefficients even more significantly than in the case of numerical metrics in Figures 1–6.
- The same conclusion holds when one compares various (single) modifications of the recursive MEWMA model. In all cases, the recursive diagonal BEKK MEWMA model dominates significantly over other considered (single) modifications (only in a limited number of results in Table 1, the results by recursive DCC MEWMA are comparable).
- As to the choice of coefficients that influence the behavior of the forgetting factor, the choice $\alpha_0 = 0.950$, $\tilde{\alpha} = 1.000$ seems to be again the best in the prevailing number of cases.

4.4 Extended benchmark

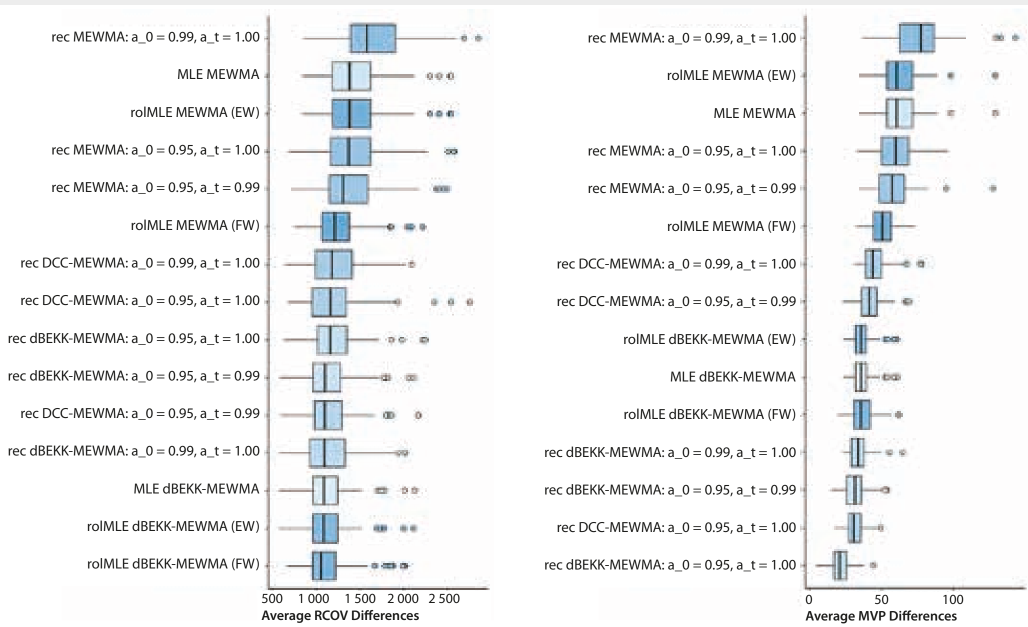
To compare the proposed recursive methods with more advanced benchmarks, we have conducted an extended simulation study for S&P500 data and 100 sub-portfolios of dimensions 10. Specifically, we compare our recursive approaches against a broader set of methods that include:

- *MLE MEWMA:* MEWMA estimated by maximum likelihood estimation (MLE),
- *rolMLE MEWMA (EW):* MEWMA estimated by rolling MLE (expanding window),
- *rolMLE MEWMA (FW):* MEWMA estimated by rolling MLE (fixed-width window of 100 observations),
- *MLE dBEKK-MEWMA:* diagonal BEKK MEWMA estimated by MLE,
- *rolMLE dBEKK-MEWMA (EW):* diagonal BEKK MEWMA estimated by rolling MLE (expanding window),
- *rolMLE dBEKK-MEWMA (FW):* diagonal BEKK MEWMA estimated by rolling MLE (fixed-width window of 100 observations).

The performance results are summarized in Figure 7. It displays (a) boxplots of 100 M^{Frob} according to (24) (i.e., averages of Frobenius norms of deviations from realized covolatilities) and (b) boxplots of 100 M^{minvar} according to (25) (i.e., averages of deviations from realized minimum portfolio variances) corresponding to particular techniques. The evaluation criteria based on the M^{Frob} and M^{minvar} measures provide the following conclusions:

- In terms of M^{Frob}, the non-recursive diagonal BEKK MEWMA methods listed above slightly outperform the recursive diagonal BEKK MEWMA approach.
- In terms of M^{minvar}, the recursive diagonal BEKK MEWMA method performs better than its non-recursive counterparts.

Figure 7 Boxplots of 100 averages of (a) Frobenius norms of deviations from realized covolatilities (left panel) and (b) deviations from realized minimum portfolio variances (right panel) for S&P500 data, $m = 10$ and an extended benchmark (original values multiplied by 10^6)

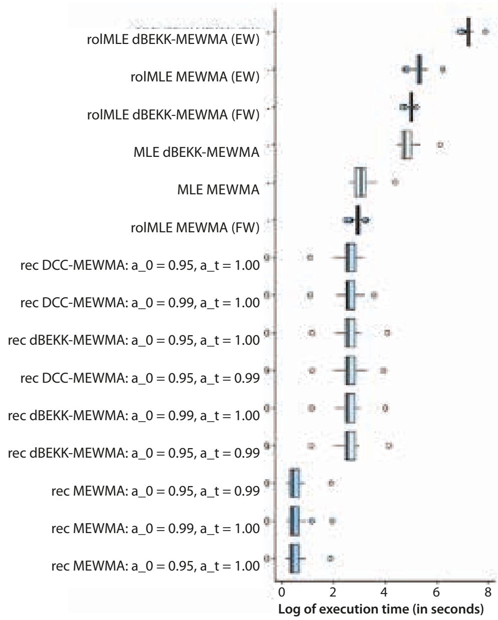


Source: Authors' calculation

A key consideration when recursive techniques are studied is computational efficiency. As shown in Figure 8, that presents the logarithm of the computational time in seconds required for one simulation per method. Our results demonstrate that recursive methods are significantly more efficient compared to the non-recursive approaches, particularly those relying on re-estimation of lambda via MLE in each step.

Given the increased computational complexity of the herein assumed non-recursive methods, we restricted our extended simulation comparison to the S&P 500 dataset with a 10-asset portfolio only.

Figure 8 Execution time in seconds per 1 simulation for S&P500 data, m = 10 and an extended benchmark (log scale)



Source: Authors' calculation

Expanding the analysis further to higher-dimensional settings would cause substantial computational challenges.

These additional results justify our arguments that recursive estimation provides a competitive trade-off between accuracy and computational feasibility, and make it be a practical choice for large-scale applications.

CONCLUSION

The MEWMA procedure is a numerically simple method recommended for projections of conditional covariances (covolatilities), mainly if the dimensions of conditional covariance matrices are large. Nevertheless, the paper shows that it is worthwhile to approach the MEWMA more dynamically and estimate its parameter(s) recursively in time, as such strategies improve performance indicators.

Moreover, a substantial improvement is achieved when each component of the corresponding multivariate process is handled by applying the separate EWMA procedure with the corresponding parameter also estimated separately (but still recursively in time). In particular, the diagonal BEKK recursive MEWMA model can be recommended in this context.

There are suggestions for further research, e.g., to combine various candidate models to improve the overall performance, to combine recursive and iterative approach to estimation and to robustify suggested recursive methods. Moreover, the covolatility projection problem should be examined from the economic point of view in the context of portfolio optimization and reported in a more economically oriented context.

References

BAUWENS, L., LAURENT, S., ROMBOUTS, J. (2006). Multivariate GARCH Models: a Survey [online]. *Journal of Applied Econometrics*, 21: 79–109. <https://doi.org/10.2139/ssrn.411062>.

BECKER, R., CLEMENTS, A. E., DOOLAN, M. B., HURN, A. S. (2015). Selecting Volatility Forecasting Models for Portfolio Allocation Purposes [online]. *International Journal of Forecasting*, 31: 849–861. <https://doi.org/10.1016/j.ijforecast.2013.11.007>.

CALDEIRA, J. F., MOURA, G. V., NOGALES, F. J., SANTOS, A. A. P. (2017). Combining Multivariate Volatility Forecasts: an Economic-Based Approach [online]. *Journal of Financial Econometrics*, 15: 247–85. <https://doi.org/10.2139/ssrn.2664128>.

CAPORIN, M., McALEER, M. (2013). Ten Things You Should Know about the Dynamic Conditional Correlation Representation [online]. *Econometrics*, 115–126. <https://doi.org/10.3390/econometrics1010115>.

CHIRIAC, R., VOEV, V. (2011). Modeling and Forecasting Multivariate Realized Volatility [online]. *Journal of Applied Econometrics*, 2011, 26, pp. 922–947. <https://doi.org/10.2307/23018257>.

CIPRA, T. (2020). *Time Series in Economics and Finance*. Berlin: Springer.

CIPRA, T., HENDRYCH, R. (2019). Modelling of Currency Covolatilities [online]. *Statistika: Statistics and Economy Journal*, 99(3): 259–271. <https://csu.gov.cz/docs/107508/a6a9d90c-259a-1aca-79f8-798b52fa26af/32019719q3_259_cipra_analyses.pdf?version=1.0>.

CLEMENTS, A., SCOTT, A., SILVENNOINEN, A. (2012). Forecasting Multivariate Volatility in Larger Dimensions: Some Practical Issues. *Working Paper 80*, NCER, Queensland University of Technology.

CREAL, D., KOOPMAN, S. J., LUCAS, A. (2013). Generalized Autoregressive Score Models with Applications [online]. *Journal of Applied Econometrics*, 28: 777–795. <https://doi.org/10.1002/jae.1279>.

ČECH, F., BARUNÍK, J. (2017). On the Modelling and Forecasting of Multivariate Realized Volatility: Generalized Heterogeneous Autoregressive (GHAR) Model [online]. *Journal of Forecasting*, 36: 181–206. <https://doi.org/10.1002/for.2423>.

ENGLE, R. F., COLACITO, R. (2006). Testing and Valuing Dynamic Correlations for Asset Allocation [online]. *Journal of Business & Economic Statistics*, 24: 238–253. <https://doi.org/10.1198/073500106000000017>.

ENGLE, R. F., SHEPHARD, N., SHEPPARD, K. (2008). Fitting Vast Dimensional Time-Varying Covariance Models [online]. *Journal of Business & Economic Statistics*, 39(3): 1–39. <https://doi.org/10.1080/07350015.2020.1713795>.

ENGLE, R. F., SHEPPARD, K. (2007). *Evaluating the Specification of Covariance Models for Large Portfolios*. Working Paper, New York: New York University.

HAFNER, C. M., FRANSES, P. H. (2009). A Generalized Dynamic Conditional Correlation Model: Simulation and Application to Many Assets [online]. *Econometric Reviews*, 28: 612–631. <https://doi.org/10.1080/07474930903038834>.

HANSEN, P. R., LUNDE, A., NASON, J. M. (2011). The Model Confidence Set. *Econometrica*, 79: 453–497.

HENDRYCH, R., CIPRA, T. (2016). On Conditional Covariance Modelling: an Approach Using State Space Models [online]. *Computational Statistics & Data Analysis*, 100: 304–317. <https://doi.org/10.1016/j.csda.2014.09.019>.

HENDRYCH, R., CIPRA, T. (2018). Self-Weighted Recursive Estimation of GARCH Models [online]. *Communications in Statistics – Simulation and Computation*, 47: 315–328. <https://doi.org/10.1080/03610918.2015.1053924>.

HENDRYCH, R., CIPRA, T. (2019). Recursive Estimation of EWMA Model [online]. *Journal of Risk*, 21: 43–67. <https://doi.org/10.21314/JOR.2019.413>.

LAURENT, S., ROMBOUTS, J. V. K., VIOLANTE, F. (2012). On the Forecasting Accuracy of Multivariate GARCH Models [online]. *Journal of Applied Econometrics*, 27: 934–955. <https://doi.org/10.2307/23355908>.

LEDOIT, O., WOLF, M. (2004). A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices [online]. *Journal of Multivariate Analysis*, 88: 365–411. <https://doi.org/10.1016/S0047259X(03)00096-4>.

LJUNG, L., SÖDERSSTRÖM, T. (1983). *Theory and Practice of Recursive Identification*. Cambridge, MA: MIT Press.

RISKMETRICS. (1996). *Technical Document*. New York, NY: JP Morgan/Reuters.

APPENDIX: DERIVATION OF RECURSIVE ALGORITHM (12)-(16)

This algorithm is based on relations (5)-(9) and (11) for $\theta = \lambda$. It suffices to derive the formulas for $\mathbf{F}_t'(\hat{\lambda}_{t-1})$ and $\widetilde{\mathbf{F}}_t''(\hat{\lambda}_{t-1})$, where:

$$F_t(\lambda) = \ln\left|\mathbf{H}_t(\lambda)\right| + \boldsymbol{r}_t^{\mathrm{T}}\mathbf{H}_t(\lambda)^{-1}\boldsymbol{r}_t. \tag{A1}$$

Using the matrix differential calculus for a general regular matrix $\mathbf{U}(x)$, namely:

$$\frac{\partial \ln\left|\mathbf{U}(x)\right|}{\partial x} = \mathrm{tr}\left(\mathbf{U}^{-1}(x)\frac{\partial \mathbf{U}(x)}{\partial x}\right), \quad \frac{\partial \mathbf{U}^{-1}(x)}{\partial x} = -\mathbf{U}^{-1}(x)\frac{\partial \mathbf{U}(x)}{\partial x}\mathbf{U}^{-1}(x), \tag{A2}$$

one can write:

$$\mathbf{F}_t'(\lambda) = \mathrm{tr}\left(\mathbf{H}_t^{-1} \frac{\partial \mathbf{H}_t}{\partial \lambda} \right) - \boldsymbol{r}_t^{\mathrm{T}} \mathbf{H}_t^{-1} \frac{\partial \mathbf{H}_t}{\partial \lambda} \mathbf{H}_t^{-1} \boldsymbol{r}_t \,, \tag{A3}$$

so that (12) follows substituting (A3) to (5). Further, it holds:

$$\mathbf{F}_t''(\lambda) = \mathrm{tr}\left(-\mathbf{H}_t^{-1} \frac{\partial \mathbf{H}_t}{\partial \lambda} \mathbf{H}_t^{-1} \frac{\partial \mathbf{H}_t}{\partial \lambda} + \mathbf{H}_t^{-1} \frac{\partial^2 \mathbf{H}_t}{\partial \lambda^2} \right) + 2\boldsymbol{r}_t^{\mathrm{T}} \mathbf{H}_t^{-1} \frac{\partial \mathbf{H}_t}{\partial \lambda} \mathbf{H}_t^{-1} \frac{\partial \mathbf{H}_t}{\partial \lambda} \mathbf{H}_t^{-1} \boldsymbol{r}_t$$
$$- \boldsymbol{r}_t^{\mathrm{T}} \mathbf{H}_t^{-1} \frac{\partial^2 \mathbf{H}_t}{\partial \lambda^2} \mathbf{H}_t^{-1} \boldsymbol{r}_t, \tag{A4}$$

which can be approximated by:

$$\widetilde{\mathbf{F}}_t''(\lambda) = \mathrm{tr}\left(\mathbf{H}_t^{-1} \frac{\partial \mathbf{H}_t}{\partial \lambda} \mathbf{H}_t^{-1} \frac{\partial \mathbf{H}_t}{\partial \lambda} \right). \tag{A5}$$

This approximation fulfills the condition:

$$\mathrm{E}\left(\mathbf{F}_t''(\hat{\lambda}_{t-1}) - \widetilde{\mathbf{F}}_t''(\hat{\lambda}_{t-1}) \middle| \Omega_{t-1} \right) = 0\,, \tag{A6}$$

see Formula (9), since it holds:

$$\mathrm{E}\left[2\boldsymbol{r}_t^{\mathrm{T}} \mathbf{H}_t^{-1} \frac{\partial \mathbf{H}_t}{\partial \lambda} \mathbf{H}_t^{-1} \frac{\partial \mathbf{H}_t}{\partial \lambda} \mathbf{H}_t^{-1} \boldsymbol{r}_t - \boldsymbol{r}_t^{\mathrm{T}} \mathbf{H}_t^{-1} \frac{\partial^2 \mathbf{H}_t}{\partial \lambda^2} \mathbf{H}_t^{-1} \boldsymbol{r}_t \,\middle|\, \Omega_{t-1} \right] =$$
$$= \mathrm{tr}\left(2\mathbf{H}_t^{-1} \frac{\partial \mathbf{H}_t}{\partial \lambda} \mathbf{H}_t^{-1} \frac{\partial \mathbf{H}_t}{\partial \lambda} - \mathbf{H}_t^{-1} \frac{\partial^2 \mathbf{H}_t}{\partial \lambda^2} \right). \tag{A7}$$