Forecasting portfolio-Value-at-Risk with mixed factorial hidden Markov models

Mohamed Saidane[*]

Department of Management Information Systems and Production Management
College of Business and Economics, Qassim University
P.O. Box 6666 - Buraidah: 51452, Kingdom of Saudi Arabia
E-mail: ⟨M.Saidane@qu.edu.sa⟩

Abstract. This paper is concerned with the statistical modeling of the latent dependence and co-movement structures of multivariate financial data using a new approach based on mixed factorial hidden Markov models, and their applications in Value-at-Risk (VaR) valuation. This approach combines hidden Markov Models (HMM) with mixed latent factor models. The HMM generates a piece-wise constant state evolution process and the observations are produced from the state vectors by a mixture of factor analyzers observation process. This new switching specification provides an alternative, compact, model to handle intra-frame correlation and unobserved heterogeneity in financial data. For maximum likelihood estimation we have proposed an iterative approach based on the Expectation-Maximisation (EM) algorithm. Using a set of historical data, from the Tunisian foreign exchange market, the model parameters are estimated. Then, the fitted model combined with a modified Monte-Carlo simulation algorithm was used to predict the VaR of the Tunisian public debt portfolio. Through a backtesting procedure, we found that this new specification exhibits a good fit to the data, improves the accuracy of VaR predictions and can avoid serious violations when a financial crisis occurs.

Keywords: mixed latent factor models, hidden Markov models, unobserved heterogeneity, EM algorithm, Value-at-Risk

---

1. Introduction

The Value-at-Risk (VaR) is one of the main risk indicators for management of financial portfolios (see [5]). It is defined as the money-loss in a portfolio that is expected to occur over a pre-determined horizon and with a pre-determined degree of confidence. More precisely, it is the threshold above which a loss over a chosen time horizon occurs with at most a given level of confidence.

Calculation of portfolio VaR is often based on some assumptions, such as returns follow a conditional normal distribution and a diffusion process. Vast academic literatures indicate that the unconditional return distribution of financial assets is non-normal and tends to have fat tails and high peaks. Besides, the returns are often characterized by a number of stylized facts, and the risk profile of an asset does not remain constant overtime. A variety of systematic and unsystematic events may change financial risk of the asset significantly. In particular, when financial markets are subjected to a major regime shift, many volatilities and correlations can be expected to shift as well, perhaps substantially. These shifts in risk are not predictable; they should be regarded as random events. Also, the effect of these shifts should be taken

---

[*]Corresponding author.

into account by risk managers in the assessment of market risk and capital allocation and by regulators in the definition of capital requirements.

The purpose of this paper is to resolve the puzzle by considering a new multivariate regime switching approach for portfolio VaR estimation. This approach combines hidden Markov Models (HMM) with mixed latent factor models, MFA (see [8]). Thus, a natural generalization of the standard MFA to a multi-state model is achieved by allowing for model transitions that are governed by a hidden Markov chain on a set of possible models that describe the different states of volatility.

This paper is organized as follows. In the next section, the proposed model is presented, followed by a description of a maximum likelihood estimation procedure, based on the EM algorithm. Section 3 develops a modified Monte-Carlo simulation algorithm for calculating portfolio VaR within a mixed factorial HMM framework, which constitutes the major contribution of this paper. To see the advantage of this method, we compare in section 4 our results with VaR results obtained from other benchmark methods. In this work the assessment of the exchange rate risk, associated with the Tunisian public debt portfolio, through the VaR methodology is considered as the basis for an application to our model. Section 5 concludes this article.

2. The mixed factorial hidden Markov model

It is well known that the return series of different assets are correlated with each other, i.e. the assets follow common influences on their returns. This can be used to reduce the number of parameters to be estimated. In the financial literature, various forms of factor models such as the CAPM and the APT are often used. In general, factor models postulate that the return of an asset is composed as the sum of an expected and an unexpected part. The unexpected part of the return is assumed to consist of a systematic portion which cannot be diversified and an unsystematic portion which is specific to the single asset.

Financial theory states that there are common influences such as macroeconomic data which drive the returns of different assets. These are known factors. The systematic unexpected part of the return is assumed to follow a factor structure. Furthermore, financial theory also states that the assumption of normal probability distribution belongs to the biggest imperfections of estimating VaR. In point of fact, the returns of financial time series are rather distributed leptokurtic than normally. Moreover, the empirical distributions are often skewed. In these cases, the assumption of normal distribution results in over/or underestimation of VaR especially when the quantiles are very high/low. Therefore it is necessary to put emphasis on respecting the leptokurtic and skewed return distribution. To take into account all these assumptions and the possibility of regime switching in stock market returns, we propose a model that combines:

1. an HMM structure in order to take into account different states of the world that can affect the evolution of the time series, and

2. a mixture of probabilistic factor analyzers with constant regime parameters to take into account the unobserved heterogeneity and the assumptions of leptokurtic and skewed return distributions.

The use of this new specification solves various problems related to the changes of the internal and unobservable structure of financial data. These are problems of the type:

(i) Can we distinguish different heterogeneous regimes in stock market returns?

(ii) How do the regimes differ?

(iii) How frequent are regime switches and when do they occur?

(iv) What is the impact of a particular regime on the common and specific variances?

(v) Are regime switches predictable?

2.1. Basic model and factor structure

Our proposed model is a dynamic state space generalization of a multiple component factor analysis system. The k-dimensional state vectors are generated by a standard diagonal co-variance Gaussian mixture HMM. The q-dimensional observation vectors are generated by a multiple noise component factor analysis observation process. A generative model for our mixed factorial HMM can be described as follows:

$$y_t = A_{s_t} f_t + \varepsilon_{s_t} \tag{1}$$

where $\forall\, t = 1, ..., T$ and $\forall\, i, j = 1, ..., m$

$$S_t \sim P(S_t = j | S_{t-1} = i) \tag{2}$$

is an homogenous hidden Markov chain indicating the state or the regime at the date t. The common latent factor vector is given by

$$f_t = w_{s_t} \tag{3}$$

where

$$w_j \sim \sum_n c_{jn}^{(f)} \mathcal{N}\left(\theta_{jn}^{(f)}, \Sigma_{jn}^{(f)}\right) \tag{4}$$

and the vector of specific factors is as follow

$$\varepsilon_j \sim \sum_m c_{jm}^{(y)} \mathcal{N}\left(\theta_{jm}^{(y)}, \Sigma_{jm}^{(y)}\right) \tag{5}$$

The HMM state transition probabilities from state i to state j are represented by p_{ij} and the state and observation space mixture distributions are described by the mixture weights $c_{jn}^{(f)}$, $c_{jm}^{(y)}$; the mean vectors $\theta_{jn}^{(f)}$, $\theta_{jm}^{(y)}$ and the diagonal variance-covariance matrices $\Sigma_{jn}^{(f)}$ and $\Sigma_{jm}^{(y)}$.

2.2. Likelihood Calculation

An important aspect of any generative model is the complexity of the likelihood calculations. The generative model in equations (1-5) can be expressed by the following two Gaussian distributions

$$p(f_t | S_t = j, w_t^{(f)} = n) = \mathcal{N}\left(\theta_{jn}^{(f)}, \Sigma_{jn}^{(f)}\right) \tag{6}$$

$$p(y_t | f_t, S_t = j, w_t^{(y)} = m) = \mathcal{N}\left(A_j f_t + \theta_{jm}^{(y)}, \Sigma_{jm}^{(y)}\right) \tag{7}$$

The likelihood of an observation y_t given the state $S_t = j$, state space component $w_t^{(f)} = n$ and observation noise component $w_t^{(y)} = m$ can be obtained by integrating the state vector f_t out of the product of the above Gaussians. The resulting likelihood is also a Gaussian and can be written as

$$b_{jmn}(y_t) = p(y_t|S_t = j, w_t^{(y)} = m, w_t^{(f)} = n) = \mathcal{N}\left(\theta_{jmn}, \Sigma_{jmn}\right) \tag{8}$$

where

$$\theta_{jmn} = A_j \theta_{jn}^{(f)} + \theta_{jm}^{(y)} \tag{9}$$

and

$$\Sigma_{jmn} = A_j \Sigma_{jn}^{(f)} A_j' + \Sigma_{jm}^{(y)} \tag{10}$$

Hence, the conditional observation density of state j factorial HMM can be viewed as an $M^{(y)}M^{(f)}$ component full covariance matrix GMM with mean vectors given by equation 9 and covariance matrices given by equation 10.

2.3. Parameter Optimization

The maximum likelihood criterion may be used to optimize the parameters of the mixed factorial HMM. In common with standard HMM training described in [11] the EM algorithm is used.[‡] The auxiliary function for our proposed model can be written as

$$\mathcal{Q}(\Theta, \Theta^{[e]}) = \sum_{\forall S} \int p(S|\mathcal{Y}, \Theta^{[e]}) p(\mathcal{F}|\mathcal{Y}, S, \Theta^{[e]}) \log p(\mathcal{Y}, \mathcal{F}, S|\Theta) d\mathcal{F} \tag{11}$$

where all the possible discrete state and continuous state sequences of length T are included in the sum and the integral. A sequence of observation vectors is denoted by $\mathcal{Y} = \{y_1, ..., y_T\}$, $S = \{S_1, ..., S_T\}$ is a sequence of the HMM states, and $\mathcal{F} = \{f_1, ..., f_T\}$ is a sequence of latent factors. The set of current model parameters is represented by $\Theta^{[e]}$.

Sufficient statistics for the first term, $p(S|\mathcal{Y}, \Theta^{[e]})$, in the auxiliary function in equation 11 can be obtained using the standard forward-backward algorithm described in [11] with likelihoods given by equation 10. For the state transition probability optimization, two sets of sufficient statistics are needed, the posterior probabilities of being in state j at time t, $\gamma_j(t) = p(S_t = j|\mathcal{Y}, \Theta^{[e]})$, and being in state i at time $t-1$ and in state j at time t, $\xi_{ij}(t) = p(S_{t-1} = i, S_t = j|\mathcal{Y}, \Theta^{[e]})$. For the state conditional observation density parameter optimization, the component posteriors, $\gamma_{jmn}(t) = p(S_t = j, w_t^{(y)} = m, w_t^{(f)} = n|\mathcal{Y}, \Theta^{[e]})$, have to be estimated. These can be obtained within the forward-backward algorithm as follows:

$$\gamma_{jmn}(t) = \frac{1}{p(\mathcal{Y}|\Theta^{[e]})} \sum_{i=1}^{N_s} p_{ij} \alpha_i(t-1) c_{jm}^{(y)} c_{jn}^{(f)} b_{jmn}(y_t) \beta_j(t) \tag{12}$$

where N_s is the number of HMM states in the model, $\alpha_i(t-1)$ is the standard forward and $\beta_j(t)$ is the standard backward variable defined for HMMs in [11].

The second term, $p(\mathcal{F}|\mathcal{Y}, S, \Theta^{[e]})$, in the auxiliary function in equation 11 is the distribution of the vector of latent factors given the observation sequence and the HMM state sequence. Only the first and second-order statistics are required since the distributions are conditionally Gaussian given the state and the mixture components. Using the conditional independence assumptions made in the model, the posterior can be expressed as

$$p(f_t|y_t, S_t = j, w_t^{(y)} = m, w_t^{(f)} = n) = \frac{p(y_t, f_t|S_t = j, w_t^{(y)} = m, w_t^{(f)} = n)}{p(y_t|S_t = j, w_t^{(y)} = m, w_t^{(f)} = n)} \tag{13}$$

[‡]For further details about the EM algorithm, see [2] or [7].

which using equations 6, 7 and 8 simplifies to a Gaussian distribution with mean vector, $\widehat{f}_{jmnt}$, and correlation matrix, $\widehat{R}_{jmnt}$, defined by

$$\widehat{f}_{jmnt} = \theta_{jn}^{(f)} + K_{jmn}\left(y_t - A_j\theta_{jn}^{(f)} - \theta_{jm}^{(y)}\right) \tag{14}$$

$$\widehat{R}_{jmnt} = \Sigma_{jn}^{(f)} - K_{jmn}A_j\Sigma_{jn}^{(f)} + \widehat{f}_{jmnt}\widehat{f}_{jmnt}' \tag{15}$$

where $K_{jmn} = \Sigma_{jn}^{(f)}A_j'\left(A_j\Sigma_{jn}^{(f)}A_j' + \Sigma_{jm}^{(y)}\right)^{-1}$. Given the two sets of sufficient statistics above, the model parameters can be optimized by solving a standard maximization problem. The parameter update formulae for the underlying HMM parameters in mixed factorial HMMs are very similar to those for the standard HMM (see [10]) except the above state vector distribution statistics replace the observation sample moments. Omitting the state transition probabilities, the state space parameter update formulae can be written as

$$\widehat{c}_{jn}^{(f)} = \frac{\sum\limits_{t=1}^{T}\sum\limits_{m=1}^{M^{(y)}}\gamma_{jmn}(t)}{\sum\limits_{t=1}^{T}\gamma_j(t)} \tag{16}$$

$$\widehat{\theta}_{jn}^{(f)} = \frac{\sum\limits_{t=1}^{T}\sum\limits_{m=1}^{M^{(y)}}\gamma_{jmn}(t)\widehat{f}_{jmnt}}{\sum\limits_{t=1}^{T}\sum\limits_{m=1}^{M^{(y)}}\gamma_{jmn}(t)} \tag{17}$$

$$\widehat{\Sigma}_{jn}^{(f)} = diag\left\{\frac{\sum\limits_{t=1}^{T}\sum\limits_{m=1}^{M^{(y)}}\gamma_{jmn}(t)\widehat{R}_{jmnt}}{\sum\limits_{t=1}^{T}\sum\limits_{m=1}^{M^{(y)}}\gamma_{jmn}(t)} - \widehat{\theta}_{jn}^{(f)}\widehat{\theta}_{jn}^{(f)'}\right\} \tag{18}$$

where $diag(.)$ sets all the off-diagonal elements of the matrix argument to zeros. The cross terms including the new state space mean vectors and the first-order accumulates have been simplified in equation 18. This can only be done if the mean vectors are updated during the same iteration, and the covariance matrices and the mean vectors are tied on the same level.

The new observation matrix, $\widehat{A}_j$, has to be optimized row by row as in [4]. The l-th row vector $\widehat{a}_{jl}$ of the new observation matrix can be written as

$$\widehat{a}_{jl} = k_{jl}'G_{jl}^{-1} \tag{19}$$

where the k by k matrix G_{jl} and the k-dimensional column vector k_{jl} are defined as follows

$$G_{jl} = \sum_{m=1}^{M^{(y)}}\frac{1}{\sigma_{jml}^{(y)2}}\sum_{t=1}^{T}\sum_{n=1}^{M^{(f)}}\gamma_{jmn}(t)\widehat{R}_{jmnt} \tag{20}$$

$$k_{jl} = \sum_{m=1}^{M^{(y)}}\frac{1}{\sigma_{jml}^{(y)2}}\sum_{t=1}^{T}\sum_{n=1}^{M^{(f)}}\gamma_{jmn}(t)\left(y_{tl} - \theta_{jml}^{(y)}\right)\widehat{f}_{jmnt} \tag{21}$$

where $\sigma_{jml}^{(y)2}$ is the l-th diagonal element of the observation covariance matrix $\Sigma_{jm}^{(y)}$, y_{tl} and $\theta_{jml}^{(y)}$ are the l-th elements of the current observation and the observation noise mean vectors, respectively.

Given the new observation matrix, the observation noise parameters can be optimized using the following formulae

$$\widehat{c}_{jm}^{(y)} = \frac{\sum\limits_{t=1}^{T}\sum\limits_{n=1}^{M^{(f)}}\gamma_{jmn}(t)}{\sum\limits_{t=1}^{T}\gamma_j(t)} \tag{22}$$

$$\widehat{\theta}_{jm}^{(y)} = \frac{\sum\limits_{t=1}^{T}\sum\limits_{n=1}^{M^{(f)}}\gamma_{jmn}(t)\left(y_t - \widehat{A}_j\widehat{f}_{jmnt}\right)}{\sum\limits_{t=1}^{T}\sum\limits_{n=1}^{M^{(f)}}\gamma_{jmn}(t)} \tag{23}$$

$$\widehat{\Sigma}_{jm}^{(y)} = \frac{\sum\limits_{t=1}^{T}\sum\limits_{n=1}^{M^{(f)}}\gamma_{jmn}(t)diag\left\{y_t y_t' - \Delta_{jmnt} - \Delta_{jmnt}' + \Omega_{jmnt}\right\}}{\sum\limits_{t=1}^{T}\sum\limits_{n=1}^{M^{(f)}}\gamma_{jmn}(t)} \tag{24}$$

where

$$\Delta_{jmnt} = \left[\ \widehat{A}_j\ \widehat{\theta}_{jm}^{(y)}\ \right]\left[\begin{array}{c}\widehat{f}_{jmnt}y_t' \\ y_t'\end{array}\right]$$

and

$$\Omega_{jmnt} = \left[\ \widehat{A}_j\ \widehat{\theta}_{jm}^{(y)}\ \right]\left[\begin{array}{cc}\widehat{R}_{jmnt}\ \widehat{f}_{jmnt} \\ \widehat{f}_{jmnt}'\ \ \ 1\end{array}\right]\left[\begin{array}{c}\widehat{A}_j' \\ \widehat{\theta}_{jm}^{(y)'}\end{array}\right]$$

3. Value-at-Risk Computation

To estimate the mixed factorial HMM VaR, we will use a Monte-Carlo simulation approach. This method is an accurate way to calculate the VaR measurement for all financial instruments linear and nonlinear. Monte-Carlo estimates the VaR by simulating random scenarios based on the predicted hidden states obtained from the mixed factorial HMM.

3.1. Making Predictions with mixed factorial HMM

In this paper, the prediction process is carried out together with the training at the same time. In other words, at the end of each trading day, when the market return of that day is known to the public, we include it in the training system to absorb the most updated information. This is exactly what human analysts and professional traders do. The mixed factorial HMM is then trained again with the newest time series and the parameters are updated.

When the training is done and parameters of the mixed factorial HMM are updated, the optimal forecast of the next state S_{t+1} (based on the information set available at time t), can be obtained as follows:

$$\widehat{S}_{t+1|t} = \arg\max_j p(S_{t+1} = j|\mathcal{Y}_{1:t};\Theta), \quad 1 \le t \le n-1$$

where

$$p(S_{t+1} = j | \mathcal{Y}_{1:t}; \Theta) = \sum_{i=1}^{m} P(S_{t+1} = j, S_t = i | \mathcal{Y}_{1:t})$$

$$= \frac{\sum_{i=1}^{m} p(\mathcal{Y}_{1:t}, S_t = i) p(S_{t+1} = j | S_t = i)}{p(\mathcal{Y}_{1:t})} = \frac{\sum_{i=1}^{m} \alpha_t(i) p_{ij}}{\sum_{i=1}^{m} \alpha_t(i)}$$

In this paper, we use the hard competition. In hard competition, it is believed that only one model is responsible for describing the observation at time $t + 1$. This is done by considering only the model j with the highest predicted probability $p(S_{t+1} = j | \mathcal{Y}_{1:t}; \Theta)$.

3.2. Monte-Carlo simulations

We now come to the core part of the problem: how to compute the VaR of a portfolio using Monte Carlo simulation. The return of the portfolio at present time t will be denoted by R_t^p. Let us assume that R_t^p depends on q risk factors (foreign exchange rates in our case), then a Monte Carlo computation of the VaR would consist of the following steps:

1. Choose the confidence level $1 - \alpha$ to which the VaR refers.

2. Simulate the evolution of the mixed common latent risk factors f_t from time t to time $t+1$ by generating n-tuples of pseudo random numbers with appropriate joint distribution that describe the behavior of the common risk factors in the optimal futur hidden state $\widehat{S}_{t+1|t}$.

   (a) Generate $\widetilde{f}_t$ from the multivariate normal distribution $\mathcal{N}(0, I_k)$.

   (b) Generate a random number U from the uniform distribution $\mathcal{U}(0, 1)$.

   (c) Return

   $$f_t^s = \sum_{n=1}^{M^{(f)}} \left(\widehat{\mu}_{jn}^{(f)} + \widehat{\Sigma}_{jn}^{*(f)} \cdot \widetilde{f}_t \right) \mathbb{I}_{\left\{ \sum_{l=1}^{n-1} \widehat{c}_{jl}^f \leq U < \sum_{l=1}^{n} \widehat{c}_{jl}^f \right\}}$$

   where $\sum_{l=1}^{0} \widehat{c}_{jl}^f = 0$. $\mathbb{I}$ is the indicator function and $\widehat{\Sigma}_{jn}^{*(f)}$ a lower triangular matrix obtained from the Cholesky decomposition of $\widehat{\Sigma}_{jn}^{(f)}$.

3. Simulate the evolution of the mixed specific risk factors ε_t from time t to time $t + 1$ by generating n-tuples of pseudo random numbers with appropriate joint distribution that describe the behavior of the specific risk factors in the optimal future hidden state $\widehat{S}_{t+1|t}$.

   (a) Generate $\widetilde{\varepsilon}_t$ from the multivariate normal distribution $\mathcal{N}(0, I_q)$.

   (b) Generate another random number V from the uniform distribution $\mathcal{U}(0, 1)$.

   (c) Return

   $$\varepsilon_t^s = \sum_{m=1}^{M^{(y)}} \left(\widehat{\mu}_{jm}^{(y)} + \widehat{\Sigma}_{jm}^{*(y)} \cdot \widetilde{\varepsilon}_t \right) \mathbb{I}_{\left\{ \sum_{o=1}^{m-1} \widehat{c}_{jo}^y \leq V < \sum_{o=1}^{m} \widehat{c}_{jo}^y \right\}}$$

where $\sum\limits_{l=1}^{0} \widehat{c}_{jl}^{y} = 0$ and $\widehat{\Sigma}_{jm}^{*(y)}$ is a lower triangular matrix obtained from the Cholesky decomposition of $\widehat{\Sigma}_{jm}^{(y)}$.

Note that the number M_s of these n-tuples is a critical variable when using this procedure. Obviously, the larger M_s is, the more accurate the VaR will be. However, the simulation can be time-consuming, especially when conducted in a recursive or rolling forecasting scheme (see [1]). Fantazzini (2008) suggests a choice of 100.000 simulations, while Bastianin (2009) suggests only 5.000 and Lu et al., (2014) uses 10.000. For the accuracy of VaR, We use 25.000 simulations, which represents a good compromise between accuracy and efficiency.

4. Compute the portfolio returns, $R_{t+1,1}^{p}$, $R_{t+1,2}^{p}$, ..., R_{t+1,M_s}^{p}, using the simulated vectors of the common latent and specific risk factors at time $t+1$:

$$R_{t+1,s}^{p} = \delta_1 y_{1t+1}^{s} + \delta_2 y_{2t+1}^{s} + ... + \delta_q y_{qt+1}^{s}$$

where $\delta_1, \delta_2, ..., \delta_q$ are the weights of the q assets of our portfolio, whose returns are simulated thanks to the following model:

$$y_{t+1}^{s} = A_j f_t^{s} + \varepsilon_t^{s}$$

5. Ignore the fraction of the α worst returns $R_{t+1,s}^{p}$. The minimum of the remaining returns is the VaR of our portfolio, $VaR(\alpha, t, t+1)$.

As soon as the time evolves from t to $t+1$, we can backtest $VaR(\alpha, t, t+1)$ by comparing it with the effective portfolio return at time $t+1$, R_{t+1}^{p}.

4. Numerical example

In this section, we report on an analysis of the correlation structure of three representative Tunisian foreign exchange rates using our new mixed factorial HMM. Second, we calculate risk measures of the portfolio from the well-fitted latent structure. Lastly, through the VaR exceedance test based on the best fitted model and alternatives, we evaluate how well the best-fitted model describes the dependence structure of the foreign exchange rate series.

4.1. Dataset description

All the competitor models presented in this work are applied to the Tunisian external public debt portfolio composed by the main currencies, namely the American dollar, the European euro, and the Japanese yen expressed in terms of the Tunisian dinar (TND/USD, TND/EUR and TND/JPY). Our database contains 2250 daily exchange rates from January 08, 2011 to December 30, 2018. This period is considered as a transitional stage from fixed to floating exchange rate regime. This transition period has experienced a succession of severe crises and was characterized by significant fluctuations and recurrent shocks in volatility, which peaked at the end of September 2016.

Note here that our approach can be applied to model the switching latent dependence and co-movement structures of the returns of any kind of risky financial assets, but in this work we have especially chosen this dataset to test the ability of this new specification to accurately reproduce observed patterns and to identify the major dimensions of change over the period from 2011 to 2018, including the impact of the crisis and the changing of the foreign exchange

rate regimes over time on the common and specific variances and the correlation structure of the different series. The data were downloaded from the PACIFIC Exchange Rate Service[§]. For evaluating our VaR-methodology and the other competing models, we transformed the exchange rates into log-returns.

Using basic descriptive statistics, we tested the empirical skewness and kurtosis of the log-returns against the values of normal distributions. We also performed the Shapiro-Wilk test of normality. All the results are summarized in Table 1. There are also p-values of relevant test in brackets.

Table 1: *Basic descriptive statistics of the daily log-returns*

Descriptive statistics	USD/TND	EUR/TND	JPY/TND
Minimum	-3.56%	-1.93%	-4.04%
Maximum	2.71%	2.04%	6.02%
Mean	0.0198%	0.0149%	0.0211%
Median	0.0071%	0.0159%	0.0090%
Standard deviation	0.51%	0.32%	0.78%
Skewness	-0.2362	-0.0343	0.3341
	(0.0000)	(0.158)	(0.0001)
Kurtosis	6.5780	5.3086	7.5014
	(0.0000)	(0.0000)	(0.0000)
S-W test	(0.0000)	(0.158)	(0.0001)

It follows from the results above that empirical distribution of the selected time series are non-normal. All the log-returns are even skewed (JPY/TND positive, USD/TND and EUR/TND negative); the kurtosis is in all cases higher than it corresponds to the normal distribution.

The first panel of figure 1 shows the histograms of the daily log-return series. We also superimpose, on each histogram, the normal density function using the same mean and the same variance. The second panel of this figure shows the normal probability plots for the different series, which displays departures from the Gaussian shape and can indicate their nature. From this figure it appears clearly, that all the log-return series are not normally distributed during this period. In particular, the normal probability plots show that the empirical probability density functions at the tails of the distributions are heavier than those of the Gaussian distribution: the left tail is above the straight line, and the right tail is below it. Therefore we can conclude that it is really reasonable assuming a mixture of distributions in our application.

4.2. Identification of the best VaR model

Now, our objective is to identify the most appropriate forecasting methodology to predict the VaR for the Tunisian public debt portfolio. In a first step, the latent correlation structure is estimated from the proposed analytical form of the jointly multivariate distribution. To compute the VaR, the data set was divided into two parts: the estimation period (In Sample) which begins on 08/01/2011 until 05/01/2012, (250 observations) and the backtesting period (Out-of-Sample), that begins on 06/01/2012 and ends on 30/12/2018 (2000 observations). To estimate the VaR for a given confidence level, we used Monte Carlo simulations and the portfolio weights: $\delta_1 = 75\%$ (for the EUR), $\delta_2 = 15\%$ (for the USD) and $\delta_3 = 10\%$ (for the JPY).

In a second step we back test the VaR, using the method of rolling sample, with the significance levels 1%, 2%, 5% and 10%. In each backtesting step 25.000 trials were simulated and VaR was calculated. In line with the backtesting procedure described in [9], the observed

---

[§]The university of British Columbia, Sauder School of Business http://fx.sauder.ubc.ca/.
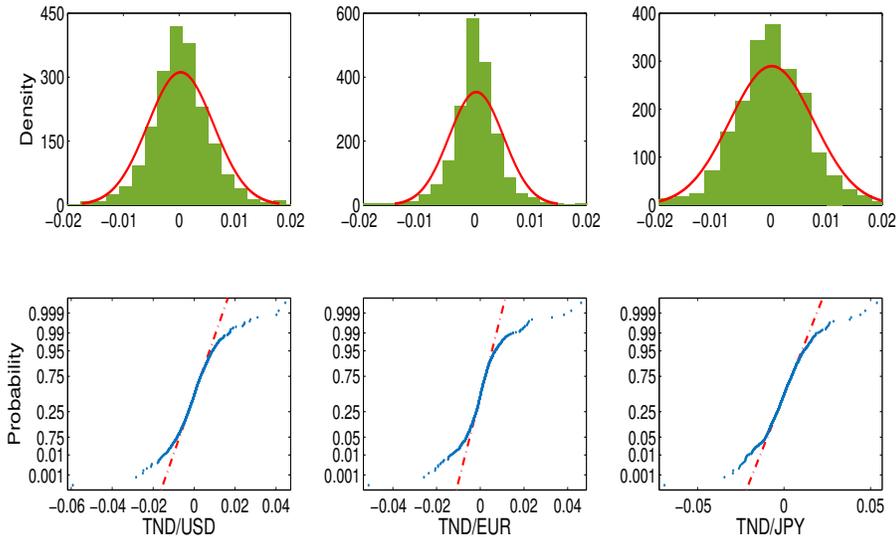
Figure 1: *Histograms (the top panels) and the Normal probability plots (the bottom panels) of the daily, log return series from 08-01-2011 to 30-12-2018.*

numbers of exceptions were recorded and p-values of LR_{PoF}, LR_{IND} and LR_{CC} tests were calculated.¶ These computations were done in many ways assuming that the data are generated from: a multivariate Gaussian distribution using the classical Monte-Carlo Method (CMC), a mixture of multivariate Gaussian distributions (GMM), a mixture of factor analyzers model (MFA), a mixed hidden Markov model (MHMM), and a mixed factorial hidden Markov model (MFHMM).

All Backtesting results are presented in tables 2-6. In general, these results show that there has been an appreciable improvement in the prediction of the VaR through the Monte Carlo based-MFHMM and MFA models compared to the CMC, and the Monte Carlo based-GMM and MHMM methodologies. As expected, the CMC methodology has the worst results on all the period 2011-2018. Indeed, this approach presented poor results through underestimating the VaR for the different risk levels 1%, 2%, 5% and 10%. The results of the conditional coverage test for this method was, respectively, 8.5074, 8.8173, 9.3084 and 10.1899.

Table 2: *Backtesting results of the CMC method.*

Characteristics	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 2\%$	$\alpha = 1\%$	LR_{critic} $\alpha^* = 5\%$
Failure rate	13.22%	7.10%	3.41%	1.65%	
First violation	9	9	13	13	
LR_{PoF}	7.391	8.073	9.224	10.121	3.841
LR_{IND}	1.1164	0.7443	0.0844	0.0689	3.841
LR_{CC}	8.5074	8.8173	9.3084	10.1899	5.991

From Figure 2, we can easily see that the predicted VaR obtained by the optimal mixed factorial HMM is highly affected by the significant and recurrent shocks in volatility. Therefore, we conclude that the poor results obtained by the CMC, GMM and MHMM methods are

---

¶LR_{PoF} is the Proportion of Failure likelihood ratio; LR_{ind} is the likelihood ratio when the exceptions are distributed equally in time, i.e. without any dependence (autocorrelation); and LR_{cc} is the conditional coverage joint likelihood ratio, including both conditional coverage and independence between exceptions.

due to the abnormal fluctuations of the three currencies during this period. Accordingly, we reject the credibility of these models because of their poor fit and their lack of reactivity. To accurately capture risks in such portfolios, it is important for risk managers to incorporate the co-movements and the interactions between the different heterogeneous risk factors. Such correlation structure of the risk factors can be modeled using a mixture of linear factor models with a Markov switching structure for the parameters. This permits reflecting two defining features of the latent volatility: co-movement among financial returns and switching between different unobservable regimes. Thus, as noted before, the Monte Carlo-MFHMM approach is the most suitable to construct the joint multivariate loss distribution in this situation.
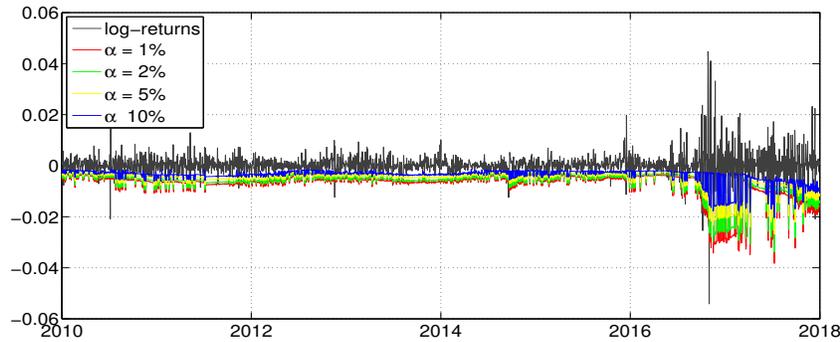
Figure 2: *Backtesting results of the Tunisian public debt portfolio, using the optimal mixed factorial HMM ($N_s = 2$, $N_{mix} = 3$) for quantiles $\alpha = 1\%$, 2%, 5% and 10% and 25000 Monte Carlo steps.*

Table 3: *Backtesting results of the GMM.*

Model	Characteristics	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 2\%$	$\alpha = 1\%$	LR_{critic} $\alpha^* = 5\%$
	Failure rate	12.55%	6.25%	3.50%	1.55%	
$N_{mix} = 2$	First violation	9	10	13	13	
	LR_{PoF}	6.7508	8.1512	9.4058	9.5176	3.841
	LR_{IND}	1.2823	0.8549	0.0399	0.0941	3.841
	LR_{CC}	8.0331	9.0061	9.4457	9.6117	5.991
	Failure rate	12.25%	6.10%	3.35%	1.40%	
$N_{mix} = 3$	First violation	7	9	13	13	
	LR_{PoF}	3.459	5.626	6.147	6.084	3.841
	LR_{IND}	1.0686	0.7124	0.0285	0.0672	3.841
	LR_{CC}	4.5276	6.3384	6.1755	6.1512	5.991

As mentioned earlier, the improvement in the results of the VaR estimates (for the different confidence levels) are mainly due to the fact that our proposed method is more flexible, than the other competing models, in capturing the heterogeneous volatility fluctuations and co-movement between log-returns over time. The failure rates given in tables 2-6 and figure 2, show high correlation between VaR violations and the volatility fluctuations during the Out-of-Sample period.

We note here an improvement in the detection of the first violation. For the risk level 10%, the optimal MFHMM and MFA models ($N_{mix} = 3$, $N_s = 2$) detect a first violation on the 8-*th* day compared to the CMC approach 9-*th* day, GMM 7-*th* day and MHMM 9-*th* day. Furthermore, for the 1% and 2% risk levels the first violations occurred at the 18-*th* and 17-*th*

Table 4: *Backtesting results of the MFA.*

Model	Characteristics	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 2\%$	$\alpha = 1\%$	LR_{critic} $\alpha^* = 5\%$
	Failure rate	11.88%	6.17%	3.16%	1.83%	
$N_{mix} = 1$	First violation	11	13	17	17	
	LR_{PoF}	4.4911	5.3930	6.2578	6.4078	3.841
	LR_{IND}	1.4456	0.9638	0.0011	0.0017	3.841
	LR_{CC}	5.9367	6.3568	6.2589	6.4095	5.991
	Failure rate	11.11%	5.92%	2.74%	1.57%	
$N_{mix} = 2$	First violation	9	12	17	17	
	LR_{PoF}	2.4210	3.3368	4.4675	5.6425	3.841
	LR_{IND}	1.2391	0.8261	0.0042	0.0036	3.841
	LR_{CC}	3.6601	4.1629	4.4717	5.6461	5.991
	Failure rate	10.15%	5.25%	2.15%	1.15%	
$N_{mix} = 3$	First violation	8	12	17	17	
	LR_{PoF}	1.3508	1.8807	2.0482	3.1732	3.841
	LR_{IND}	1.0326	0.6884	0.0000	0.0000	3.841
	LR_{CC}	2.3834	2.5691	2.0482	3.1732	5.991

days. We note also that the optimal MHMM and MFA methods give similar results for the 1% coverage rate. More precisely, the results obtained by the LR_{CC} test give the same failure rate (1.15%) and the same statistic (3.1732). In this paper, both the optimal MFHMM and MFA methods give too similar VaR sequences in such a way that the backtests cannot easily discriminate between them.

Backtesting results presented in tables 4 and 6, show the validity of the risk measures at the 1% and 2% levels of the optimal MFHMM and MFA models ($N_{mix} = 3$, $N_s = 2$). Moreover, in terms of conditional and unconditional coverage tests, it appears that the optimal MFHMM model gives better results than those obtained by the MFA method.

From table 6 we can see that at the risk level 2%, the failure rate obtained by the optimal MFHMM for the tunisian external public debt portfolio has been 2.05%. This result imply that our proposed model is more accurate than the optimal MFA model (table 4). Moreover, it has approved the "unconditional coverage" and the "independence" tests, respectively (0.0923) and (0.0000), which imply a significant conditional coverage Test ($0.0923 < \chi^2 = 5.991$).

Tables 6 and 4, show also the superiority of the optimal MFHMM compared to the best MFA for the risk levels 5% and 10%. In terms of LR_{CC} tests, it appears that the MFHMM results (1.4685; 0.6244) are more significant than those of the MFA model (2.5691; 2.3834).

5. Conclusion

This paper introduces a new VaR methodology suitable for trading portfolios that are driven by correlated financial variables characterized by strong volatility clustering, and skewed and fat-tailed returns distributions. The use of the mixed factorial HMM in VaR allows the risk manager to accurately account for these characteristics that occur in real data. We test the method, with different confidence levels, on the Tunisian external public debt portfolio and show that MFHMM-VaR compares favorably to VaRs based on the classical Monte Carlo method, the Gaussian mixture model, the mixture of factor analyzers model and the mixed HMM in terms of backtesting violations. More precisely, for each confidence level the absolute deviations of the percentages of the backtesting failures from the quantiles are smaller in the case of the mixed factorial HMM compared with the other competing models. Hence, our new approach is proven to provide an objective and more accurate forecasts of returns, and it outperforms particularly

Table 5: *Backtesting results of the MHMM.*

Model	Characteristics	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 2\%$	$\alpha = 1\%$	LR_{critic} $\alpha^* = 5\%$
$N_{mix} = 1$	Failure rate	12.48%	6.61%	3.77%	2.33%	
$N_s = 2$	First violation	11	14	15	15	
	LR_{PoF}	4.2857	4.3966	5.1369	5.6425	3.841
	LR_{IND}	2.8058	2.5372	1.0184	1.0109	3.841
	LR_{CC}	7.0915	6.9338	6.1553	6.6534	5.991
	Failure rate	12.51%	6.43%	3.56%	2.18%	
$N_s = 3$	First violation	10	13	15	15	
	LR_{PoF}	4.4603	4.3683	5.1565	5.8771	3.841
	LR_{IND}	2.9339	2.3560	1.0110	1.0117	3.841
	LR_{CC}	7.3942	6.7243	6.1675	6.8888	5.991
$N_{mix} = 2$	Failure rate	12.34%	6.28%	3.35%	2.11%	
$N_s = 2$	First violation	10	12	15	15	
	LR_{PoF}	3.5666	3.4533	4.1271	4.5252	3.841
	LR_{IND}	2.8493	2.6997	1.1214	1.1103	3.841
	LR_{CC}	6.4159	6.1530	5.2485	5.6355	5.991
	Failure rate	11.93%	5.72%	3.09%	1.93%	
$N_s = 3$	First violation	11	13	14	15	
	LR_{PoF}	3.5312	3.4249	4.1174	3.4078	3.841
	LR_{IND}	2.6776	2.5184	0.1452	0.1785	3.841
	LR_{CC}	6.2088	5.9433	4.2626	3.5863	5.991
$N_{mix} = 3$	Failure rate	11.89%	5.25%	2.35%	1.15%	
$N_s = 2$	First violation	9	13	14	15	
	LR_{PoF}	4.3541	4.2833	3.0978	3.1732	3.841
	LR_{IND}	2.7184	1.8123	0.0000	0.0000	3.841
	LR_{CC}	7.0725	6.0956	3.0978	3.1732	5.991
	Failure rate	12.20%	6.32%	2.84%	1.55%	
$N_s = 3$	First violation	9	12	14	15	
	LR_{PoF}	6.4249	5.3400	4.1467	4.7598	3.841
	LR_{IND}	2.2621	2.1748	1.6228	1.2543	3.841
	LR_{CC}	8.6870	7.5148	5.7695	6.0141	5.991

during the crisis period. This supports the economic argument for integrating heterogeneity and Markov switching properties of returns into the risk measurement methodology.

To be able to decide whether or not one should prefer the mixed factorial HMM to the traditional models, a supplementary investigation of the speed of the proposed algorithm would be useful. Furthermore, our model can be generalized to one where one allows the common latent factors and the specific factors to be stochastic functions of time. Secondly, we can also think of the case where the state transition probabilities are not homogeneous in time, but depend on the previous state and the previously observed covariates levels. The study of such models would provide a further step in the extension of hidden Markov models to mixed conditionally heteroscedastic factor analysis and allow for further flexibility in value-at-risk applications.

Acknowledgements

Table 6: *Backtesting results of the MFHMM.*

Model	Characteristics	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 2\%$	$\alpha = 1\%$	LR_{critic} $\alpha^* = 5\%$
$N_{mix} = 1$	Failure rate	11.26%	6.07%	2.81%	1.78%	
$N_s = 2$	First violation	7	10	17	17	
	LR_{PoF}	0.4778	0.3823	0.1292	1.6425	3.841
	LR_{IND}	0.3963	1.6736	0.0017	0.0143	3.841
	LR_{CC}	0.8741	2.0559	0.1309	1.6568	5.991
	Failure rate	10.91%	5.66%	2.44%	1.55%	
$N_s = 3$	First violation	8	11	17	17	
	LR_{PoF}	0.4437	0.3550	0.1477	1.8771	3.841
	LR_{IND}	1.4680	1.5540	0.0019	0.0163	3.841
	LR_{CC}	1.9117	1.909	0.1496	1.8934	5.991
$N_{mix} = 2$	Failure rate	10.47%	5.38%	2.21%	1.33%	
$N_s = 2$	First violation	8	11	17	18	
	LR_{PoF}	1.5461	0.4370	0.1200	1.5252	3.841
	LR_{IND}	1.4530	1.9126	0.0016	0.0133	3.841
7	LR_{CC}	2.9991	2.3496	0.1216	1.5385	5.991
	Failure rate	10.07%	5.10%	2.14%	1.22%	
$N_s = 3$	First violation	8	12	17	18	
	LR_{PoF}	1.5120	0.4097	0.1108	1.4078	3.841
	LR_{IND}	1.4247	1.7931	0.0014	0.0122	3.841
	LR_{CC}	2.9367	2.2028	0.1122	1.4200	5.991
$N_{mix} = 3$	Failure rate	10.00%	5.10%	2.05%	1.15%	
$N_s = 2$	First violation	8	12	17	18	
	LR_{PoF}	0.3413	0.2731	0.0923	1.1732	3.841
	LR_{IND}	0.2831	1.1954	0.0000	0.0000	3.841
	LR_{CC}	0.6244	1.4685	0.0923	1.1732	5.991
	Failure rate	10.00%	5.10%	2.05%	1.15%	
$N_s = 3$	First violation	8	12	17	18	
	LR_{PoF}	1.4096	0.3277	0.1384	1.7598	3.841
	LR_{IND}	1.3397	1.4345	0.0018	0.0153	3.841
	LR_{CC}	2.7493	1.7622	0.1402	1.7751	5.991

References

[1] Bastianin, A. (2009). Modelling Asymmetric Dependence Using Copula Functions: An application to Value-at-Risk in the Energy Sector. *Fondazione Eni Enrico Mattei Working Papers*, Paper 285. https://services.bepress.com/feem/paper285

[2] Dempster, A.P., Laird, N.M., and Rubin, D.B. (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society, Series B (Methodological)*, 39(1), 1-38. doi:10.1111/j.2517-6161.1977.tb01600.x

[3] Fantazzini, D. (2008). Dynamic copula modelling for value at risk. *Frontiers in Finance and Economics*, 5, 72-108. https://ssrn.com/abstract=1542608

[4] Gales, M.J.F. (1998). Maximum likelihood linear transformations for HMM-based speech recognition. *Computer Speech and Language*, 12(2), 75-98. doi:10.1006/csla.1998.0043

[5] Jorion, P. (2011). *Financial Risk Manager Handbook Plus Test Bank*. FRMR Part I/Part II, 6th Edition - John Wiley & Sons, Inc., Hoboken, New Jersey, USA.

[6] Lu, X.F., Lai, K.K., and Liang, L. (2014). Portfolio value-at-risk estimation in energy futures markets with time-varying copula-GARCH model. *Annals of Operation Research*, 219, 333-357. doi:10.1007/s10479-011-0900-9

[7] McLachlan, G., and Krishnan, T. (2008). *The EM Algorithm and Extensions*, 2nd Edition. John

Wiley and Sons, Inc., Hoboken, New Jersey, USA. doi:10.1002/9780470191613

[8] Mosbahi, M. N., Saidane, M., and Messabeb, S. (2017). Mixture Of Probabilistic Factor An-alyzers For Market Risk Measurement: Empirical Evidence From The Tunisian Foreign Ex-change Market. *Risk governance & control: financial markets & institutions*, 7(2,1), 158-169. doi:10.22495/rgcv7i2c1p4

[9] Saidane, M. (2017). A Monte-Carlo-based Latent Factor Modeling Approach with Time-Varying Volatility for Value-at-Risk Estimation: Case of the Tunisian Foreign Exchange Market. *Industrial Engineering & Management Systems*, 16(3), 400-414. doi:10.7232/iems.2017.16.3.400

[10] Saidane, M., and Lavergne, C. (2007). Conditionally heteroscedastic factorial HMMs for time series in finance, *Applied Stochastic Models in Business and Industry*, 23(6), 503-529. doi:10.1002/asmb.687

[11] Saidane, M., and Lavergne, C. (2006). On Factorial HMMs for Time Series in Finance. *The Kyoto Economic Review*, 75, 63-90. doi:10.11179/ker.75.63