

HMM based scenario generation for an investment optimisation problem

Christina Erlwein · Gautam Mitra · Diana Roman

Published online: 30 March 2011
© Springer Science+Business Media, LLC 2011

Abstract The Geometric Brownian motion (GBM) is a standard method for modelling financial time series. An important criticism of this method is that the parameters of the GBM are assumed to be constants; due to this fact, important features of the time series, like extreme behaviour or volatility clustering cannot be captured. We propose an approach by which the parameters of the GBM are able to switch between regimes, more precisely they are governed by a hidden Markov chain. Thus, we model the financial time series via a hidden Markov model (HMM) with a GBM in each state. Using this approach, we generate scenarios for a financial portfolio optimisation problem in which the portfolio CVaR is minimised. Numerical results are presented.

Keywords Scenario generation · Hidden Markov model · Geometric Brownian motion · Asset allocation · Optimal parameter estimation

1 Introduction

Stochastic programming is an important tool for portfolio optimisation. Since the future returns of most assets are uncertain, the outcome of an investment is uncertain. Usually, the decision maker tries to achieve an optimal combination of risk and expected return.

C. Erlwein (✉)
Fraunhofer Institute, ITWM, Fraunhofer Platz 1, 67663 Kaiserslautern, Germany
e-mail: christina.erlwein@itwm.fraunhofer.de

G. Mitra · D. Roman
CARISMA, Brunel University, Uxbridge UB8 3PH, UK

G. Mitra
e-mail: gautam.mitra@brunel.ac.uk

D. Roman
e-mail: diana.roman@brunel.ac.uk

G. Mitra · D. Roman
OptiRisk Systems Ltd, Uxbridge UB9 4DA, UK

This is done by solving stochastic programs: optimisation problems in which some of the parameters (the future returns) are not certain, but described by distributions. In order for the stochastic programs to be numerically solved, the distributions involved are approximated by discrete distributions with a finite number of outcomes (scenarios). An overview of financial optimisation problems can be found in Mulvey (2001).

Thus, a crucial issue for financial optimisation problems (with portfolio selection as a particular case) is the **scenario generation for the future realisations of the time series involved**. This is closely related to a proper modelling of the financial time series.

Among the most popular models for financial time series are the Geometric Brownian motion (GBM) (see for example Ross 2002) and ARCH-GARCH models (Engle 1982; Bollerslev 1986). With GBM, stock prices are approximated by continuous time stochastic processes. **ARCH-GARCH models were designed for modelling time-dependent variance**. However, both methods are accepted only with reservation; one of the main criticism is that they underestimate the lower tails of distributions, thus the severity of unfavourable events. Both models may be used for scenario generation by sampling from the assumed distributions. A more recent method of scenario generation for general stochastic programs is the “moment-matching” approach (Hoyland et al. 2003) in which it is assumed that the moments of the “true” distribution are known; a discrete distribution with the same moments is then constructed. An overview of scenario generation for general stochastic programs is given in Kaut and Stein (2007).

Over the last years, hidden Markov models (HMM) became increasingly popular for modelling financial time series. One main idea behind these models is that some information on the market is hidden in noisy observations, e.g. price movements. The underlying hidden information can symbolise different stages of a business cycle, like expansion, peak, recession, trough and recovery, which influence the price movements. The market may switch from time to time from a state to another; each state is characterised by different parameters of the model for the observable stochastic process. In this case, one single model is valid only for short periods; regime switching models offer a better way of modelling.

A **hidden Markov model (HMM) is a particular regime-switching model**, in which there are two stochastic processes involved: apart from the one of interest, that is observable (e.g. asset prices), there is an underlying stochastic process describing the system’s state over time, that is not observable, i.e. “hidden”. HMMs were applied to a variety of fields, in particular to speech recognition (Rabiner 1989); a review of HMMs and their applications is given in Rabiner (2002).

A review of the applications of HMM for financial time series modelling and forecasting is given in Zhang (2004). In Hamilton and Susmel (1994) weekly time series were modelled using regime switching with ARCH-type models within regimes. A simpler regime switching model for monthly returns with log-normal distributions is proposed in Hardy (2001). Messina and Toscani (2008) use autoregressive HMM for modelling univariate financial time series and for generating scenarios to describe their possible future evolution. In order **to validate the HMM based scenario generation method**, they show that the resulting Monte Carlo-sampled distributions can replicate with good approximation the empirical distribution of the observed data. In Roman et al. (2010) **multi-variate financial time series are modelled by using HMMs and mixtures of multi-variate normal distributions**; based on this approach, scenarios are **generated for a mean-downside risk optimisation problem**. It is shown that the scenario generator based on this method has **a reasonable stability and can capture extreme movements of asset prices**. However, the resulting model is computationally demanding and can only handle a limited number of time series.

In this paper, we use HMM and Geometric Brownian motion for modelling financial time series. The motivation behind is that, although GBM is a traditional and popular method, it

has recognised limitations. It fails to capture some important observed aspects, like volatility clustering and extreme behaviour. The limitations of GBM are mainly due to the fact that the parameters of the GBM (the drift and the volatility) are assumed to be constant. In our approach we aim to address this shortcoming by allowing the GBM parameters to switch between different states.

These parameters are estimated using a parameter estimation filtering approach (Elliott et al. 1995). Information of the Markov chain is filtered out of the observation process and optimal parameter estimates are derived through a filter-based Expectation-Maximisation (EM) algorithm. Once optimal parameters are estimated using historical information of the prices, scenario paths can be generated according to the parameter estimates and the filtered transition probability of the underlying Markov chain.

In our numerical experiments, we generate scenarios for the future returns of the FTSE100 index and of gold spots. The quality of the scenario generator is tested in a portfolio optimisation problem, in which the mean-CVaR approach is used, namely, the CVaR of the portfolio is minimised while a minimal portfolio expected return is imposed. We conclude that the scenario generation method proposed is reasonably stable.

The rest of the paper is organised as follows. In Sect. 2 the proposed approach for modelling financial time series is described. The basics of GBM and of HMM are presented. Once the parameter estimation and the future Markov chain estimation are made, the model can be used as a scenario generator. In Sect. 3 we describe the portfolio selection problem and the mean-CVaR optimisation model. Section 4 presents the numerical results. The scenarios generated with the approach presented in Sect. 2 are used as input in a mean-CVaR model; the stability of this scenario generator is investigated. Conclusions are presented in Sect. 5.

2 Modelling financial time series using HMM and GBM

2.1 The Geometric Brownian motion

The Geometric Brownian motion (GBM) is a continuous time stochastic process that is widely used as a reasonable approximation of stock price dynamics. Although stock prices are stochastic processes in discrete time, this approximation is generally accepted by practitioners for short—medium time periods. Formally, a stochastic process S_t follows a GBM if it satisfies the following stochastic differential equation:

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad (1)$$

where W_t is a Wiener process, and μ (the drift), σ (the volatility) are parameters of the model, assumed to be constant.

Equation (1) has the following analytic solution, where S_0 is the initial value (at time $t = 0$):

$$S_t = S_0 \exp \left\{ \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W_t \right\}. \quad (2)$$

One shortcoming of the model is that, in this form, it cannot account for observable phenomena in financial markets, like extreme behaviour and volatility clustering.

It is known that a Wiener process W_t can be simulated as $\sqrt{t}z_t$, where z_t is $N(0, 1)$ distributed. Thus, the GBM can be used for generating possible scenarios for S_t by sampling from the standard normal distribution and using (2).

2.2 Hidden Markov models with GBM as observation processes

With the HMM approach, there are two stochastic processes involved (both in discrete time): in addition to the one of interest, which is observable (e.g. stock prices), there is another one describing the “state of the system”, which is not directly observable. The underlying, hidden stochastic process describing the system’s state is a Markov chain, i.e. the system’s state at one point in time depends only on the system’s state at the previous point in time and not on the entire history. At each point in time, the system is in one of possible n states and it may switch from the current state to another (or it may stay within the same state) according to the transition probabilities of the underlying Markov chain. Another assumption is that the transition probabilities are not time dependent.

Each state generates observations for the stochastic process of interest according to its own probability density. In our approach, we assume that the observations of the stock prices are generated according to discretised Geometric Brownian motions; each state in the Markov chain has different parameters of the GBM. This approach may address the shortcomings of the GBM, described at the previous section.

We work on a complete probability space (Ω, \mathcal{F}, P) . The log-return process $y_k = \ln \frac{S_k}{S_{k-1}}$ has the following dynamics in discrete time:

$$y_{k+1} = f(\mathbf{x}_k) + \sigma(\mathbf{x}_k)z_{k+1}, \quad (3)$$

where $f(\mathbf{x}_k) = (\mu(\mathbf{x}_k) - \frac{\sigma^2(\mathbf{x}_k)}{2})$. The parameters f and σ are governed by the Markov chain \mathbf{x} in discrete time and are therefore able to switch between different regimes. The z_k ’s are a sequence of independent, identically distributed (IID) standard normal random variables independent of \mathbf{x} .

We denote the scalar product with $\langle \cdot, \cdot \rangle$, and with n the number of states of the Markov chain. The parameters associated with GBMs are of the form $\mathbf{f} = (f_1, f_2, \dots, f_n)^\top$ and $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_n)^\top$ such that $f(\mathbf{x}_k) = \langle \mathbf{f}, \mathbf{x}_k \rangle$ and $\sigma(\mathbf{x}_k) = \langle \sigma, \mathbf{x}_k \rangle$ where σ_i ’s are all positive for every $1 \leq i \leq n$. The observation at time $k+1$ depends on the state of \mathbf{x} at time k . This is a one-step delay model and is reasonable as y may not react to \mathbf{x} immediately.

We consider a homogeneous Markov chain \mathbf{x}_k with finite state in discrete time, $k = 0, 1, \dots$. The state space of \mathbf{x}_k is associated with the canonical basis $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ of \mathbb{R}^n , where $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)^\top \in \mathbb{R}^n$ and \top denotes the transpose of a row vector, we suppose that the distribution of \mathbf{x}_0 is known. Let $\mathcal{F}_k^0 = \sigma\{\mathbf{x}_0, \dots, \mathbf{x}_k\}$ be the σ -field generated by $\mathbf{x}_0, \dots, \mathbf{x}_k$, and \mathcal{F}_k be the complete filtration generated by \mathcal{F}_k^0 . Furthermore denote by \mathcal{Y}_k the complete filtration generated by the sequence of increasing sub-sigma fields $\sigma(y_1, y_2, \dots, y_k)$ from the observation process and let $\mathcal{G}_k = \mathcal{F}_k \vee \mathcal{Y}_k$ be the global filtration.

Under the real world probability measure P , the Markov chain \mathbf{x} has the dynamics

$$\mathbf{x}_{k+1} = \Pi \mathbf{x}_k + \mathbf{v}_{k+1} \quad (4)$$

where $\mathbf{v}_{k+1} := \mathbf{x}_{k+1} - \Pi \mathbf{x}_k$ is a martingale increment and $\Pi = (\pi_{ji})_{j,i=1,\dots,n}$ is the transition probability matrix with $\pi_{ji} = P(\mathbf{x}_{k+1} = \mathbf{e}_j | \mathbf{x}_k = \mathbf{e}_i)$, i.e., the probability of transitioning from state i to state j (as stated before, the assumption is that these transition probabilities are independent of time).

The parameters to be estimated are: the number of Markov states n , the transition probabilities π_{ji} , $j, i = 1, \dots, n$ and the parameters of the discretised GBM \mathbf{f} and σ . While optimal values can be derived recursively for \mathbf{f} , σ and Π , the number of states is supposed to be determined in advance; it is estimated separately from the rest of the parameters. In practice,

the state dimension of the hidden Markov chain may be dictated by the actual application or may be determined in an empirical manner by visual inspection of plots of the data (Messina and Toscani 2008; Geyer and Ziemba 2008). In Roman et al. (2010) the AIC criterion is used (Akaike 1974) in order to determine the number of Markov states.

Recursive parameters can be derived for the remaining parameters \mathbf{f} , σ and Π , these parameter estimates utilise adaptive filters for the Markov chain. Filters (the conditional expectations of adapted processes under the filtration \mathcal{Y}_k) are obtained for the state space of the Markov chain and related quantities, e.g. the jump and the occupation time process. The jump process of the Markov chain, namely the number of times the Markov chain jumps from state i to state j up to a given time k is denoted by $J_k^{(ji)}$,

$$J_k^{(ji)} = \sum_{l=1}^k \langle \mathbf{x}_{l-1}, \mathbf{e}_i \rangle \langle \mathbf{x}_l, \mathbf{e}_j \rangle. \quad (5)$$

The occupation time process which indicates how long the Markov chain stays in state i up to time k is denoted by $O_k^{(i)}$,

$$O_k^{(i)} = \sum_{l=1}^k \langle \mathbf{x}_{l-1}, \mathbf{e}_i \rangle. \quad (6)$$

Furthermore we need an auxiliary process $T_k^{(i)}$ for filtering the Markov chain,

$$T_k^{(i)}(g) = \sum_{l=1}^k \langle \mathbf{x}_{l-1}, \mathbf{e}_i \rangle g(y_l), \quad (7)$$

where g is a function that will take the form $g(y) = y$ or $g(y) = y^2$.

Information of the Markov chain through these processes is filtered out of our observation process y . We shall derive the adaptive filters under a reference probability measure \tilde{P} following an HMM filtering method by Elliott et al. (1995). The change of measure technique is widely used in filtering applications and was introduced for stochastic filtering by Zakai (1969). Elliott (1995) utilises this change of measure technique, which is based on a discrete-time version of Girsanov's theorem to derive optimal filters. This technique enables us to make calculations under a mathematically "ideal" measure, which we shall term as a reference probability measure. This new "ideal" probability measure is equivalent to the real world measure, the measure under which we have the observation process. Under the new measure all components of the observation process are $\mathcal{N}(0, 1)$ iid random variables and independent from the Markov chain. The Markov chain follows the same dynamics under both the reference probability measure and the real world measure. Changing from the real measure to the ideal measure leads to easier ways of calculating filters as Fubini-type results can be employed instead of direct calculations that require hard semi-martingale methods. So we construct (through a discretised version of Girsanov's Theorem) a new reference probability measure \tilde{P} under which all components of the observation process are $\mathcal{N}(0, 1)$ iid random variables and independent from the Markov chain, $\frac{d\tilde{P}}{dP}|_{\mathcal{G}_k} := \tilde{\Lambda}_k$. The Markov chain has the same dynamics as under the real measure. A reverse measure change back to the real world measure P allows us then to utilise the adaptive filters for the optimal parameter estimates.

We suppose we start under the reference probability measure \tilde{P} . We wish to construct an equivalent measure P , so that under P we have that $z_{k+1} := \frac{y_{k+1} - f(\mathbf{x}_k)}{\sigma(\mathbf{x}_k)}$ is a sequence of

$N(0, 1)$ iid random variables. To construct the real world measure P from \tilde{P} , we define the processes λ_l and Λ_l by

$$\lambda_l = \frac{\phi(\sigma(\mathbf{x}_{l-1})^{-1}(y_l - f(\mathbf{x}_{l-1})))}{\sigma(\mathbf{x}_{l-1})\phi(y_l)} \quad \text{and} \quad (8)$$

$$\Lambda_k = \prod_{l=1}^k \lambda_l, \quad k \geq 1, \quad \Lambda_0 = 1$$

where $\phi(z)$ is the probability density function of the standard normal distribution. We refer to Λ_k as the Radon-Nikodým derivative of P with respect to \tilde{P} , $\frac{dP}{d\tilde{P}}|_{\mathcal{G}_k} = \Lambda_k$.

We shall derive the filters under the reference probability measure \tilde{P} . For any \mathcal{G} -adapted process $\{H_k\}$ Bayes' Theorem gives us

$$E[H_k|\mathcal{Y}_k] = \frac{\tilde{E}[\Lambda_k H_k|\mathcal{Y}_k]}{\tilde{E}[\Lambda_k|\mathcal{Y}_k]}, \quad (9)$$

so that the conditional expectation under the real world probability measure can be easily obtained.

Notation We shall use the notation $\hat{H}_k := E[H|\mathcal{Y}_k]$ for any \mathcal{Y} -adapted process H and $\gamma(H)_k := \tilde{E}[\Lambda_k H_k|\mathcal{Y}_k]$.

Utilising a filter theorem by Elliott (1994), we derive recursive filters for the jump process, the occupation time process and auxiliary processes. The filters $\gamma(H)_k$ are derived under the reference probability measure, therefore we use Λ_k , the Radon-Nikodým derivative of P , componentwise in our filter equations. For this, write \mathbf{D} for the diagonal matrix whose entries d_{ij} are

$$d_{ij} = \begin{cases} \frac{\phi(\frac{y_{k+1}-f_i}{\sigma_i})}{\sigma_i \phi(y_{k+1})}, & \text{for } i = j, \\ 0, & \text{otherwise.} \end{cases}$$

The entries on the diagonal for the case $i = j$ are the componentwise elements from the process λ_l .

First we would like to estimate the state of \mathbf{x} , given the observations under P , the real world probability. Let $\hat{\mathbf{x}}^i = P(\mathbf{x}_k = \mathbf{e}_i|\mathcal{Y}_k) = E[\langle \mathbf{x}_k, \mathbf{e}_i \rangle|\mathcal{Y}_k]$ be the conditional distribution of \mathbf{x}_k given \mathcal{Y}_k under P . We define $\xi_k := \tilde{E}[\Lambda_k \mathbf{x}_k|\mathcal{Y}_k]$. Noting that $\sum_{i=1}^n \langle \mathbf{x}_k, \mathbf{e}_i \rangle = 1$, we have

$$\sum_{i=1}^n \tilde{E}[\langle \Lambda_k \mathbf{x}_k, \mathbf{e}_i \rangle] = \tilde{E} \left[\Lambda_k \sum_{i=1}^n \langle \mathbf{x}_k, \mathbf{e}_i \rangle \middle| \mathcal{Y}_k \right] = \tilde{E}[\Lambda_k|\mathcal{Y}_k] = \sum_{i=1}^n \langle \xi_k, \mathbf{e}_i \rangle.$$

Again noting that $\hat{\mathbf{x}}_k = E[\mathbf{x}_k|\mathcal{Y}_k]$, we get an explicit form for the conditional distribution, $\hat{\mathbf{x}}_k = \frac{\xi_k}{\sum_{i=1}^n \langle \xi_k, \mathbf{e}_i \rangle}$.

With Elliott's filtering theorem (see Elliott 1994), the following filters are derived:

$$\begin{aligned} \gamma(J^{(ji)}\mathbf{x})_l &= \Pi \mathbf{D}(y_l) \gamma(J^{(ji)}\mathbf{x})_{l-1} \\ &\quad + \langle \xi_{l-1}, \mathbf{e}_i \rangle \frac{\phi(\sigma_i^{-1}(y_l - f_i))}{\sigma_i \phi(y_l)} \pi_{ji} \mathbf{e}_j, \end{aligned} \quad (10)$$

$$\begin{aligned}\gamma(O^{(i)}\mathbf{x})_l &= \mathbf{PD}(y_l)\gamma(O^{(i)}\mathbf{x})_{l-1} \\ &+ \langle \xi_{l-1}, \mathbf{e}_i \rangle \frac{\phi(\sigma_i^{-1}(y_l - f_i))}{\sigma_i \phi(y_l)} \mathbf{Pe}_i,\end{aligned}\quad (11)$$

$$\begin{aligned}\gamma(T^{(i)}(g)\mathbf{x})_l &= \mathbf{PD}(y_l)\gamma(T^{(i)}(g)\mathbf{x})_{l-1} \\ &+ \langle \xi_{l-1}, \mathbf{e}_i \rangle \frac{\phi(\sigma_i^{-1}(y_l - f_i))}{\sigma_i \phi(y_l)} g(y_l) \mathbf{Pe}_i,\end{aligned}\quad (12)$$

with $\xi_{k+1} = \mathbf{PD}\xi_k$. Filters are derived for $\gamma(H_k\mathbf{x}_k)_k$, but the estimate for $\gamma(H_k)$ is easily obtained by summing the components of $\gamma(H_k\mathbf{x}_k)_k$. More details and proofs for these filter derivations can be found in Mamon et al. (2008). Note that the filters are adaptive, the filter at time l depends on the filter at time $l - 1$.

These derived adaptive filters for the processes of the Markov chain are used for the recursive model parameter estimates. The recursive formulas for the model parameters are derived with the help of a filter-based Expectation-Maximisation (EM) algorithm. The maximum likelihood estimation (MLE) for the parameters within the EM-algorithm makes use of the derived adaptive filters of the observation process. In the calculation of the MLE, the filters substitute terms involving the processes of the Markov chain. The following recursive optimal parameter estimates for the transition probabilities π_{ji} , the mean f_i and the variance σ_i of the observation process are derived (see Mamon et al. 2008 for a proof).

$$\hat{\pi}_{ji} = \frac{\hat{J}_k^{(ji)}}{\hat{O}_k^{(i)}} = \frac{\gamma(J^{(ji)})_k}{\gamma(O^{(i)})_k}, \quad (13)$$

$$\hat{f}_i = \frac{\hat{T}_k^{(i)}}{\hat{O}_k^{(i)}} = \frac{\gamma(T^{(i)}(y))_k}{\gamma(O^{(i)})_k} \quad (14)$$

and

$$\hat{\sigma}_i = \sqrt{\frac{\hat{T}_k^{(i)}(y^2) - 2\hat{f}_i\hat{T}_k^{(i)}(y) + \hat{f}_i^2\hat{O}_k^{(i)}}{\hat{O}_k^{(i)}}}. \quad (15)$$

Since here the adaptive filters are utilised, the obtained optimal parameter estimates contain information of the Markov chain at time k and can be updated whenever new information is available on the market. Parameters are re-estimated after a pre-specified time interval. The algorithm runs on these batches, new information available from the market is filtered and in turn used for updated parameter estimates. When suitable initial values are chosen (according to statistical characteristics of the observed time series), the parameter estimates converge.

Once the parameters are estimated, this method can be used as a scenario generator for the future price returns, after the Markov chain for the next time period is estimated. As derived above, the conditional distribution of the Markov chain for the next time step is given by $E[\mathbf{x}_{k+1}|\mathcal{Y}_k] = \mathbf{P}\hat{\mathbf{x}}_k$ with $\hat{\mathbf{x}}_k = E[\mathbf{x}_k|\mathcal{Y}_k] = \frac{\xi_k}{\sum_{i=1}^n \langle \xi_k, \mathbf{e}_i \rangle}$. Scenarios are then generated by sampling from the corresponding distributions, according to (3).

2.3 Algorithmic description

The optimal parameter estimates are calculated through the following HMM algorithm:

1. Choose initial parameter values \mathbf{f}_0 , σ_0 and $\mathbf{\Pi}_0$. Here, the initial values are chosen with regard to the mean and standard deviation of the modelled time series, the initial transition probabilities are equally distributed.
2. Set initial values for the processes $J_k^{(ji)}$, $O_k^{(i)}$ and $T_k^{(i)}(g)$. Here, the initial values of the processes are set to zero in each component.
3. Choose length of batch (min ~ 10 data points) and number of algorithm runs depending on length of time series.
4. Calculate the diagonal matrix \mathbf{D} with entries from the Radon-Nikod m derivative Λ_k .
5. Calculate the state estimator and filters for the jump and the occupation time process as well as auxiliary processes (10), (11) and (12).
6. After filters are calculated for the current batch, updates of the parameter estimates are obtained through (13), (14) and (15) utilising the adaptive filters. The resulting parameter estimates are taken as initial values in the next algorithm run.
7. Repeat the previous two steps on the next data batch, until the algorithm ran the chosen number of times.
8. The obtained optimal parameter estimates are utilised to generate scenarios for the returns of the financial time series.

The scenarios are generated through the following scenario generation algorithm:

1. Generate standard normal distributed random variables, which are used for simulating the white noise part (it is known that a Wiener process W_t can be simulated as $\sqrt{t} \cdot N(0, 1)$).
2. **Generate scenarios for the next time step:**
 - (a) The conditional distribution of the Markov chain for the next time step is calculated. As derived above, we have $E[\mathbf{x}_{k+1}|\mathcal{Y}_k] = \mathbf{\Pi}\hat{\mathbf{x}}_k$ with $\hat{\mathbf{x}}_k = E[\mathbf{x}_k|\mathcal{Y}_k] = \frac{\xi_k}{\sum_{i=1}^n \xi_k \cdot \mathbf{e}_i}$.
 - (b) The parameter values for the next time step are calculated as the scalar product between the conditional expectation of the Markov chain $\hat{\mathbf{x}}_{k+1}$ and the estimated optimal parameters,

$$f_{scen} = \langle f, \hat{\mathbf{x}}_{k+1} \rangle \quad \text{and} \quad \sigma_{scen} = \langle \sigma, \hat{\mathbf{x}}_{k+1} \rangle.$$

- (c) The scenarios for the asset returns in the next time step are created with the discretised version of the geometric Brownian motion using the parameters stated above

$$S_{scen} = C(k) * \exp((f_{scen} - 0.5 * \sigma_{scen}^2) + \sigma_{scen} * w_{scen}),$$

where $C(k)$ denotes the last actual data point of the observation process and w denotes the IID standard normal distributed random variable independent of the Markov chain for scenario $scen$.

- (d) The required log returns of the data series are then calculated for each scenario using the estimated prices calculated in the different scenarios above.

In our numerical experiments we consider two different frameworks for the scenario generations. In the first setting both time series are assumed to be independent, the optimal parameters for each process are derived in separate algorithms. The second setting assumes a dependency between the time series. Both observation processes are governed by the same Markov chain and are therefore estimated as vector observations. A thorough discussion on the filter derivation and optimal parameter estimation for vector observations can be found in Elliott et al. (1995) and Erlwein et al. (2009).

3 The mean-CVaR model for portfolio optimisation

A crucial issue in financial optimisation is the choice criterion used. The result of a financial decision (say a portfolio return) is uncertain and regarded as a random variable (with a discrete distribution, using scenario generation). Thus a natural question is what criterion to use for making choices among random returns. Various models for optimal portfolio selection have been proposed, starting with the popular Mean-Variance approach (Markowitz 1952), in which decisions (i.e., portfolio weights) are taken such that the portfolio's expected return is maximised while a risk measure (the variance) is minimised. More recently portfolio optimisation problems include more sophisticated risk measures. A popular risk measure in industry at the moment is Value-at-Risk (VaR), which indicates the maximum amount to be lost at a particular confidence level. One drawback of this measure in the application to stochastic optimisation problems is that it is non-smooth and non-convex and can therefore lead to multiple local extrema (Artzner et al. 1999). On the other hand, the Conditional Value-at-Risk (CVaR), also known as Mean Excess Loss or Mean shortfall, which is approximately equal to the conditional expectation of losses beyond VaR, is a risk measure with good theoretical and computational properties (Pflug 2000). CVaR models were developed for portfolio optimisation, where in one framework (Rockafellar and Uryasev 2000) risk is minimised with a given minimum level of portfolio return, while in other frameworks (Krokhmal et al. 2002) the return is maximised with a given maximum level of risk, or a weighted combination of risk and return is optimised. All formulations lead to the same efficient frontier (Krokhmal et al. 2002).

Let R_x be a random variable representing the return of a portfolio x over a given holding period and $A\% = \alpha \in (0, 1)$ a percentage which represents a sample of “worst cases” for the outcomes of R_x (usually, $\alpha = 0.01 = 1\%$ or $\alpha = 0.05 = 5\%$). The feasible set of decision vectors is denoted by X and assumed to be a convex polyhedron (usually, defined simply as $X = \{(x_1, \dots, x_n) \mid \sum_{j=1}^n x_j = 1, x_j \geq 0, \forall j = 1, \dots, n\}$).

The definition of CVaR at the specified level α is the mathematical transcription of the concept “average of losses in the worst $A\%$ of cases” (Acerbi and Tasche 2002), where the “loss” associated with R_x is usually described by the random variable $-R_x$, meaning a negative return is a loss, a positive return is a “gain” (however, the definition of CVaR allows for a general loss function).

For a detailed definition of CVaR, see Rockafellar and Uryasev (2002). CVaR is approximately equal to the average of losses greater than or equal to VaR (at the same confidence level α); in some cases, the equality is exact.

An important result, proved in Rockafellar and Uryasev (2002), is that the CVaR of a portfolio x can be calculated by solving a convex optimisation problem. Moreover and very importantly, CVaR can be optimised over the set of feasible decision vectors (feasible portfolios) and this is also a convex optimisation problem. In Rockafellar and Uryasev (2002), an auxiliary function is used, $F : X \times \mathbb{R} \rightarrow \mathbb{R}$,

$$F_\alpha(x, v) = \frac{1}{\alpha} E[-R_x + v]^+ - v,$$

where $[u]^+ = u$ if $u \geq 0$ and 0 otherwise. They proved that minimising CVaR over X can be done by minimising F_α over $X \times \mathbb{R}$.

When the random returns under consideration are represented as discrete random variables (via scenario generation), the CVaR optimisation problem is a LP.

The algebraic formulation of the mean-CVaR model is given below:

$$\begin{aligned}
 \text{(M-CVaR)} \quad & \min v + \frac{1}{\alpha S} \sum_{i=1}^S y_i \\
 \text{Subject to:} \quad & \sum_{j=1}^N -r_{ij}x_j - v \leq y_i, \quad \forall i \in \{1, \dots, S\}, \\
 & y_i \geq 0, \quad \forall i \in \{1, \dots, S\}, \\
 & \sum_{j=1}^N \mu_j x_j \geq d; \quad x \in X,
 \end{aligned}$$

where the parameters of the model are:

- S = the number of scenarios,
- N = the number of assets,
- r_{ij} = return of asset j under scenario i , $i = 1, \dots, S$, $j = 1, \dots, N$,
- μ_j = the expected return of asset j , $j = 1, \dots, N$ ($\mu_j = \frac{1}{S} \sum_{i=1}^S r_{ij}$),
- d = the desired expected return of the portfolio.

The decision variables of the model are:

- x_j = the fraction of portfolio wealth invested in asset j , $j = 1, \dots, N$,
- v = the negative of an α -quantile of the portfolio return distribution (it is an approximation for the α -VaR of the portfolio),
- y_i = the magnitude of the negative deviations of the portfolio return from the α -quantile, for every scenario $i \in \{1, \dots, S\}$ (they are 0 if the portfolio return is higher than the α -quantile).

4 Numerical experiments: a case study for an investment problem

4.1 Data set and computational set-up

We consider the following portfolio optimisation problem. An investor faces the problem of finding the optimal ratio between investing in gold spots, FTSE100 stocks or putting his money into a bank account. We consider a one time-step problem: decisions have to be taken now in order to give the best return at the next time period.

An investment in gold is traditionally seen as a long-term investment that can play a role in hedging against inflation and political or economical problems. Like other commodities, gold is traded on spot and future markets, but it has specific characteristics which are not common within commodity markets. Because of its role as a global currency, prices on spot markets are global. Gold supply and demand does not depend on seasonality, it has a low risk of supply interruption and low storage and insurance costs. Furthermore gold has no risk of spoilage and the consumption level relative to inventory is low. Unlike price processes for other commodities, a convenience yield is not included in the price process due to the distinctive gold features.

We generate scenarios for the future prices of gold and FTSE100 index according to the scenario generation method proposed in Sect. 2. The scenarios obtained are used as input in the mean-CVaR optimisation problem (the scenarios are considered to have equal

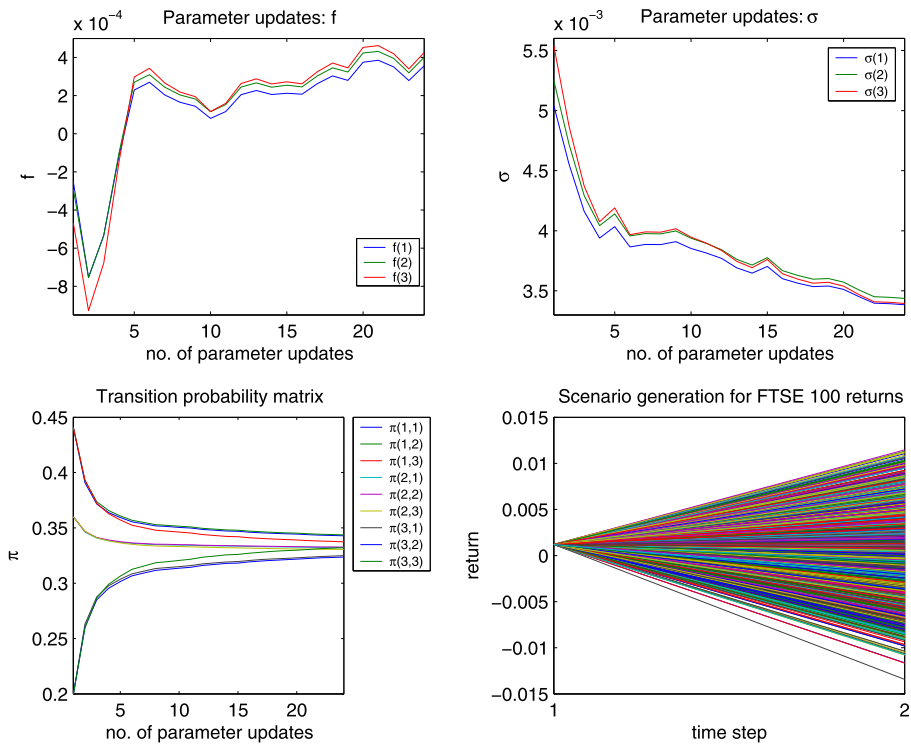


Fig. 1 Parameter estimation and scenario generation for FTSE 100

probability). We consider CVaR at confidence level $\alpha = 1\%$. The daily return on the bank account is assumed to be 0.0001. The minimal required level of daily portfolio return is set to $d = 0.0005$. The optimisation problems are formulated in AMPL (Fourer et al. 1989) and solved with the FortMP solver (Ellison et al. 1999).

The dataset used for generating the scenarios are historical prices of gold spots and FTSE 100 index, monitored daily over a 1-year period (in our numerical examples we use data from the year 2006). We considered a number of $n = 3$ possible states for the Markov chain. This is consistent with Roman et al. (2010), Messina and Toscani (2008), Geyer and Ziemba (2008), who, either by inspecting data or by using statistical techniques, estimated the number of Markov states to 3. It is in fact more or less agreed that three states usually give a good representation of asset prices, without overfitting the model.

We generate scenarios under two different frameworks. First, we consider that the two time series of interest are independent and thus governed by different Markov chains. Secondly, we consider them governed by the same Markov chain; observations are generated as vectors.

The algorithm for estimating the parameters runs 24 times on the data sets. For each data set the parameters are updated when new information arrives after batches of 10 data points. Scenarios are generated starting from the last data point. Figures 1 and 2 depict the optimal parameter estimates for the gold spot prices and for FTSE 100, respectively

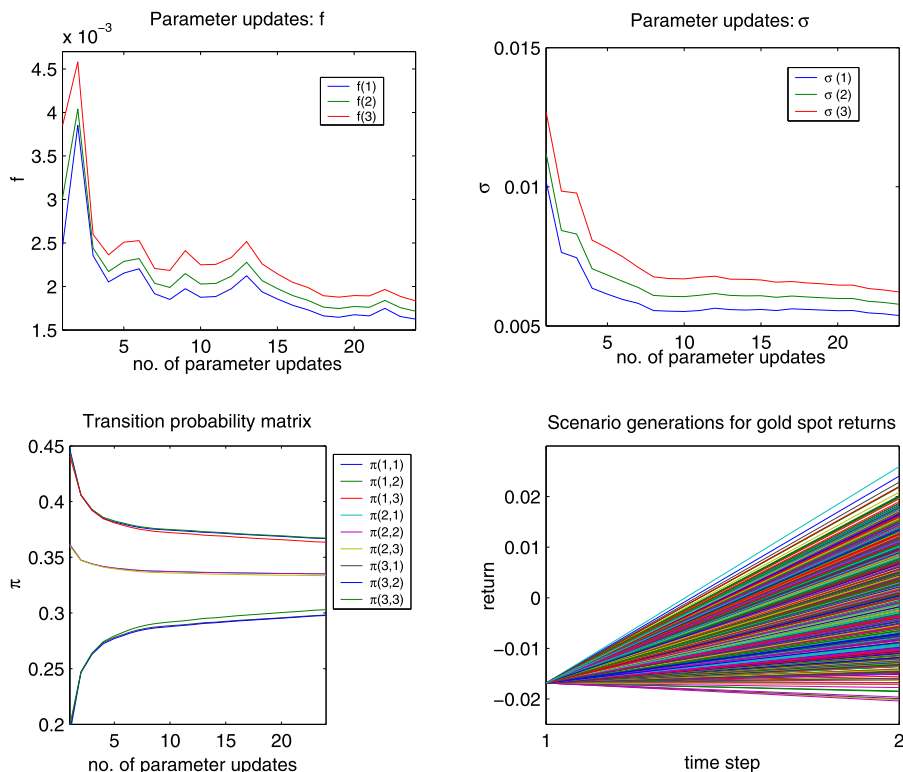


Fig. 2 Parameter estimation and scenario generation for gold spots

(estimated independently). Scenarios for the next time period are generated following the parameter estimation.

Under the assumption of correlated time series, we consider that both time series are governed by the same hidden Markov chain, which is filtered out from both processes. Under this framework we estimate the optimal parameters of both observation processes simultaneously and we generate vector observations. Figure 3 shows the parameter estimation of the observation processes for the vector observation case. Possible scenarios for both time series are then generated and depicted in Fig. 4.

4.2 Comparison of HMM and GBM scenarios with historical data

We generate 10,000 paths for both time series, the gold spot prices and FTSE 100, considering a 3-state Markov chain which follows the estimated transition probabilities. We simulate paths of the asset prices which take into account the optimal parameter estimates of the time series. For the purposes of comparison, we also consider a 1-state Markov chain, i.e. we consider the two time series as following a GBM. A comparison of the descriptive statistics of these paths and the actual time series can be found in Table 1; ten sample paths for each time series are plotted in Fig. 5.

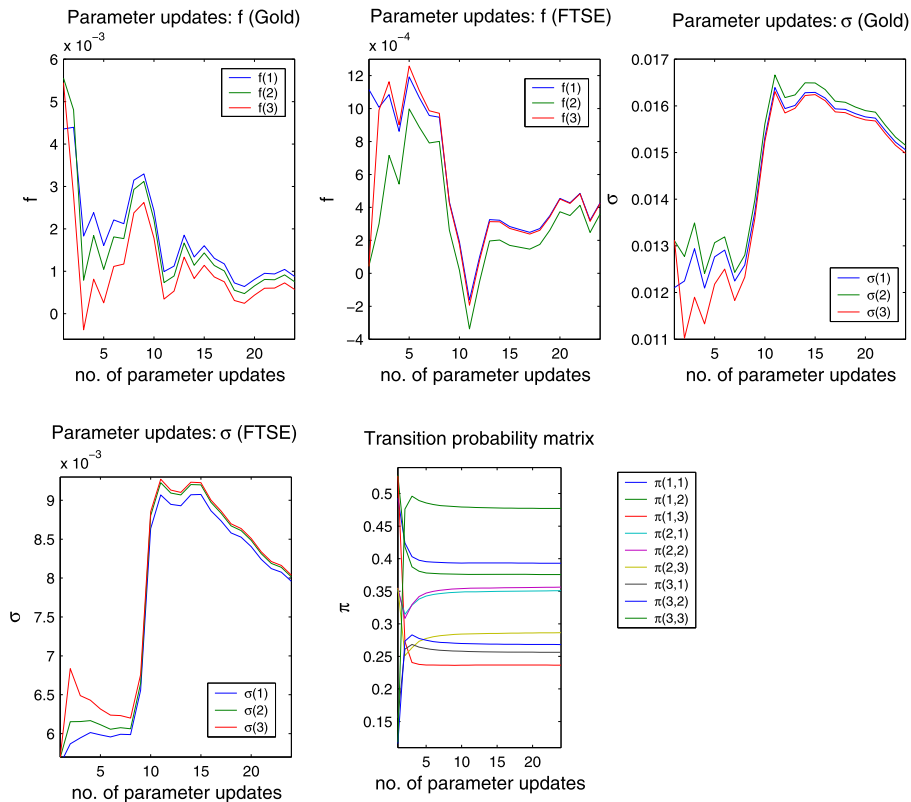


Fig. 3 Simultaneous parameter estimation for observation process of gold spots and FTSE 100 returns

It can be observed that, by considering a 3-state HMM, extreme downside price fluctuations are not under-represented—unlike in the GBM case, which gives rise to rather “over-optimistic” scenarios.

4.3 Stability of the scenario generators

Stability is one of the most important requirements from a scenario generator. Since any scenario generator has an element of randomness, the final outcome depends on the values drawn from the corresponding distributions (in our case, the standard normal distribution). Thus, every time a scenario generator is run, different sets of scenarios are obtained. Stability guarantees that the optimal solution of the optimisation problem of interest does not depend (but, possibly, only to a small extent) on the specific scenario set chosen.

For each framework (independent/correlated time series) we use the same methodology. We generate 10 sets, each with 5000 scenarios for the future returns of FTSE100 and gold.

The optimal solutions and objectives, for the case of independent observations, are displayed in Table 2. All values of the objective function lie between 0.002844 and 0.003325 with a mean of 0.0031 and variance of $2.7070\text{e-}008$. The in-sample stability of the scenario generator is further investigated through a variation of the size of the scenario sets.

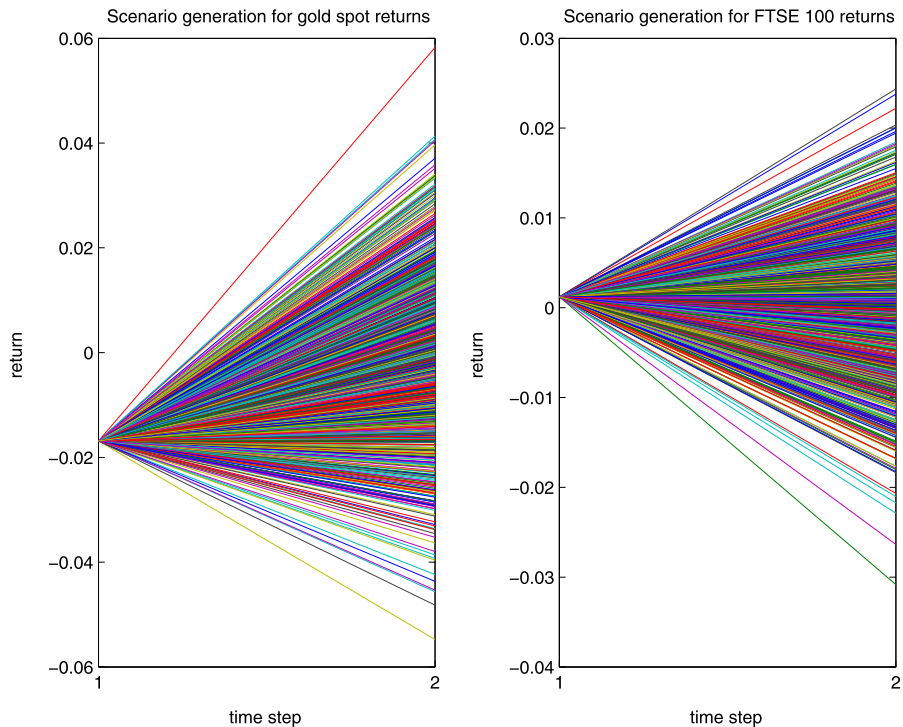


Fig. 4 Scenario generation for returns on gold spots and FTSE 100

Table 1 Descriptive statistics of generated paths for gold spot prices and FTSE

	Min	Max	Median	Mean	St.dev	Skewness	Kurtosis
Gold spot							
Actual time series	524.75	725.00	609.60	603.77	39.95	0.2123	2.5406
Mean of generated paths (3-state)	488.89	762.82	609.44	613.77	70.96	0.1681	2.3906
Mean of generated paths (1-state)	526.77	859.71	672.24	678.83	94.79	0.1701	1.9651
FTSE							
Actual time series	5506.8	6260.0	5896.8	5920.3	180.13	0.0443	2.1054
Mean of generated paths (3-state)	5281.9	6376.1	5805.0	5811.6	273.47	0.0502	2.4129
Mean of generated paths (1-state)	5637.2	6676.8	6136.0	6140.9	288.55	0.0607	2.1337

The mean, standard deviation, minimum and maximum of the resulting objective functions are depicted in Table 3. The value of the objective function is determined for ten different scenario sets of each size.

The optimal solutions and objectives, for the case of vector observations, are displayed in Table 4. All values of the objective function lie between 0.015058 and 0.027259 with a mean of 0.0193 and variance of $1.4093\text{e-}005$. The in-sample stability of the scenarios

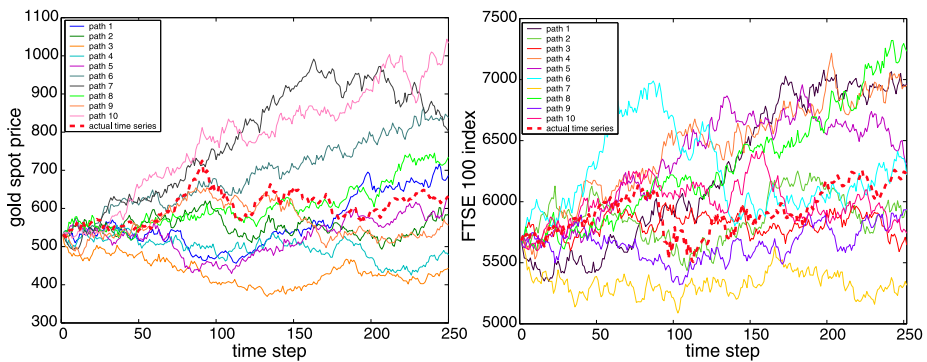


Fig. 5 Sample paths for gold spots and FTSE 100 index within 3-state framework

Table 2 Stability analysis for 10 scenario sets with independent observations

Scenario set	1	2	3	4	5	6	7	8	9	10
Objective func.	0.00333	0.00321	0.00296	0.00320	0.00322	0.00304	0.00300	0.00284	0.00293	0.00290
Investment in										
FTSE (in %)	9.77	17.64	9.14	16.88	15.27	12.03	15.84	10.38	13.63	8.81
Gold (in %)	24.37	20.71	22.26	21.80	22.76	23.08	21.08	21.44	21.80	22.38
Bond (in %)	65.96	61.65	68.60	61.32	61.97	64.89	63.08	68.18	64.57	68.81

Table 3 In-sample stability of scenario sets in independent setting

Objective function	Mean	Std dev	Min	Max
500 scen	0.00323	0.00083	0.00240	0.00491
1000 scen	0.00312	0.00044	0.00263	0.00387
2000 scen	0.00307	0.00029	0.00272	0.00352
5000 scen	0.00306	0.00016	0.00284	0.00332

generated through the vector setting is further examined through varying sizes of scenario sets. The descriptive statistics of the resulting objective functions are depicted in Table 5.

It is evident from the tables above that the scenario generator is more stable under the framework of independent observations. The stability under this framework is good, with respect to both the optimal percentages to be invested and the optimum. The similarity of optimal solutions guarantees a good out-of-sample stability of this scenario generator (i.e. stability with respect the “true”, unknown distribution of the random returns), which is one of the most important aspects. The examination of scenario sets of different sizes shows a good in-sample stability of both settings, however, the independent settings leads to lower standard deviations and a smaller range.

When generating scenarios through vector observations, there is a higher variability of both optima and optimal solutions. Economically this can be explained by the largely uncorrelated price of gold to any other stock. Due to its characteristics as a stable long-term investment and a global currency, gold prices have a low correlation to stock prices or in-

Table 4 Stability analysis for 10 scenario sets with vector observations

Scenario set	1	2	3	4	5	6	7	8	9	10
Objective func.	0.01993	0.02154	0.02154	0.01357	0.01793	0.02405	0.02148	0.02024	0.018912	0.02154
Investment in										
FTSE (in %)	21.66	35.36	21.66	47.54	41.43	33.15	21.38	56.16	48.25	21.66
Gold (in %)	52.34	45.34	52.34	23.14	41.40	62.41	55.62	42.63	41.83	52.34
Bond (in %)	26.00	19.30	26.00	29.32	17.17	4.44	23.00	1.22	9.93	26.00

Table 5 In-sample stability of scenario sets in vector setting

	Mean	Std dev	Min	Max
500 scen	0.0191	0.0090	0.0078	0.0357
1000 scen	0.0190	0.0061	0.0109	0.0319
2000 scen	0.0188	0.0045	0.0142	0.0261
5000 scen	0.0201	0.0028	0.0136	0.0241

Table 6 Descriptive statistics of generated paths in a one- and two-dimensional setting for Eurostoxx 50 and FTSE 100

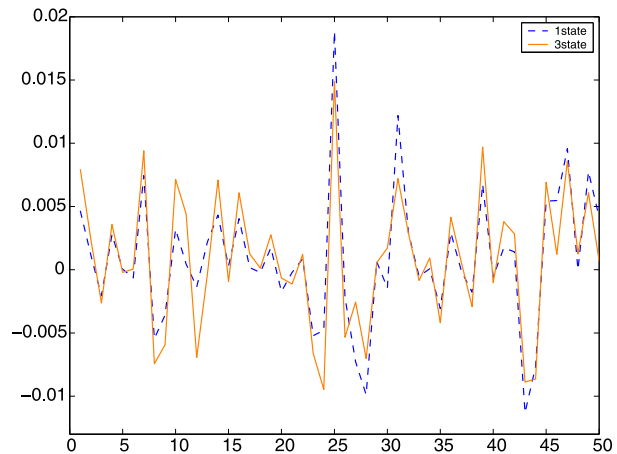
	Min	Max	Median	Mean	St.dev	Skewness	Kurtosis
Eurostoxx							
Actual time series	3408.0	4140.7	3792.0	3795.6	173.56	0.1351	2.2448
Mean of generated paths (2-dim)	3348.7	4274.3	3779.2	3787.7	235.56	0.1115	2.3728
Mean of generated paths (1-dim)	3591.0	4737.5	4119.4	4135.0	329.98	0.1074	1.9170
FTSE							
Actual time series	5506.8	6260.0	5896.8	5920.3	180.13	0.0443	2.1054
Mean of generated paths (2-dim)	5335.8	6572.0	5924.4	5930.2	315.13	0.0587	2.3782
Mean of generated paths (1-dim)	5281.9	6376.1	5805.0	5811.6	273.47	0.0502	2.4129

dices. Therefore optimal solutions calculated with independent generated scenarios are more stable.

4.4 The case of correlated assets

We extended our numerical study to two correlated indices, namely FTSE 100 and Eurostoxx 50. It has to be noted that also in that case where highly correlated indices are used in these scenario generation methods, the objective function reaches a smaller minimum with independently generated scenarios and the values of the objective functions are in smaller ranges. Nevertheless the distribution of generated paths for correlated indices are closer to the actual distributions observed in the market. Table 6 depicts the descriptive statistics of sets of 1,000 sample paths. Here we distinguish between sample paths generated from the hidden Markov model with vector observations (labelled ‘2-dim’) and sample paths generated independently for FTSE 100 and Eurostoxx 50. The characteristics of actual data is closely reached through paths generated with estimated parameters from the vector setting.

Fig. 6 Portfolio returns from rolling window backtest procedure



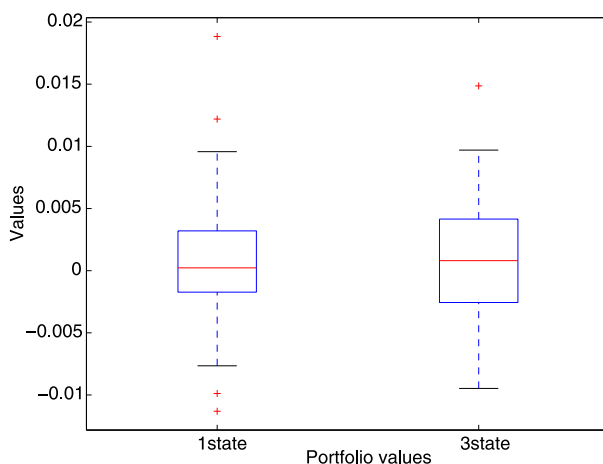
4.5 Backtesting

To backtest the performance of our scenario generator and compare it to a standard GBM setting, we perform a rolling window backtest procedure to compare the out-of-sample performance. For this, we use the historical datasets of gold spots and the FTSE 100 index to generate scenario sets for 50 time steps. Each scenario set consists of 5,000 scenarios. We generate scenario sets for gold spots and FTSE 100 from the regime-switching model with a 3-state Markov chain. Furthermore we generate scenario sets for both time series from the model with a 1-state Markov chain, which represents a discretised version of the geometric Brownian motion. For each of these 50 time steps, the optimisation is performed on the scenario sets from the regime-switching model as well as from the 1-state model, which does not allow switches between regimes. The optimal weights, which are obtained in these 100 optimisations (we obtain 50 sets of optimal weights for each model), are then utilised to create portfolios from the actual market data at that point in time.

The resulting portfolio values are plotted in Fig. 6. We see here, that the resulting out-of-sample portfolio values in these 50 time steps are similar in both models. Nevertheless, a statistical analysis of both sets of portfolio values shows an outperformance of the regime-switching model with 3 possible states. The mean of the portfolio return from the backtest procedure in a 3-state Markov chain framework is higher than that obtained through GBM scenarios, the distribution of portfolio returns over that time period is more symmetric and it contains less outliers. This can be seen in the boxplots in Fig. 7, the descriptive statistics are listed in Table 7.

4.6 The investment problem with weight constraints

In this section we modify our original investment problem and add constraints on the weights (fractions of capital) invested in the different assets. All weights now have an upper bound of 50%, i.e. no more than half of the budget is allowed to be invested in either gold spots, FTSE 100 stocks or bonds. Taking into account the results from the previous section, we only use the scenario generator that assumes independent observations. We generate five sets containing 5,000 scenarios and five sets containing 20,000 scenarios; we solve the con-

Fig. 7 Boxplots of both out-of-sample portfolio return series**Table 7** Descriptive statistics of portfolio return series

	Min	Max	Median	Mean	Standdev	Skew	Kur
Portfolio returns 1-state	-0.0113	0.0188	0.0002	0.0008	0.0053	0.5748	4.9389
Portfolio returns 3-state	-0.0095	0.0149	0.0008	0.0009	0.0054	0.0715	2.7587

Table 8 Investment problem with weight constraints, results for 10 scenario sets

Scenario set	5,000 generated scenarios					20,000 generated scenarios				
	1	2	3	4	5	1	2	3	4	5
Objective func.	0.00368	0.00354	0.00358	0.00346	0.00354	0.00352	0.00357	0.00352	0.00347	0.00352
Investment in										
FTSE (in %)	29.38	32.67	32.60	31.57	30.00	31.84	30.55	31.87	32.04	30.05
Gold (in %)	20.62	17.33	17.40	18.43	20.00	18.16	19.45	18.13	17.96	19.95
Bond (in %)	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00

strained mean-CVaR optimisation problem with this input. As it may be seen from Table 8, the stability of the solutions and objective is very good. The average objective function over all generated scenario sets is 0.003538409 with a very low variance of $3.77072\text{e-}09$. As already specified in the previous section, the similarity of the optimal solutions implies a good out-of-sample stability.

5 Conclusions

In this paper, we modelled financial time series using a hidden Markov model approach, whereby the parameters of the observation process (which we assumed to be a discretised Geometric Brownian motion) can switch between different economic regimes, that are not directly observable. Thus our method, while not departing fundamentally from traditional

approaches, tries to address a major shortcoming (the assumption that the parameters of the GBM are constant over time).

The parameters of the hidden Markov model are estimated using a filtering approach. Once the parameter estimation and the estimation of the Markov chain for the next time period are done, the model can be used for generating scenarios of future realisations for financial time series.

We tested the quality of the scenario generator in a financial decision making problem: a single period portfolio optimisation problem, in which the basket of assets is formed by FTSE 100, gold spots and a bank savings account. Scenarios are generated for the future returns of FTSE 100 and of gold spots and used as an input in a mean-CVaR optimisation problem. We investigated the following aspects: (a) the stability of the resulting scenario generator; (b) how the generated scenario paths compare to the realised (historical) price path; (c) the *realised* performance of the portfolio solutions obtained using this input.

When investigating the stability, we considered two frameworks: first, we assume that these time series are independent and follow different underlying Markov chains; secondly, we assume that the time series are governed by the same Markov chain. We concluded that the scenario generator based on independent observations gives a good in-sample stability.

For (b) and (c) above, we made comparisons with the scenario sets generated by assuming a standard GBM. We concluded that the HMM (3 states) gives a better representation of the data, especially with respect to unfavourable events. We performed a rolling window backtest; the results were better in the case of portfolio solutions obtained using the HMM scenarios. The focus of future research is to extend and test the scenario generator for multiple time step optimisation problems and to test this scenario generator in the context of multivariate (dependent) random variables.

Acknowledgements This work was carried out while Christina Erlwein was a NET ACE sponsored Intern at OptiRisk Systems; and Diana Roman was a KTP associate of Brunel University assigned to OptiRisk systems. The support of these organisations are gratefully acknowledged.

References

- Acerbi, C., & Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance*, 26, 1487–1503.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723.
- Artzner, P., Delbaen, F., Eber, J. M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3), 203–228.
- Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 2, 307–327.
- Elliott, R. (1994). Exact adaptive filters for Markov chains in Gaussian noise. *Automatica*, 30, 1399–1408.
- Elliott, R. J., Aggoun, L., & Moore, J. B. (1995). *Hidden Markov models: estimation and control*. New York: Springer.
- Ellison, E. F. D., Hajian, M., Levkovitz, R., Maros, I., & Mitra, G. (1999). *A Fortran based mathematical programming system FortMP*. Brunel University, Uxbridge, Oxford: Brunel University, NAG Ltd.
- Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation. *Econometrica*, 50(2), 987–1007.
- Erlwein, C., Mamon, R., & Davison, M. (2009). An examination of HMM-based investment strategies for asset allocation. *Applied Stochastic Models in Business and Industry*. doi:10.1002/asmb.820.
- Fourer, R., Gay, D. M., & Kernighan, B. (1989). *AMPL: a mathematical programming language*. AT&T Bell Laboratories.
- Geyer, A., & Ziemba, W. (2008). The innovest Austrian pension fund financial planning model InnoALM. *Operations Research*, 56(4), 797–810.

- Hamilton, J. D., & Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, 64(2), 307–333.
- Hardy, M. R. (2001). A regime-switching model of long-term stock returns. *North American Actuarial Journal*, 5(2), 41–53.
- Hoyland, K., Kaut, M., & Stein, W. (2003). A heuristic for moment-matching scenario generation. *Computational Optimization and Applications*, 24(2), 169–185.
- Kaut, M., & Stein, W. (2007). Evaluation of scenario-generation methods for stochastic programming. *Pacific Journal of Optimization*, 3(2), 257–271.
- Krokhmal, P., Palmquist, J., & Uryasev, S. (2002). Portfolio optimization with conditional value-at-risk objective and constraints. *The Journal of Risk*, 4(2), 21–41.
- Mamon, R., Erlwein, C., & Gopaluni, R. B. (2008). Adaptive signal processing of asset price dynamics with predictability analysis. *Information Sciences*, 178(1), 203–219.
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77–91.
- Messina, E., & Toscani, D. (2008). Hidden Markov models for scenario generation. *IMA Journal of Management Mathematics*, 19(4), 379–401.
- Mulvey, J. M. (2001). Introduction to financial optimization: mathematical programming special issue. *Mathematical Programming*, 89(2), 205–216.
- Pflug, G. C. (2000). Some remarks on the value-at-risk and the conditional value-at-risk. *Probabilistic Constrained Optimization: Methodology and Applications*, 38, 272–281.
- Rabiner, L. R. (1989). A tutorial on Hidden Markov Models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257–286.
- Rabiner, L. R. (2002). Hidden Markov processes. *IEEE Transactions on Information Theory*, 48(6), 1518–1569.
- Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *The Journal of Risk*, 2(3), 21–41.
- Rockafellar, R. T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, 26, 1443–1471.
- Roman, D., Mitra, G., & Spagnolo, N. (2010). Hidden Markov models for financial optimization problems. *IMA Journal of Management Mathematics*, 21(2), 111–129.
- Ross, S. M. (2002). *An elementary introduction to mathematical finance*. Cambridge: Cambridge University Press.
- Zakai, M. (1969). On the optimal filtering of diffusion processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 11, 230–243.
- Zhang, Y. (2004). *Prediction of financial time series with Hidden Markov Models*. M.Sc. Thesis, Simon Fraser University, China.