Ann. Inst. Statist. Math.
Vol. 53, No. 1, 50–62 (2001)

A MONTE CARLO FILTERING APPROACH FOR ESTIMATING THE TERM STRUCTURE OF INTEREST RATES

AKIHIKO TAKAHASHI[1] AND SEISHO SATO[2]

[1] *Graduate School of Mathematical Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, Japan*
[2] *The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan*

(Received April 17, 2000; revised August 10, 2000)

Abstract. We develop new methodology for estimation of general class of term structure models based on a Monte Carlo filtering approach. We utilize the generalized state space model which can be naturally applied to the estimation of the term structure models based on the Markov state processes. It is also possible to introduce measurement errors in the general way without any bias. Moreover, the Monte Carlo filter can be applied even to the models in which the zero-coupon bonds' prices can not be analytically obtained. As an example, we apply the method to LIBORs (London Inter Bank Offered Rates) and interest rates swaps in the Japanese market and show the usefulness of our approach.

Key words and phrases: Generalized state sapce model, Monte Carlo integration, interest rate model, self-organizing method.

1. Introduction

We propose a new framework of the estimation of the term structure of interest rates based on the state space model. In particular, we develop a Monte Carlo filtering approach for estimating the term structure models based on multi-dimensional Markov state variables. For example, our method can be applied to the term structure models based on the dynamic general equilibrium theory of Cox *et al.* (1985*a*, 1985*b*) which includes multi-factor CIR (Cox, Ingersoll and Ross) models used by Chen and Scott (1993), Pearson and Sun (1994), Singh (1995) and Duffie and Singleton (1997), and the stochastic volatility model developed by Longstaff and Schwartz (1992). It is well-known that at least two state variables are necessary to explain the dynamics of the term structure in the real world. Chen and Scott (1993), Pearson and Sun (1994), Singh (1995), and Duffie and Singleton (1997) concluded that one-factor models are not enough to describe the variation of the term structure by analyzing treasury or swap markets in the United States. However, multi-factor models often tested in the analysis are not necessarily the best among the candidates. For instance, in the multi-factor CIR model where the spot interest rate is described by the sum of several state variables independently following square-root processes, the intuitive interpretation of the state variables is not clear, and sometimes it seems difficult to find admissible parameters for which the state variables are non-negative over the entire sample period (see Duffie and Singleton (1997)). One of the reasons why the models which do not explain the data very well are often employed in empirical analyses is that they allow analytic solutions of zero coupon bonds' prices.

It is mainly due to the limitation of the methods applied to the estimation. In particular, Chen and Scott (1993), Pearson and Sun (1994) and Duffie and Singleton (1997) apply the maximum likelihood method while Singh (1995), and Longstaff and Schwartz (1992) employ the three-stage least square method with the principal component analysis and the generalized moment method (GMM) with GARCH, respectively. However, it is substantially difficult to apply those methods without analytic solutions of the zero-coupon bonds' prices. Moreover, existing researches tend to replace unobservable state variables such as the spot rate and the volatility by some observable variables (Chan *et al.* (1992), Longstaff and Schwartz (1992)), but they may substantially suffer from the measurement errors. While some of them take the measurement errors into account explicitly, the ways of the consideration are not natural and somewhat ad hoc (Chen and Scott (1993), Duffie and Singleton (1997)).

We propose a Monte Carlo filtering approach based on the generalized state space model to overcome the problems of existing researches. The state space model consists of the system model describing the processes of state variables and the observation model representing the functional relation between the state variables and the observational data in the real world, which implies that the method can be naturally applied to the estimation of the term structure models based on the Markov state processes. It is also possible to introduce measurement errors in the general way without any bias. Moreover, the Monte Carlo filter can be applied to much broader class of the term structure models, especially even to the models in which the zero-coupon bonds' prices can not be analytically obtained.

The paper is organized as follows. In Section 2, we will first summarize the state space model and the term structure models based on Markov state processes. Then, we will clarify the relation between interest rate models as well as observational data and the state space model. Next, we will give a concrete algorithm of the Monte Carlo filter applied to the empirical analysis. In Section 3, we will show the usefulness of the Monte Carlo filter by the analysis of LIBORs (London Inter Bank Offered Rates) and interest rate swaps in the Japanese market. In Section 4, we will give the conclusion.

2. The estimation of the term structure based on the state space modeling

2.1 *State space model for term structure*

We explain in this section the estimation method for the common factors and the parameters of term structure models based on the state space modeling.

First, we briefly explain term structure models based on Markov state processes (see Björk (1996), Duffie (1996) or Hull (1999) for the detail). Given the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$ with the time horizon $[0, T^*]$ for some $T^* < \infty$, we suppose that a N-dimensional vector of state variables denoted by Y_t follows a N-dimensional Markov process:

$$dY_t = \mu(Y_t, t)dt + S(Y_t, t)dB_t, \tag{2.1}$$

where B_t is the d-dimensional standard Brownian motion under the filtered probability space, and $\mu(Y_t, t)$ and $S(Y_t, t)$ denote real-valued functions of $R^N \times [0, T^*] \mapsto R^N$ and $R^N \times [0, T^*] \mapsto R^{N \times d}$, respectively. Suppose also that the instantaneous short-term interest rate at time t denoted by $r(Y_t, t)$ and a zero coupon bond's price at t with the maturity T denoted by $P(Y_t, t; T)$ are some functions of Y_t where $t \in [0, T^*]$ and $T \in [t, T^*]$. Here, a zero coupon bond with the maturity T means a bond with no coupons and with the face value, 1 which is redeemed at time T. We note that the set of the zero coupon bonds' prices $\{P(Y_t, t; T)\}_{T \in [t, T^*]}$ represents the term structure of interest rates at time t. The assumption reflects the idea that the whole term structure can be explained by relatively small number of factors, Y_t while what the factors represent depends on the specification of a model.

Then, based on the arbitrage-free argument of financial economics $P(Y_t, t; T)$ satisfies a partial differential equation (PDE)

$$\frac{1}{2}\,\mathrm{trace}(SS^{'}P_{YY}) + [\mu - \phi]'P_Y + P_t - rP = 0,$$

with the terminal boundary condition, $P(Y_t, T; T) = 1$, where $P_{YY} \equiv \frac{\partial^2 P}{\partial Y \partial Y'}$, $P_Y \equiv \frac{\partial P}{\partial Y}$. Here, R^N-valued vector, ϕ denotes so called the risk premium which is a function of Y_t and t, $\phi(Y_t, t)$ and for instance, it can be determined by the general equilibrium asset pricing theory of economics such as CIR (Cox *et al.* (1985*a*)). Moreover, it is well known that the solution of this PDE is represented by the conditional expectation given information at time t,

$$P(Y_t, t; T) = \boldsymbol{E}^Q[e^{-\int_t^T r(Y_u, u)du} \mid \mathcal{F}_t], \tag{2.2}$$

where $\boldsymbol{E}^Q[\cdot \mid \mathcal{F}_t]$ denotes the conditional expectation operator given information at time t under the risk-neutral probability measure Q (see Björk (1996)). It is also known that under the measure Q, the vector of state variables Y_t follows a stochastic differential equation,

$$dY_t = \{\mu(Y_t, t) - \phi(Y_t, t)\}dt + S(Y_t, t)dB_t^* \tag{2.3}$$

where B^* denotes the d-dimensional standard Brownian motion under the measure Q (see Björk (1996) and Chapter 7 and Appendix E of Duffie (1996)).

For application of these term structure models to observed data, we introduce the general form of state space models (see Kitagawa and Gersh (1996) for the detail). A state space model consists of the following system model and the observation model. That is,

$$\begin{cases} Y_t = \boldsymbol{F}(Y_{t-\Delta t}, v_t) & \text{system model} \\ Z_t = \boldsymbol{H}(Y_t, u_t) & \text{observation model} \end{cases} \tag{2.4}$$

where Y_t, Z_t and Δt denote a N-dimensional state vector, a M-dimensional observation vector at time t and the time interval of observational data respectively while v_t and u_t denote the N-dimensional system noise and the M-dimensional observational noise whose density functions are given respectively by $q(v)$ and $\psi(u)$. $\boldsymbol{F}(\cdot,\cdot)$ and $\boldsymbol{H}(\cdot,\cdot)$ are generally non-linear functions, and the initial state vector Y_0 is assumed to be a random variable whose density function is given by $p_0(Y)$.

Next, we clarify how the state space models can be applied to the estimation of term structure models. When Δt is sufficiently small, the Euler approximation to the equation (2.1) can be used for the system model, $Y_t = \boldsymbol{F}(Y_{t-\Delta t}, v_t)$. That is,

$$Y_t = Y_{t-\Delta t} + \mu(Y_{t-\Delta t}, t - \Delta t)\Delta t + S(Y_{t-\Delta t}, t - \Delta t)v_t\sqrt{\Delta t} \tag{2.5}$$

where the system noise v_t follows the N-dimensional standard normal distribution. Of course, the other approximation schemes could be applied to the discretization of (2.1) (for instance, see section D and notes of chapter 11 in Duffie (1996)). Moreover, when Y_t

is explicitly solved given $Y_{t-\Delta t}$ as in the case of a linear stochastic differential equation, it is better to use that representation: Namely, suppose that in (2.1), Y_t is represented by a linear stochastic differential equation.

$$dY_t = (AY_t + \beta^*(t))dt + SdB_t \tag{2.6}$$

where $\beta^*(t)$ and A denote R^N-valued functions of the time parameter t and $N \times N$ matrix with constant elements respectively, and S denotes an $N \times d$ matrix with constant elements so that $\Sigma = SS'$ is positive definite. Then, given $Y_{t-\Delta t}$, Y_t can be expressed as

$$\begin{aligned} Y_t &= e^{\Delta t A} Y_{t-\Delta t} + \int_{t-\Delta t}^{t} e^{(t-s)A}\beta^*(s)ds + v_t^{(\Delta t)} \\ &= FY_{t-\Delta t} + \beta(t) + v_t^{(\Delta t)} \end{aligned} \tag{2.7}$$

where $F \equiv e^{\Delta t A}$ is an $N \times N$ matrix with constant elements and $\beta(t) \equiv \int_{t-\Delta t}^{t} e^{(t-s)A}\beta^*(s)ds$ is an $N \times 1$ vector function of time t. Here, $v_t^{(\Delta t)}$ follows the normal distribution with the mean zero and the variance covariance matrix $\Sigma_{\Delta t}$ defined by

$$\Sigma_{\Delta t} \equiv \int_0^{\Delta t} e^{sA}\Sigma e^{sA'} ds. \tag{2.8}$$

In this case, the system model is given by (2.7) and the density function $q(v)$ of the system noise v_t is the normal distribution with the variance covariance matrix specified by (2.8).

In the observation model, the vector of the observation at time t denoted by Z_t, can be expressed as a function of $k(\geq 1)$ units of zero-coupon bonds' prices and the observation noise vector u_t

$$Z_t = h(P(Y_t, t; t+T_1), \ldots, P(Y_t, t; t+T_k)) + u_t. \tag{2.9}$$

That is, each element of Z_t is an observed bond price or interest rate which is represented by a function of zero coupon bonds with possibly different maturities $(T_i, i = 1, \ldots, k)$ and a measurement error. Moreover, Z_t can be also expressed as a function of Y_t, because each $P(Y_t, t; T_i)$ is a function of Y_t,

$$Z_t = H(Y_t) + u_t. \tag{2.10}$$

In $h(\cdot)$, $P(Y, t; t + T_i)$ can be evaluated by the computation of the equation (2.2) under the process (2.3). In addition, we assume hereafter that the density function $\psi(u)$ of the noise vector u is given by that of the multi-dimensional normal distribution with the mean 0, and the variance-covariance matrix Σ_u. Here, Σ_u is assumed to be a $M \times M$ diagonal matrix with positive diagonal elements where M is defined to be the dimension of the observational data at each time.

LIBORs and interest rate swap rates at time t are the typical examples of the observation vector Z_t. In this case, $h(\cdot)$ can be specified by the theoretical relation between LIBORs/interest rate swap rates and the zero-coupon bonds' prices. That is, LIBOR with the term τ_n denoted by $L_t(Y_t, \tau_n)$ and swap rates with the term τ_n denoted by $S_t(Y_t, \tau_n)$ are expressed by the zero-coupon bonds' prices as follows (see for instance

Björk (1996), Duffie (1996) and Duffie and Singleton (1997)).

$$L_t(Y_t, \tau_n) = \left(\frac{1}{P(Y_t, t; t + \tau_n)} - 1 \right) \frac{1}{\tau_n} \tag{2.11}$$

$$S_t(Y_t, \tau_n) = \frac{1 - P(Y_t, t; t + \tau_n)}{\delta \sum_{i=1}^{\tau_n/\delta} P(Y_t, t; t + i\delta)}, \tag{2.12}$$

where δ denotes the interval of cash flows. For example, $\delta = 0.5$ is standard in the swap market of Japanese yen. We will explain actual data of LIBORs and swap rates used for the empirical analysis in the next section.

From the discussion above, we show that a term structure model represented by (2.1)–(2.3) can be re-interpreted within the framework of the generalized state space form (2.4). Hence we can obtain the estimates of Y_t in a term structure model by estimating states in the corresponding state space model.

2.2 *Estimation of term structure by the Monte Carlo filter*

We discuss about our estimation method in detail. We note that the standard Kalman filter cannot be applied to the estimation because both the system model and the observation model described above are generally non-linear. Thus we utilize the Monte Carlo filter. While several approaches are proposed for the Monte Carlo filter (see Doucet *et al.* (1995), Durbin and Koopman (1997), Gordon *et al.* (1993), Tanizaki (1993), for instance), we adopt the approach developed by Kitagawa (1996). In the following, we describe the outline of the algorithm of the Monte Carlo filter applied to the empirical analysis in the next section. First, we summarize the notation following Kitagawa (1996). $p(Y_t \mid Z_{t-\Delta t})$, called "one step ahead predictor" denotes the conditional density function of Y_t given $Z_{t-\Delta t}$ where Δt is the interval of time series. $p(Y_t \mid Z_t)$, called "filter" denotes the conditional density function of Y_t given Z_t. $\{p_t^{(1)}, \dots, p_t^{(m)}\}$ and $\{f_t^{(1)}, \dots, f_t^{(m)}\}$ represent the vectors of the realization of m trials of Monte Carlo from $p(Y_t \mid Z_{t-\Delta t})$ and $p(Y_t \mid Z_t)$, respectively. Then, if we set $\{f_0^{(1)}, \dots, f_0^{(m)}\}$ as the realization of random draws from $p_0(Y)$, the density function of the initial state vector Y_0, the algorithm of the Monte Carlo filter is as follows.

[The algorithm of the Monte Carlo filter]

(i) Generate the initial state vector $\{f_0^{(1)}, \dots, f_0^{(m)}\}$.

(ii) Apply the following steps (a)–(d) to each time $t = 0, \Delta t, 2\Delta t, \dots, (T_* - \Delta t), T_*$ where T_* denotes the final time point of the data.

(a) Generate the system noise $v_t^{(j)}$, $j = 1, \dots, m$ according to the density function $q(v)$.

(b) Compute for each $j = 1, \dots, m$, $p_t^{(j)} = F(f_{t-\Delta t}^{(j)}, v_t^{(j)})$.

(c) Evaluate $\alpha_t^{(j)} = n[x_t; 0, \Sigma_u]$, where $n[x_t; 0, \Sigma_u]$ the density function of $N(0, \Sigma_u)$ at $x_t = Z_t - H(p_t^{(j)})$, for $j = 1, \dots, m$. Here, $H(\cdot)$ represent for instance, the equation (2.11) and (2.12) which are regarded as functions of the state vector Y. The prices of zero-coupon bonds in those equations are computed through the equation (2.2) by using the process (2.3), and if it is not evaluated analytically, some numerical method such as Monte Carlo simulation is implemented.

(d) Implement resampling of $\{f_t^{(1)}, \ldots, f_t^{(m)}\}$ from $\{p_t^{(1)}, \ldots, p_t^{(m)}\}$. More precisely, obtain $f_t^{(i)}$, $i = 1, \ldots, m$ by the sampling with replacement from $\{p_t^{(1)}, \ldots, p_t^{(m)}\}$ with the probability

$$\text{Prob.}(f_t^{(i)} = p_t^{(j)} \mid Z_t) = \frac{\alpha_t^{(j)}}{\sum_{k=1}^{m} \alpha_t^{(k)}}, \qquad j = 1, \ldots, m, \quad i = 1, \ldots, m.$$

The estimation of unknown parameters is based on the maximum likelihood method. If μ denotes the vector representing all unknown parameters, the log-likelihood $l(\mu)$ is given by

$$l(\mu) = \log p(Z_{\Delta t}, \ldots, Z_{T_*} \mid \mu) = \sum_{k=1}^{T_*/\Delta t} \log p(Z_{k\Delta t} \mid Z_{\Delta t}, \ldots, Z_{(k-1)\Delta t}, \mu)$$

where $p(Z_{\Delta t} \mid Z_0) = p_0(Z_{\Delta t})$. Since each term in the log-likelihood can be approximated as

$$p(Z_{k\Delta t} \mid Z_{\Delta t}, \ldots, Z_{(k-1)\Delta t}, \mu) \simeq \frac{1}{m} \sum_{j=1}^{m} \alpha_{k\Delta t}^{(j)},$$

the log-likelihood $l(\mu)$ is computed within the framework of the Monte Carlo filter by

$$l(\mu) \simeq \sum_{k=1}^{T_*/\Delta t} \left(\log \sum_{j=1}^{m} \alpha_{k\Delta t}^{(j)} \right) - \frac{T_*}{\Delta t} \log m.$$

Then, the maximum likelihood estimator $\hat{\mu}$ is obtained by maximizing $l(\mu)$ with respect to μ. For maximization, grid search or a self-organizing method is applied. (See Kitagawa (1998) for details of a self-organizing state-space model.) Finally, we utilize AIC (Akaike Information Criterion) as a criterion to select the term structure model if there are several candidates. That is, the model with the smaller AIC can be regarded as the better model (Akaike (1973)).

3. Analysis of the term structures in the Japanese market

In this section, we examine the validity of our method using the time series of interest rates in the Japanese market. The data used for the analysis is summarized as follows.

- The period and the frequency of the data: daily data of 1997/1/1–1999/7/22 (662 observations).
- Japanese yen LIBORs: six-month, twelve-month.
- Japanese yen swap rates: two-year, three-year, four-year, five-year, seven-year, ten-year.

Figures 1(1)–(3) show the observational data for the period of the analysis; (1) and (2) show the series of LIBORs and those of swap rates respectively while (3) shows the spread between ten-year and two-year.

For an interest rate model, we use Hull and White (1994) in which the dynamics of a state vector Y can be represented by a linear stochastic differential equation (2.6), where $Y_t = (Y_{it})$, $i = 1, 2$ is a two dimensional state vector and $\beta^*(t)$ is a R^2-valued

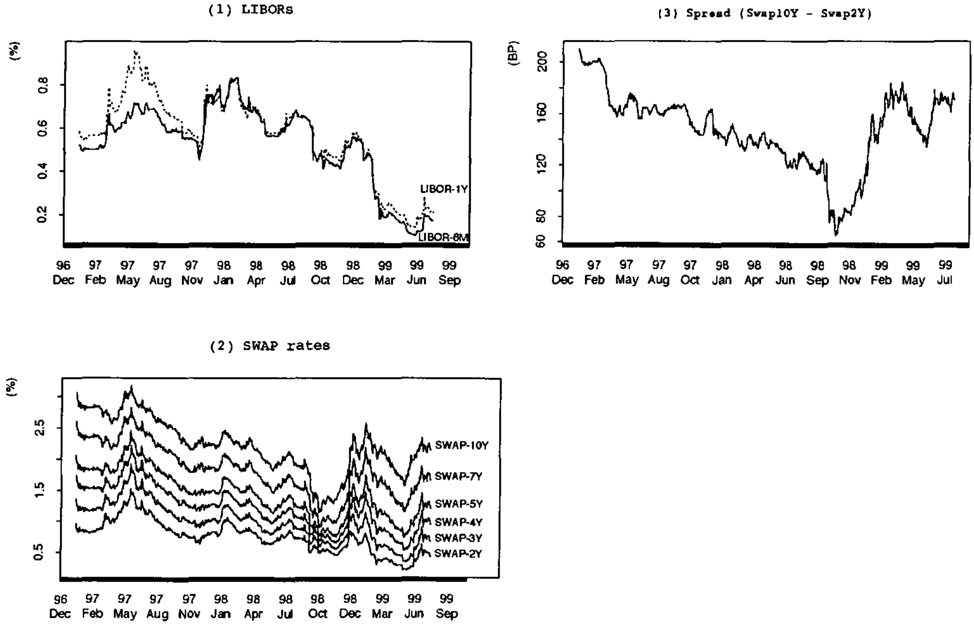
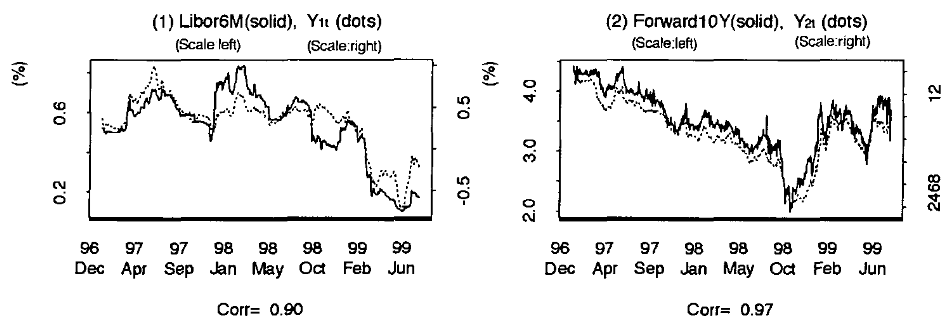


Fig. 1. Time series of observed LIBORs and swap rates.



– **Parameter estimates:**

$$A = \begin{pmatrix} -0.08 & 0.08 \\ 0 & -0.13 \end{pmatrix}, \Sigma = \begin{pmatrix} 0.32 \times 10^{-4} & -0.16 \times 10^{-3} \\ -0.16 \times 10^{-3} & 0.17 \times 10^{-2} \end{pmatrix},$$

$$\Sigma_u = \mathrm{diag}(0.11 \times 10^{-7}, 0.56 \times 10^{-8}, 0.78 \times 10^{-9}, 0.63 \times 10^{-9}, 0.63 \times 10^{-9}, 0.53 \times 10^{-9}, 0.58 \times 10^{-9}, 0.16 \times 10^{-8}).$$

Fig. 2. Estimated factors (two-factor case).

function of the time parameter t. They also assume that the spot rate r is expressed as a function of Y_{1t}, $r = g(Y_{1t})$ where $g(\cdot)$ is some real-valued function. For a functional form of $r = g(Y_{1t})$, if we take $g(Y_{1t}) = Y_{1t}$, we allow negative interest rates because of the normality of the spot rate while we can obtain an analytic solution of $P(Y_t, t; T)$, which substantially reduce the computational burden. In particular, the model implies relatively high probability of the negative interest rates in such low interest rates environment of recent Japan and this seems inappropriate. Hence, we specify $g(Y_{1t})$ based

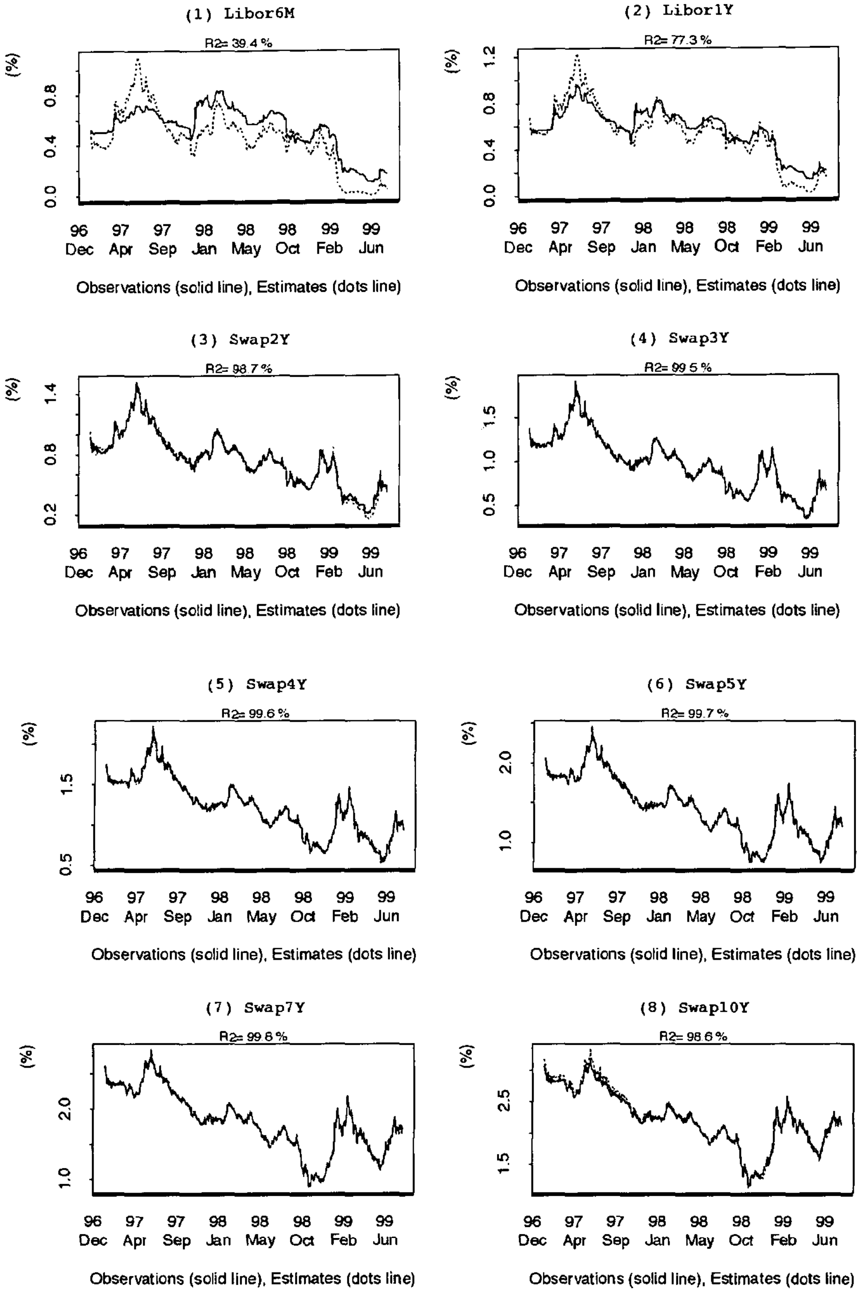


Fig. 3. Observations and estimates (two-factor case); the explanation power (R2) is defined by $\max\left[\left(1 - \frac{\text{the variance of residuals}}{\text{the variance of observations}}\right), 0\right]$ (%).

on Hull and White ((1997), chapter 21, pp. 588) such that $g(Y_{1t}) = Y_{1t}$ for $Y_{1t} \geq \varepsilon$, and $g(Y_{1t}) = \varepsilon e^{\frac{(Y_{1t}-\varepsilon)}{\varepsilon}}$ for $Y_{1t} < \varepsilon$, where ε is some predetermined positive constant. Clearly, $g(\cdot)$ is positive, monotonically increasing, and $\lim_{y \to -\infty} g(y) = 0$. We note that from above specification, Y_{1t} can be considered to be a factor which has a large impact on the short-term sector of the term structure. On the other hand, Y_{2t} will be characterized after the fitting of the model to the data.

We next determine the observation model as the equation (2.9) with the equations (2.11) and (2.12). Moreover, the zero-coupon bonds' prices in $h(\cdot)$ of (2.9) are computed by the equation,

$$P(Y_t, t; T) = E^Q[e^{-\int_t^T g(Y_{1u})du} \mid Y_t]. \tag{3.1}$$

We note that (3.1) should be computed by some numerical method such as Monte Carlo simulations because it can not be evaluated analytically. We apply the algorithm of the Monte Carlo filter described in the previous section to the estimation. The state space model applied to this empirical analysis is described by

(The system model)

$$Y_t = FY_{t-\Delta t} + \beta(t) + v_t^{(\Delta t)}, \tag{3.2}$$

with $Y_t = (Y_{1t}, Y_{2t})'$ and

(The observational model)

$$Z_t = (Z_{1,t}, \ldots, Z_{8,t})' \tag{3.3}$$

with

$$Z_{n,t} = \begin{cases} L_t(Y_t, \tau_n) + u_{n,t} & (\text{for } n = 1, 2, \text{ with } \tau_n = 0.5, 1) \\ S_t(Y_t, \tau_n) + u_{n,t} & (\text{for } n = 3, \ldots, 8, \text{ with } \tau_n = 2, 3, 4, 5, 7, 10), \end{cases} \tag{3.4}$$

where the dimension of the observational data is $M = 8$.

In the computation of $P(Y_t, t, T)$ in $L_t(Y_t, T_n)$ and $S_t(Y_t, T_n)$, we apply a numerical integration for the integral, $\int_t^T g(Y_{1u})du$ in (3.1) and a Monte Carlo integration for the conditional expectation: Namely,

$$P(Y_t^{(i)}, t : T) \simeq \frac{1}{J} \sum_{j=1}^{J} \exp\left(- \sum_{l=0}^{T/\Delta t} g(Y_{1,t+l\Delta t}^{(i,j)})\Delta t \right), \tag{3.5}$$

where $Y_t^{(i)}$ denotes the value of Y_{1t} for the i-th particle of the state vector in the Monte Carlo filter, and $Y_{1,t+l\Delta t}^{(i,j)}$ denotes the value of $Y_{1,t+l\Delta t}$ in the j-th path of the state vector starting from $Y_t^{(i)}$, which are generated from

$$Y_{t+l\Delta t}^{(i,j)} = FY_{t+(l-1)\Delta t}^{(i,j)} + \beta(t + l\Delta t) + v^{(i,j)}(t + l\Delta t), \qquad Y_t^{(i,j)} = Y_t^{(i)}.$$

In the Monte Carlo integration, $J = 300$, the random numbers are generated by ran2 in *Numerical Recipes* (2nd ed.) by Press *et al.* (1992) and the method of antithetic variates (Kalos and Whitlock (1986)) is utilized, in which we set $v_{t+l\Delta t}^{(i,j)} = -v_{t+l\Delta t}^{(i,j-1)}$, $l = 1, 2, \ldots, (T/\Delta t)$ for even j (i.e. $j = 2, 4, \ldots, 300$).

Moreover, the number of particles in the Monte Carlo filter is $m = 5000$ and we use 80 parallel computer for this computation (SGI 2800 system).

In estimation, it is hard to implement standard numerical optimization methods such as quasi-Newton method because computational difficulty arises due to the nonlinearity and large number of parameters in the model. Hence, for ε in $g(\cdot)$ we set $\varepsilon = 0.0005$ which is smaller than the lowest level of six-month LIBOR observed during the sample period so that the choice of ε does not have large impact on the shape of the yield curve, and for the other parameters we adopt the following estimation method.

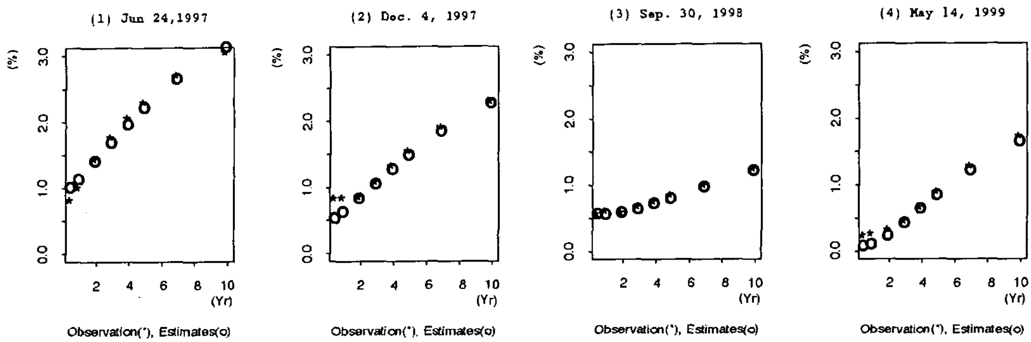


Fig. 4. The term structures (two-factor case).

*Step*1. Apply Kalman-Filter as if the interest rate function were $g(r) = r$.
*Step*2. Apply the self-organizing method using the estimates obtained in Step1 as initial values.
*Step*2'. Implement the self-organizing method again setting the average of each parameter over the sample period as the initial value. The procedure is iterated until the log-likelihood is not significantly improved because slight improvement is not reliable due to the random nature of the estimated log-likelihood function.
*Step*3. Apply a grid search around the estimates obtained in Step2' until the log-likelihood is not significantly improved.

Here, the self-organizing method was proposed by Kitagawa (1998); in this method, the original state vector, Y_t is augmented with the unknown parameter vector, μ as $Y_t^* = (Y_t, \mu)'$, and we assume μ is time-varying such that $\mu_t = \mu_{t-\Delta t} + v_t^{(\mu)}$ where $v_t^{(\mu)}$ is normally distributed (for example, see Higuchi and Kitagawa (2000) or Kitagawa and Sato (2000)).

Finally, we briefly explain the result of this analysis. First we investigate the time series of the estimated factors. Figures 2(1)–(2) show the relation between the estimated factors Y_{it}, $i = 1, 2$ obtained by averaging all the samples of filters and the forward rates computed from observed LIBORs and swaps. The sample correlation between the level of each factor and the corresponding forward rate is also listed below of each figure. From the graphical observation and the sample correlation, there seems to be strong relation between Y_{2t} and ten-year forward rate which is defined to be the rates with the term 0.5 year starting from 9.5 years forward, and the variation of Y_{1t} is similar to that of six-month LIBOR. Next, we examine the fitting of the model to the data. The time series of each observation and corresponding estimate are shown in Figs. 3(1)–(8), and the explanation power of each estimate which is defined by $\max[(1 - \frac{\text{the variance of residuals}}{\text{the variance of observations}}), 0](\%)$ is listed above corresponding figure. We note that the estimates of LIBORs and swap rates are based on estimated parameters and estimated factors. They show that the model fits very well to the swap rates while it does not to LIBORs: The explanation powers are more than 98.5% for all the swap rates while they are 39.4% and 77.3% for six-month and twelve-month LIBORs respectively.

In the two-factor model, estimated term structures are generally fitted well to the real ones except to LIBORs. In Figs. 4(1)–(4), we show the examples in which the fitting to LIBORs are good (9/30/98), average (6/24/97), and bad (12/4/97, 5/14/99). Hence,

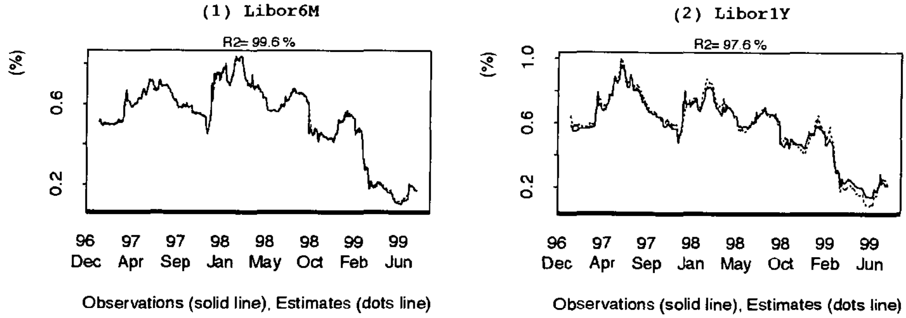


Fig. 5. Observations and estimates of LIBORs (three-factor case).

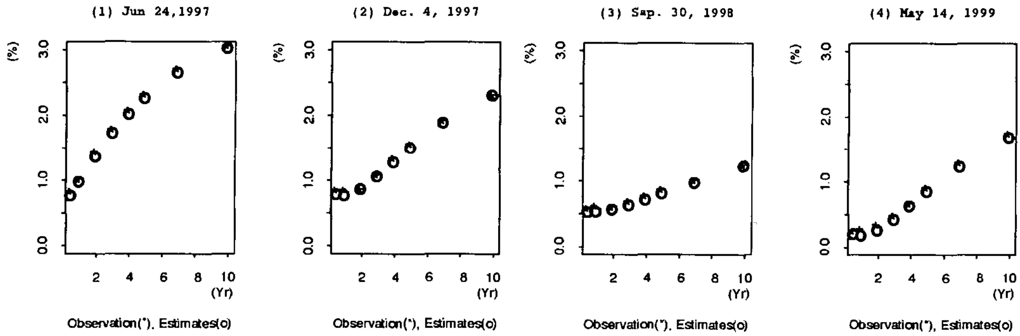


Fig. 6. The term structures (three-factor case).

we can conclude that the model explains the variations of the swap rates very well while it does not explain those of LIBORs.

The result of the two-factor case implies that another factor may be necessary to improve the fitting to LIBORs. Hence, we next implement the case in which the state vector $Y_t = (Y_{it})$, $i = 1, 2, 3$ is three dimensional, and has the same form of (3.2).

We report that the explanation powers are more than 97.5% for all the rates and more than 99% except for twelve-month LIBOR and two-year swap rate. Especially, we note that the model remarkably improves the fitting to LIBORs (see Fig. 5). Figures 6(1)–(4) show the observations and estimates of the term structures at four dates (97/6/24, 97/12/4, 98/9/30, 99/5/14, which are same as the two-factor case), which implies that the model can replicates the real term structures including LIBORs very closely. Moreover we note that AIC in this case (AIC$=-72921.23$) is substantially improved by more than 7000. However, when we implement more subtle method of model diagnostics described in Kim *et al.* (1998), it turns out that there are some problems for the innovation of the model. Especially, the independence of the innovations and the normality of the transformed innovations are rejected by using Box-Ljung statistic and Bowman and Shenton (1975) normality statistic. We show the results for six-month LIBOR and five-year swap rate in Fig. 7. Thus it seems necessary to investigate the broader class of interest rate models.

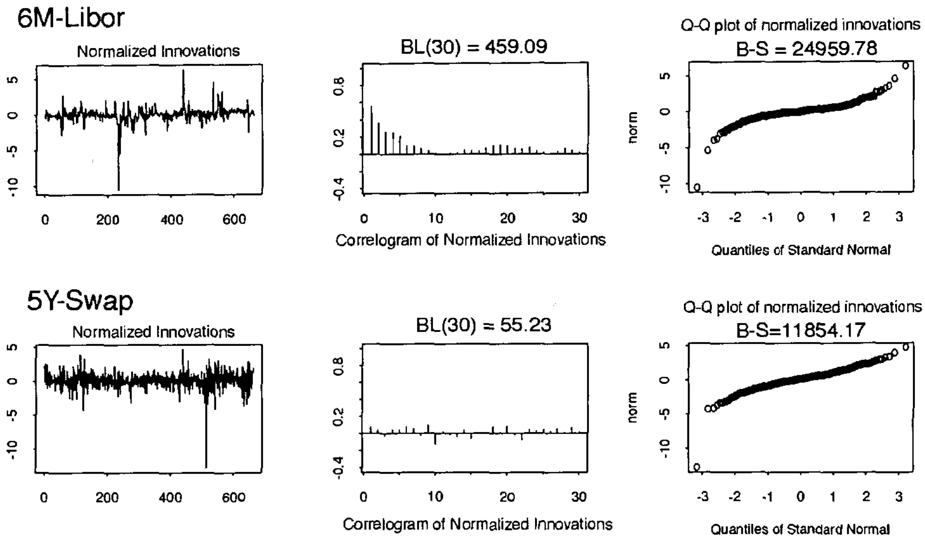


Fig. 7. The result of model diagnostics for three-factor case (we show the results for only six month LIBOR and five year swap), where innovations are computed by $u_{t+1}^M = \Pr(Z_{k,t+1} \leq Z_{k,t+1}^o \mid Z_t) \simeq \frac{1}{M}\sum_{i=1}^M \Pr(Z_{k,t+1} \leq Z_{k,t+1}^o \mid Y_{t+1}^{(i)})$, normalized innovations are $n_t^M = \Omega^{-1}(u_t^M)$ (the inverse of the normal distribution), $Z_{k,t+1}^o$ means observed value, and BL(30) and B-S is denoted as Box-Ljung statistic with 30 lags and Bowman and Shenton (1975) normality statistic ($\sim \chi^2(2)$), respectively (see Kim *et al.* (1998)).

4. Conclusion

We develop a new framework for the empirical analysis of the term structure of interest rates based on the generalized state space model. Our approach is useful for the estimation and the model diagnostics of various types of interest rate models, which are usually considered tough task. As an example, we apply the Monte Carlo filter to the time series of LIBORs and of interest rate swaps in the Japanese market, and confirm the validity of our method.

Furthermore, we will utilize this approach to analyze more complicated models.

Acknowledgements

We thank editors and anonymous referees for useful comments on the previous version.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle, *Second International Symposium on Information Theory* (eds. B. N. Petrov and F. Csáki), 267–281, Akademiai Kiádo, Budapest.

Björk, T. (1996). Interest rate theory, *Financial Mathematics Bressanone* (eds. W. J. Runggaldier), Springer, Belrin.

Bowman, K. O. and Shenton, L. R. (1975). Omnibus test contours for departures from normality based on $\sqrt{b_1}$ and b_2, *Biometrika*, **62**, 243–250.

Chan, K. C., Karolyi, G. A., Longstaff, F. A. and Sanders, A. B. (1992). An empirical comparison of alternative models of the short-term interest rate, *Journal of Finance*, **47**, 1209–1227.

Chen, R. and Scott, L. (1993). Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates, *Journal of Fixed Income*, December, 14–31.

Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985*a*). An intertemporal general equilibrium model asset prices, *Econometrica*, **53**, 363–384.

Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985*b*). A theory of the term structure of interest rates, *Econometrica*, **53**, 385–408.

Doucet, A., Barat, E. and Duvaut, P. (1995). A Monte Carlo approach to recursive Bayesian state estimation, *Proceedings IEEE Signal Processing/Athos Workshop on Higher Order Statistics*, Girona, Spain.

Duffie, D. (1996). *Dynamic Asset Pricing Theory*, 2nd ed., Princeton University Press, Princeton, New Jersey.

Duffie, D. and Singleton, K. J. (1997). An econometric model of the term structure of interest-rate swap yields, *Journal of Finance*, **52**, 1287–1321.

Durbin, J. and Koopman, S. J. (1997). Monte Carlo maximum likelihood estimation for non-Gaussian state space models, *Biometrika*, **84**, 669–684.

Gordon, N., Salmond, D. J. and Smith, A. F. M. (1993). Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proceedings-F*, **140** (2), 107–113.

Higuchi, T. and Kitagawa, G. (2000). Knowledge discovery and self-organizing state space model, *IEICE Transactions on Information and Systems*, **E83-D** (1), 36–43.

Hull, J. (1999). *Options, Futures, and Other Derivatives*, 4th ed., Prentice-Hall, Upper Saddle River, New Jersey.

Hull, J. and White, A. (1994). Numerical procedures for implementing term structure models II: Two-factor models, *Journal of Derivatives*, **2**, 37–48.

Hull, J. and White, A. (1997). Taking rates to the limits, *Risk*, December, 1997, 168–169.

Kalos, M. H. and Whitlock, P. A. (1986). *Monte Carlo Methods Volume I*, Wiley-Interscience Publication, New York.

Kim, S., Shephard, N. and Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH model, *Rev. Econom. Stud.*, **65**, 361–393.

Kitagawa, G. (1996). Monte Carlo filter and smoother for non-Gaussian nonlinear state space models, *J. Comput. Graph. Statist.*, **5** (1), 1–25.

Kitagawa, G. (1998). A self-organizing state-space model, *J. Amer. Statist. Assoc.*, **93**, 443, 1203–1215.

Kitagawa, G. and Gersch, W. (1996). *Smoothness Prior Analysis of Time Series*, Lecture Notes in Statist, No. 116, Springer, Belrin.

Kitagawa, G. and Sato, S. (2000). Nonlinear state space model approach to financial time series with time-varying variance, *Proceedings of the Hong Kong International Workshop on Statistics and Finance: An Interface* (eds. W. S. Chan, W. K. Li and H. Tong), Imperial College Press, London.

Longstaff, F. and Schwartz, E. S. (1992). Interest rate volatility and the term structure: A two-factor general equilibrium model, *Journal of Finance*, **47**, 1259–1282.

Pearson, N. D. and Sun, T. S. (1994). Exploiting the conditional density in estimating the term structure: An application to the Cox, Ingersoll, and Ross model, *Journal of Finance*, **49**, 1279–1304.

Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling, W. T. (1992). *Numerical Recipes in C*, 2nd ed., Cambridge University Press, Cambridge.

Singh, M. K. (1995). Estimation of multifactor Cox, Ingersoll, and Ross term structure model, *Journal of Fixed Income*, September, 8–28.

Tanizaki, H. (1993). *Nonlinear Filters*, Lecture Notes in Econom. and Math. Systems, No. 400, Springer, Belrin.