On Parameter Estimation of the Schwartz-Smith Short-Term/Long-Term Model

Xin Tai and Minyue Fu

Abstract— **The short-term/long-term model proposed by Schwartz and Smith in 2000 is widely used in modeling commodity prices. A key and nontrivial problem in this modeling technique is how to estimate the model parameters. This paper considers the parameter estimation problem based on the maximum likelihood criterion and proposes a method to simplify the task. Two components are contained in the proposed method: one to do with re-parametrization and one to do with separating the parameter set so that one part can be solved directly using least-squares and another part using nonlinear optimization. The effectiveness of the proposed method is demonstrated via numerical tests.**

I. INTRODUCTION

Dynamic behaviors of commodities prices are important determinants of economic fundamentals. Hence, modeling and analysis of commodity price dynamics receive a great amount of attention by both academics and practitioners, and have become an important area of financial economics. Since the Black-Scholes model proposed in 1973 [1], many alternative models have been developed for modeling commodity prices.

Brennan and Schwartz [2] proposed in 1985 a one-factor model to describe the behavior of copper prices, and the optimal decisions for managing the exploration of a mine was analyzed based on this model. Gibson and Schwartz [3] proposed the first two-factor model in commodities in 1990 to analyze oil prices, in which stochastic factors were used to describe the convenience yield and the spot price. In 1997, Schwartz [4] used a set of models including one-, two- and three-factor model to analyze the behavior of commodity prices. Schwartz and Smith [5] proposed an alternative two-factor model, which becomes very popular. Cortazar and Naranjo [6] proposed an N-factor Gaussian model to analyze the oil futures prices in 2006. However, how to estimate the model parameters becomes an important problem in this modeling technique. This has been studied by many researchers, e.g. [7] and [8].

In this paper we propose a simplified method to estimate the parameters in the short-term/long-term model based on the maximum likelihood criterion. Our method has the advantage of giving the true maximum likelihood but having low computational complexity. We first reparameterize the model parameters and then separate the new parameters into two subsets, one sovlable directly using least-suqares and one solvable using nonlinear optimization. As a result, very accurate parameter estimates can be given without heavy computation. A numerical test is given to illustrate its reliability.

This work was supported by Centre of Complex Dynamic Systems and Control at University of Newcastle, Australia.

X. Tai is with State Key Lab. of Industrial Control Technology, Zhejiang University, 310027, P. R. China. taixin@zju.edu.cn

M. Fu is with the School of Electrical Engineering and Computer Science, University of Newcastle, Callaghan, NSW 2308, Australia. Minyue.Fu@newcastle.edu.au

II. SHORT-TERM/LONG-TERM MODEL

The two-factor model proposed by Schwartz and Smith [5] assumes that the logarithm of the price S_t (including spot and futures prices) is separated into two components: a short-term deviation (known as mean-reverting) part χ_t and a long-term equilibrium part ξ_t, i.e.,

$$\ln S_t = \chi_t + \xi_t, \tag{1}$$

and that both parts are driven by a first-order linear stochastic process. More specifically,

$$d\chi_t = -(\kappa\chi_t + \lambda_\chi)dt + \sigma_\chi dz_{\chi,t} \tag{2}$$
$$d\xi_t = (\mu_\xi - \lambda_\xi)dt + \sigma_\xi dz_{\xi,t} \tag{3}$$

where $\kappa > 0$ is the mean-reverting parameter representing the time constant of the transient response to price disturbances, $\mu_\xi > 0$ represents the expected (linear) growth rate of the log price $\ln(S_t)$, λ_χ and λ_ξ present the investment risks due to interest costs for borrowing, both of which equal to zero in the spot price case, $dz_{\chi,t}$ and $dz_{\xi,t}$ are increments of standard Brownian motions representing fluctuations in the spot price, and the two are allowed to correlate with

$$dz_{\chi,t}dz_{\xi,t} = \rho_{\chi\xi}dt,\ \rho_{\chi\xi} \in (-1,\ 1) \tag{4}$$

We denote the state and output of the model by $x_t = [\chi_t\ \xi_t]'$. Without loss of generality, it is assumed that the price sampling interval $\triangle t = 1$. Then the continuous-time model (2) and (3) can be converted into a discrete-time model as follows:

$$\begin{aligned} x_t &= c + Gx_{t-1} + w_t, \\ y_t &= d + F'x_t + v_t, \end{aligned} \tag{5}$$

where

$$x_t = \begin{bmatrix} \chi_t \\ \xi_t \end{bmatrix},\ c = \begin{bmatrix} 0 \\ \mu_\xi \end{bmatrix},\ G = \begin{bmatrix} e^{-\kappa} & 0 \\ 0 & 1 \end{bmatrix} \tag{6}$$

978-89-956056-9-1/09/©2009 ACA

$$y_t = \begin{bmatrix} \ln E\{S_{T_1}\} \\ \ln E\{S_{T_2}\} \\ \vdots \\ \ln E\{S_{T_n}\} \end{bmatrix}, \; d = \begin{bmatrix} a(T_1) \\ a(T_2) \\ \vdots \\ a(T_n) \end{bmatrix}$$

$$F' = \begin{bmatrix} e^{-\kappa T_1} & 1 \\ e^{-\kappa T_2} & 1 \\ \vdots & \vdots \\ e^{-\kappa T_n} & 1 \end{bmatrix} \tag{7}$$

$$\begin{aligned} a(T) &= \mu_\xi^\star T - (1-e^{-\kappa T})\frac{\lambda_\chi}{\kappa} + (1-e^{-2\kappa T})\frac{\sigma_\chi^2}{4\kappa} \\ &\quad + (1-e^{-\kappa T})\frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{\kappa} + \frac{\sigma_\xi^2}{2}T \end{aligned} \tag{8}$$

The mean of the state variables can be obtained as:

$$m_t = \begin{bmatrix} e^{-\kappa t}\chi_0 \\ \xi_0 + \mu_\xi t \end{bmatrix} \tag{9}$$

where χ_0 and ξ_0 are initial values of the state variables. The output variable y_t is the vector of observed (log) futures prices with time maturities T_i $i = 1,2,\ldots,n$, where n is the number of futures prices, and $T = 0$ for spot price. w_t and v_t are zero-mean Gaussian white noises with covariances

$$\begin{aligned} Cov\{w_t\} &= \begin{bmatrix} (1-e^{-2\kappa})\frac{\sigma_\chi^2}{2\kappa} & (1-e^{-\kappa})\frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{\kappa} \\ (1-e^{-\kappa})\frac{\rho_{\chi\xi}\sigma_\chi\sigma_\xi}{\kappa} & \sigma_\xi^2 \end{bmatrix} \\ Cov\{v_t\} &= \mathrm{diag}\{s_1^2, s_2^2, \ldots, s_n^2\} \end{aligned} \tag{10}$$

where $\sigma_\chi, \sigma_\xi, \rho_{\xi\chi}$ and $s_i, i = 1,2,\ldots,n$ are the parameters related to the noise variances. There are $n+8$ parameters in this model. All of them need to be estimated before the short-term/long-term model is used.

III. Maximum Likelihood Estimation

The parameters in short-term/long-term model need to be estimated from a sequence of given observations (past spot and futures prices) $Y = [y_1 \; y_2 \; \cdots \; y_N]'$. Schwartz and Smith [5] suggests to use maximum likelihood estimation. The likelihood function used by Schwartz and Smith is not explicitly given in their paper. This, however, can be worked out from the estimation algorithm used in the paper. Indeed, the initial state x_0 is assumed to have a Gaussian distribution with mean $\hat{x}_0$ and covariance C_0, then the likelihood function can be computed as:

$$L(\theta, \hat{Y}) = \mathrm{Prob}(y_t = \hat{y}_t, t = 1,2,\cdots,N) \tag{11}$$

where $\theta = \{\kappa, \lambda_\chi, \mu_\xi, \mu_\xi^\star, \sigma_\chi, \sigma_\xi, \rho_{\chi\xi}, s_1, s_2, \cdots, s_n, \hat{x}_0, C_0\}$ denotes the set of model parameters, $\hat{Y} = [\hat{y}_1 \; \hat{y}_2 \; \cdots \; \hat{y}_N]'$, and $\hat{y}_t$ is the one-step-ahead estimate of y_t based on the observations $y_1, y_2, \cdots, y_{t-1}$ as well as $\hat{x}_0$ and C_0.

The resulting maximum likelihood estimation problem involves two parts. The first part is about optimal prediction where the set of model parameters θ is assumed to be known and the task is to maximize the likelihood function $L(\theta, \hat{Y})$. This is the well-known Kalman filtering problem and the solution is readily available. The second part is parameter estimation which aims to optimize the model parameters, which is the focus of this paper.

A. Kalman Filtering

The solution to the optimal one-step-ahead prediction is given by a recursive Kalman filter as follows:

$$\begin{aligned} \hat{x}_{t|t-1} &= c + G\hat{x}_{t-1} \\ \hat{y}_t &= d + F'\hat{x}_{t|t-1} \\ R_t &= GC_{t-1}G' + W \\ Q_t &= F'R_tF + V \\ A_t &= R_tFQ_t^{-1} \\ \hat{x}_t &= \hat{x}_{t|t-1} + A_t(y_t - \hat{y}_t) \\ C_t &= R_t - A_tQ_tA_t' \end{aligned} \tag{12}$$

where $\hat{x}_{t|t-1}$ is the one-step-ahead prediction of x_t with R_t being the covariance of the corresponding prediction error, $\hat{y}_t$ is the one-head prediction of y_t with Q_t being the covariance of the corresponding prediction error, A_t is the Kalman gain, $\hat{x}_t$ is the updated estimate of x_t with C_t being the covariance of the corresponding estimation error.

One important feature of recursive Kalman filtering is that the recursion of C_t and R_t is independent of the observation data. Indeed, from (12), we get

$$R_{t+1} = G(R_t - R_tF(F'R_tF + V)^{-1}F'R_t)G' + W$$

Using the fact that

$$R_t - R_tF(F'R_tF + V)^{-1}F'R_t = (R_t^{-1} + FV^{-1}F')^{-1}$$

(the well-known matrix inversion lemma), the above becomes

$$R_{t+1} = G(R_t^{-1} + FV^{-1}F')^{-1}G' + W \tag{13}$$

It is also straightforward to show that A_t can be rewritten as

$$A_t = (R_t^{-1} + FV^{-1}F')^{-1}FV^{-1} \tag{14}$$

which also implies that

$$I - A_tF' = (R_t^{-1} + FV^{-1}F')^{-1}R_t^{-1} \tag{15}$$

Using (13)-(15), the recursion in (12) can be simplified to

$$\begin{aligned} \hat{x}_t &= (R_t^{-1} + FV^{-1}F')^{-1} \\ &\quad \cdot\{R_t^{-1}(c + G\hat{x}_{t-1}) + FV^{-1}(y_t - d)\} \\ \hat{y}_t &= d + F'(c + G\hat{x}_{t-1}) \\ R_{t+1} &= G(R_t^{-1} + FV^{-1}F')^{-1}G' + W \end{aligned} \tag{16}$$

with $R_1 = GC_0G' + W$, $t = 1,2,\cdots,N$.

Since G is a marginally stable matrix (with eigenvalues at $e^{-\kappa}$ and 1), R_t converges asymptotically to some constant R which is positive definite. So the recursion in (12) can be simplified to the following steady-state Kalman filter:

$$\begin{aligned} \hat{x}_t &= (R^{-1} + FV^{-1}F')^{-1} \\ &\quad \cdot\{R^{-1}(c + G\hat{x}_{t-1}) + FV^{-1}(y_t - d)\} \\ \hat{y}_t &= d + F'(c + G\hat{x}_{t-1}) \end{aligned} \tag{17}$$

where R is the solution of the following nonlinear equation:

$$R = G(R^{-1} + FV^{-1}F')^{-1}G' + W \tag{18}$$

B. Parameter Estimation

We now consider the parameter estimation problem. We will take the assumption that the data set is sufficiently long so that the recursive Kalman filter can be approximated using the steady-state Kalman filter.

The proposed method contains two components: one to do with re-parametrization and one to do with separating the parameter set so that one part can be solved directly using least-squares and another part using nonlinear optimization. As a result, the parameter estimation problem is much more manageable.

Before we proceed further, we need to work out the likelihood function after Kalman filtering. It is a well-known property of the recursive Kalman filter that the prediction error

$$\varepsilon_t = y_t - \hat{y}_t \tag{19}$$

is a zero-mean Gaussian white noise sequence with covariance $Q_t = F'R_tF + V$ which becomes $Q = F'RF + V$ at steady state. Therefore, the probability of $Y = \hat{Y}$ is proportional to

$$\prod_{t=1}^{N} \frac{1}{\sqrt{\det(Q_t)}} \exp\left(-\frac{1}{2}(y_t - \hat{y}_t)Q_t^{-1}(y_t - \hat{y}_t)\right) \tag{20}$$

Taking the log of the above, maximizing the likelihood function is the same as minimizing

$$\sum_{t=1}^{N} \{\ln(\det(Q_t)) + (y_t - \hat{y}_t)'Q_t^{-1}(y_t - \hat{y}_t)\} \tag{21}$$

To account for the potential errors in the transient period when a steady-state Kalman filter is used, we can truncate the sequence of $\{\varepsilon_t\}$ to start from some $N_0 > 0$. We also replace Q_t by Q, so the resulting maximum likelihood estimation problem becomes

$$\begin{aligned} \min_{\theta} J(\theta) &= (N - N_0 + 1)\ln(\det(Q)) \\ &\quad + \sum_{t=N_0}^{N} (y_t - \hat{y}_t)'Q^{-1}(y_t - \hat{y}_t) \end{aligned} \tag{22}$$

1) Re-parametrization of W: Examining (18) shows that R and W form a one-one mapping in the set of 2×2 positive definite matrices. Indeed, for any $W > 0$, there is a unique solution $R > 0$ to (18), as explained before. On the other hand, for any $R > 0$, $(R^{-1} + FV^{-1}F')^{-1} < R$, which means that the solution to W in (18), which is given by

$$W = R - G(R^{-1} + FV^{-1}F')^{-1}G' \tag{23}$$

is unique and positive definite.

The observation above means that we can use R instead of W as parameters. This has the advantage that instead of solving the nonlinear equation (18), we only need to solve (23), which is much easier.

2) Re-parametrization of x_t: Recall from (9) that y_t has the mean of $e^{-\kappa t}\chi_0 + \xi_0 + \mu_\xi t$. To bring this term out explicitly in $\hat{y}_t$, we define

$$\begin{aligned} \tilde{\chi}_t &= \chi_t - e^{-\kappa t}\chi_0 \\ \tilde{\xi}_t &= \xi_t - \xi_0 - \mu_\xi t \end{aligned} \tag{24}$$

and the new state $\tilde{x}_t = [\tilde{\chi}_t \ \tilde{\xi}_t]'$. Then, the discrete-time two-factor model (5) becomes

$$\tilde{x}_t = G\tilde{x}_{t-1} + w_t, \quad \tilde{x}_0 = 0 \tag{25}$$

$$y_t = \tilde{d}_t + F'\tilde{x}_t + v_t \tag{26}$$

where

$$\tilde{d}_t = d + (e^{-\kappa t}\chi_0 + \xi_0 + \mu_\xi t)h \tag{27}$$

with $h = [1 \ 1 \cdots \ 1]'$. The corresponding steady-state Kalman filter is further simplified to

$$\begin{aligned} \hat{\tilde{x}}_t &= K\hat{\tilde{x}}_{t-1} + A(y_t - \tilde{d}_t), \ \hat{\tilde{x}}_0 = 0, \\ \hat{y}_t &= \tilde{d}_t + D\hat{\tilde{x}}_{t-1} \end{aligned} \tag{28}$$

where

$$\begin{aligned} K &= (R^{-1} + FV^{-1}F')^{-1}R^{-1}G \\ A &= (R^{-1} + FV^{-1}F')^{-1}FV^{-1} \\ D &= F'G \end{aligned} \tag{29}$$

3) Partial Parameter Estimation via Least-Squares: We now separate the set of parameters into two subsets: $\theta_1 = [\mu_\xi \ \mu_\xi^\star \ \lambda_\chi \ \chi_0 \ \xi_0]'$ and $\theta_2 = \{\kappa, R, V\}$. It turns out that θ_1 can be efficiently optimized via least-squares. To see this, we note that θ_1 enters only into $\tilde{d}_t$ and do so linearly (see (7), (10) and (27)), the filter (28) operates linearly on $\tilde{d}_t$. Therefore, $\hat{y}_t$ is linear in θ_1. In computing $\hat{y}_t$, we can easily express it as

$$\hat{y}_t = \phi_t\theta_1 + \psi_t \tag{30}$$

where ϕ_t is a $n \times 5$ matrix and ψ_t is a $n \times 1$ vector. The exact expressions for ϕ_t and ψ_t can be worked out from (28), but omitted here.

The least-squares solution to the minimization of $J(\theta)$ with respect to θ_1 is then given by

$$\theta_1^\star = \left(\sum_{t=N_0}^{N} \phi_t'Q^{-1}\phi_t\right)^{-1} \sum_{t=N_0}^{N} \phi_t'Q^{-1}(y_t - \psi_t) \tag{31}$$

and the resulting $J(\theta)$ is given by

$$\tilde{J}(\theta_2) = (N - N_0 + 1)\ln(\det(Q)) + \Upsilon(Q) \tag{32}$$

where

$$\begin{aligned} \Upsilon(Q) &= \sum_{t=N_0}^{N} (y_t - \psi_t)'Q^{-1}(y_t - \psi_t) \\ &\quad - \left(\sum_{t=N_0}^{N} (y_t - \psi_t)'Q^{-1}\phi_t'\right)\left(\sum_{t=N_0}^{N} \phi_t'Q^{-1}\phi_t\right)^{-1} \\ &\quad \cdot \left(\sum_{t=N_0}^{N} \phi_tQ^{-1}(y_t - \psi_t)\right) \end{aligned} \tag{33}$$

We caution that the parameters χ_0 and ξ_0 can not be estimated accurately using the method above. The reason is that the effect of the initial condition x_0 on the estimate $\hat{y}_t$ is transient and will decay exponentially. When the number of samples N becomes large, the cost function $J(\theta)$ becomes insensitive to x_0, which makes the accurate estimate of x_0 difficult. More precisely, we first see from (27) that the effect

of χ_0 on $\tilde{d}_t$ diminishes exponentially. Then, we see from (28) that the effect of $\tilde{d}_\tau$ on $\hat{y}_{t+\tau}$ also diminishes exponentially (because K has eigenvalues strictly within the unit circle). The two facts above together mean that the effect of χ_0 to $\hat{y}_t$ diminishes exponentially. To see that the effect of ξ_0 on $\hat{y}_t$ also diminishes exponentially, we note that the component of ξ_0 in $\tilde{d}_t$ is a constant vector $h\xi_0$. At steady state, the response of $\hat{y}_t$ to this component is

$$\hat{y}_{\xi_0} = (I - D(I-K)^{-1}A)h\xi_0$$

which can be obtained from (28) by setting $\hat{\hat{x}}_t = \hat{\hat{x}}_{t-1}$. Using (29) and the fact that $h\xi_0 = F'G[0\ \xi_0]'$, it is straightforward to verify that $\hat{y}_{\xi_0} = 0$ for any ξ_0. This confirms the claim that the effect of ξ_0 on $\hat{y}_t$ diminishes exponentially.

Inaccurate estimation of χ_0 is typically not problematic because its effect on y_t (and hence $\hat{y}_t$) is transient only. This difficulty is also inherent because its effect is transient only. The difficulty in estimating ξ_0 is not inherent, it is due to the use of the Kalman filter based method. In other words, accurate estimation of ξ_0 is possible because its information is permanently present in y_t. Despite the fact that χ_0 are ξ_0 can be inaccurately estimated, it does not affect computing the minimum of $J(\theta)$ over θ_1.

4) Recursive Partial Parameter Estimation: The solution to θ_1 in (31) can be computed recursively using a standard recursive least-squares method, as explained below.

Defining the recursion

$$P_t = (P_{t-1}^{-1} + \phi_t' Q^{-1} \phi_t)^{-1} \tag{34}$$

with $P_{N_0} = \gamma I$ for some large γ, it is straightforward to verify that

$$P_t = P_{t-1} - P_{t-1}\phi_t'(Q + \phi_t P_{t-1}\phi_t')^{-1}\phi_t P_{t-1} \tag{35}$$

(which follows the matrix inversion lemma) and that

$$P_t \to \left(\sum_{\tau=N_0}^{t} \phi_\tau' Q^{-1} \phi_\tau \right)^{-1}, \text{ as } \gamma \to \infty \tag{36}$$

when t is large. Defining

$$\theta_{1,t} = P_t \sum_{\tau=N_0}^{t} \phi_\tau' Q^{-1}(y_\tau - \psi_\tau) \tag{37}$$

and using (35), we have

$$\theta_{1,t} = (I - \alpha_t \phi_t)\theta_{1,t-1} + \alpha_t(y_t - \psi_t) \tag{38}$$

with $\theta_{1,N_0} = 0$, where

$$\alpha_t = P_{t-1}\phi_t'(Q + \phi_t P_{t-1}\phi_t')^{-1} \tag{39}$$

It is clear that as $\gamma \to \infty$, $\theta_{1,N} \to \theta_1^\star$ in (31).

C. Parameter Estimation Algorithm

The proposed parameter estimation method is summarised in the algorithm below:

Step 1: Initialize θ_2, i.e., initialize κ, R and V.

Step 2: Compute F, G, W, K, A and D using (6), (7), (23) and (29). Also compute $Q = F'RF + V$.

Parameters	Values	Parameters	Values
κ	1.49	σ_χ	0.286
λ_χ	0.157	μ_ξ	-0.0125
σ_ξ	0.145	μ_ξ^*	0.0115
$\rho_{\xi\chi}$	0.300	χ_0	0.5
ξ_0	2.5	$s_i, i = 1,2,\ldots,n$	0.005

TABLE I
PARAMETERS' VALUES OF GENERATION MODEL.

Step 3: Express $\tilde{d}_t = \Delta_t \theta_1 + \delta$ using (7) and (27), then compute Δ_t and δ.

Step 4: Run the steady-state Kalman filter (28) to compute ϕ_t and ψ_t in (30) for $t = N_0, 2, \cdots, N$ (where N_0 can be chosen to be, say 10% of N).

Step 5: Compute the optimal θ_1 and $\tilde{J}(\theta_2)$ using (31) and (32).

Step 6: Tune κ, R and V using any nonlinear optimization method (e.g., Newton Gradient search) until further reduction of $\tilde{J}(\theta_2)$ is negligible.

IV. NUMERICAL TESTS

In this section some simulation results will be given to illuminate how this method works. We frist test the method on a set of artificially generated data using a known two-factor model. The purpose of this exercise is to see how well the method estimated the paremeters. We then apply the proposed method on the NYMEX crude oil futures contract data which were used in [5], and we show that our method gives somewhat better estimates than those given in [5].

A. Generation of Artifical Data

A set of artifical data are generated using the two-factor model (equation (5)), with the values of parameters as in Table I. Fig. 1 shows the state variables, and Fig. 2 shows the outputs of the futures contracts with maturities $(0, 1, 5, 9, 13, 17)$ months (Obviously, the first one is the spot price). We generate 500 data samples (i.e., $N = 500$).

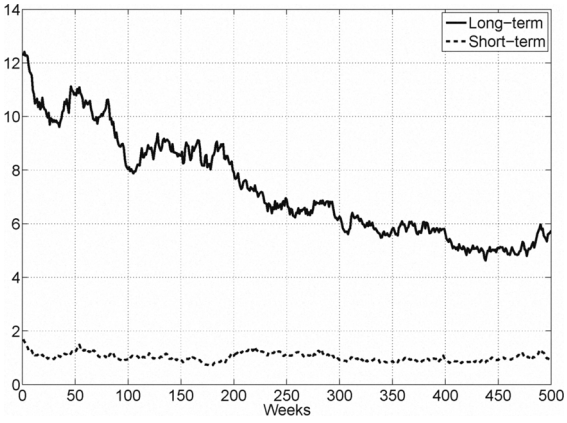

Fig. 1. Generated States of the Two-Factor Model.

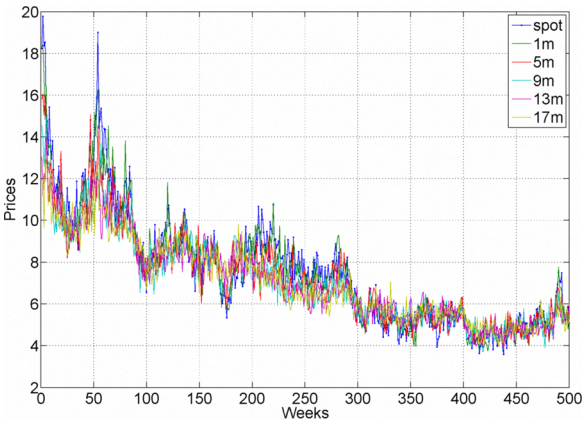

Fig. 2. Generated Outputs of the Two-Factor Model.

θ_2	Correct Value	Size of Data 500	1000	2000
κ	1.49	1.4625	1.4629	1.4924
σ_χ	0.286	0.2551	0.2551	0.3002
σ_ξ	0.145	0.1296	0.1296	0.1329
$\rho_{\xi\chi}$	0.300	0.5058	0.5059	0.2761
s_i	0.005	0.00494	0.00494	0.00497

TABLE II
ESTIMATION RESULTS OF θ_2.

B. Testing Estimation of θ_1

We first want to see how the estimation of θ_1 works. For this purpose, we choose θ_2 to be the correct values and apply the least squares method to estimate θ_1. The results are shown in Fig. 3 to Fig. 7, from which we can see that the estimated values converge well as time evolves.

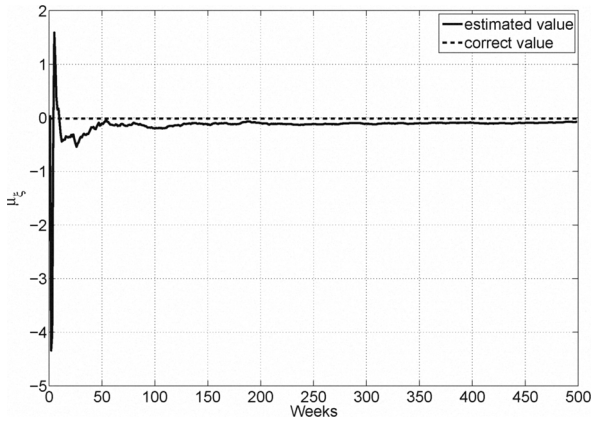

Fig. 3. Estimation of μ_ξ.

C. Testing Estimation of θ_2

With some initial values, the parameters in θ_2 were estimated via Newton Graident method. From the results in Table II, we can see that, when the size of the data increases, the estimated values of θ_2 also approach their correct values. However, it should be noted that in this table we give the results of $\sigma_\chi, \sigma_\xi, \rho_{\xi\chi}$ directly instead of R or W.

Parameters	Values given in [5]	Estimated Values
κ	1.49	1.50
σ_χ	28.6%	31.39%
λ_χ	15.7%	12.17%
μ_ξ	-1.25%	-1.72%
σ_ξ	14.5%	14.45%
μ_{ξ_*}	1.15%	0.94%
$\rho_{\xi\chi}$	0.300	0.227
χ_0	0.117	0.158
ξ_0	3.01	2.99
s_1	0.042	0.00181
s_2	0.006	0.00003
s_3	0.003	0.00001
s_4	0.000	0.00000
s_5	0.004	0.00002
cost function $J(\theta)$	-6.8853×10^3	-1.0626×10^4

TABLE III
PARAMETER ESTIMATION RESULTS BASED ON NYMEX DATA

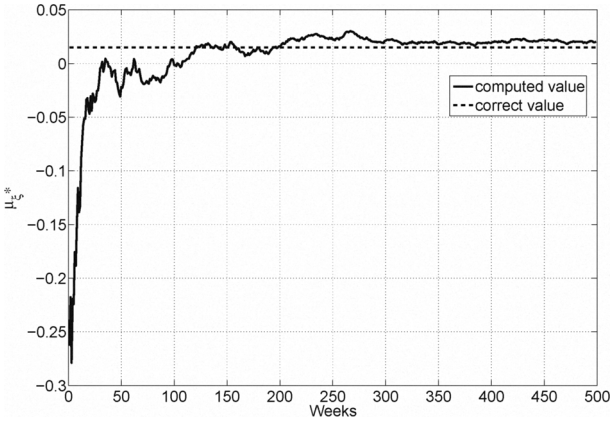

Fig. 4. Estimation of μ_ξ^*.

D. Application to NYMEX Crude Oil Prices

Now we apply the proposed method to the NYMEX crude oil futures contracts with maturities in 1,5,9,13 and 17 months, from 1/2/90 to 2/17/95. This is the same set of data used in [5]. The purpose of our exercise is to see how well we can estimate the parameters in a two-factor model.

We note that [5] gives a set of estimated parameters using a numerical method. However, their extimates do not contain the values for χ_0 and ξ_0. In order to see how our method works, we first apply our method to estimate all the paremeters. We then take the parameter values given in [5] and use the least-squares method to obtain the optimal values for χ_0 and ξ_0. Finally, the cost function $J(\theta)$ is compared using the two sets of parameters. The results are shown in Table III, from which we see that the second set of parameters has a less value of the likelihood function, which means it can let the model to be more likely to the actual data. (i.e., a smaller value for the cost function).

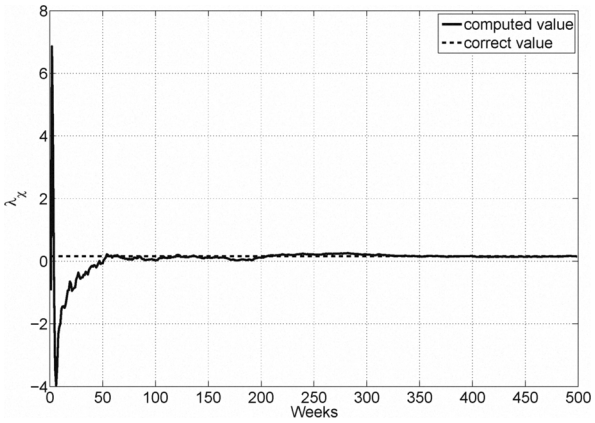

Fig. 5. Estimation of λ_χ.

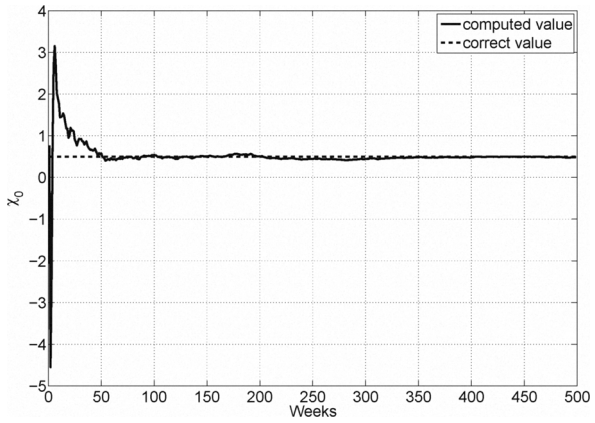

Fig. 6. Estimation of χ_0.

V. Conclusions

This paper proposes a method to simplify the parameter estimation problem in the short-term/long-term model of Schwartz and Smith. The resulting optimization problem is still nonlinear, but with much less number of parameters to search. Some numerical tests are given to show this method works. The proposed method is tested on a set of NYMEX oil data used in [5] and favorable comparison is shown.

References

[1] R. C. Merton, "An intertemporal captial asset pricing model," *Econometrica*, vol. 41, no. 5, pp. 867-887, Sept 1973.

[2] M.J. Brennan, E.S. Schwartz, "Evaluating natural resources investments," *Journal of Business*, vol. 58, pp. 133, 1985.

[3] R. Gibson, E.S. Schwartz, "Stochastic convenience yield and the pricing of oil contingent claims," *The Journal of Finance*, vol. 45, no. 3, pp. 959-976, 1990.

[4] E.S. Schwartz, "The stochastic behavior of commodity prices: implications for valuation and hedging," *The Journal of Finance*, vol. 52, no. 3, pp. 923-973, 1997.

[5] E. Schwartz, J. Smith, "Short-term variations and long-term dynamics in commodity prices," *Management Science*, Vol. 46, no. 7, pp. 893-911, July 2000.

[6] G. Cortazar, L. Naranjo, "An N-factor Gaussian model of oil futures prices," *The Journal of Futures Markets*, vol. 26, no. 3, pp. 243-268, 2006.

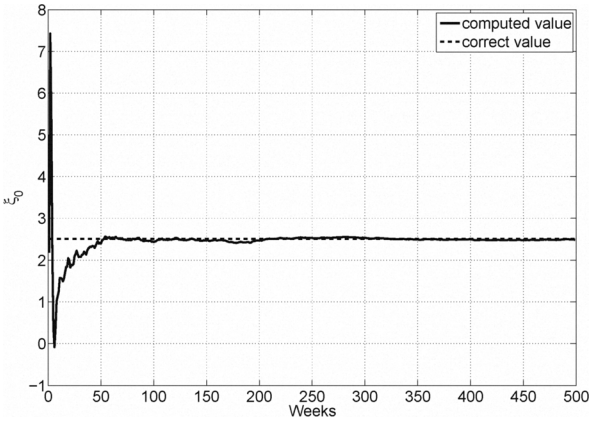

Fig. 7. Estimation of ξ_0.

[7] R.J. Elliott, C.B. Hyndman, "Parameter estimation in commodity markets: A filtering approach," *Journal of Economic Dynamics & Control*, vol. 31, pp.2350-2373, 2006.

[8] F.A.L. Aiube, T.K.N. Baidya and E.A.H. Tito, "Analysis of commodity prices with the particle filter," *Energy Economics*, vol. 30, pp. 597-605, 2006.

[9] R.S. Mamon, C. Erlwein and R.B. Gopaluni, "Adaptive signal processing of asset price dynamics with predictability analysis," *Information Sciences*, vol. 178, pp.203-219, 2008.