Parameter Estimation of Electricity Spot Models from Futures Prices

S. Aihara ∗ A. Bagchi ∗∗ E. Imreizeeq ∗∗

∗ Tokyo University of Science, Suwa, Toyohira 5000-1, Chino, Nagano, Japan (e-mail: aihara@rs.suwa.twu.ac.jp)
** FELab, Department of Applied Mathematics and Department of Finance and Accounting, University of Twente, P.O.Box 217, 7500AE Enschede, The Netherlands (e-mails: a.bagchi@ewi.utwente.nl, e.s.n.imreizeeq@ewi.utwente.nl)

Abstract: We consider a slight perturbation of the Schwartz-Smith model for the electricity futures prices and the resulting modified spot model. Using the martingale property of the modified price under the risk neutral measure, we derive the arbitrage free model for the spot and futures prices. We estimate the parameters of the model by the method of maximum likelihood using the Kalman filter’s estimate of the unobservable state variables, coupled with the usual statistical techniques. The main advantage of the new model is that it avoids the inclusion of artificial noise to the observation equation for the implementation of Kalman filter. The extra noise is built within in the model in an arbitrage free setting.

Keywords: Parameter Estimation; Finance; Kalman filter; Maximum Likelihood estimators; Electricity Spot Model.

1. INTRODUCTION

The recent deregulation of gas and electricity markets led to the creation of a network of energy exchanges, where the electricity is quoted almost as any other commodity. However, electricity cannot be stored. Furthermore, electricity produced at any moment should be immediately consumed, so that supply and demand of electricity have to be matched at any moment of time. Hence, the power prices present a much higher volatility than equity prices. A precise mathematical model of electricity spot price behavior is required for energy risk management, and for pricing of electricity-related futures and options.

There are two approaches for the modeling of electricity. One approach is modeling the spot price dynamics, from which the forward dynamics can be constructed, see Miltersen and Schwartz [1998] and Schwartz and Smith [2000]. The other approach follows Heath Jarrow Morton (HJM) framework, see Heath et al. [1992] which describes the forward curve dynamics directly via the use of volatility function(s), see Clewlow and Strickland [1999]. In these approaches, the model parameters may be calibrated using the maximum likelihood method. In order to derive the likelihood functional, the Kalman filter is constructed. However in spite of the mathematically elegant derivation of the futures prices, which are the observed data, one needs to add some ad hoc observation noise in order to derive the Kalman filter. This assumption has been made by numerous authors, either in the commodity or interest rates market, see Elliott and Hyndman [2007]. The additional noise in the observation has been interpreted to take into account bid-ask spreads, price limits, non-simultaneity of the observations, or errors in the data. The argument is clearly forced and unconvincing. We approach the modeling differently. In our setup, on one hand, the added measurement noise is built in within the model. On the other hand, the modeling of the correlation structure between the futures (observation) is a natural component of our formulation.

This paper is organized as follows: A review of Schwartz and Smith [2000] model is presented in Section 2. In Section 3, we present the new model for the future price for one maturity date $T$, by using the idea proposed in Aihara and Bagchi [2008]. In the following Section, we focus our attention on the electricity futures situation. The futures contracts on electricity are based on the arithmetic averages of the spot prices over a delivery period. Although we are able to derive the explicit formula of the futures prices for the arithmetic average case, the explicit formula is not easy to handle. So, we use the geometric average as an approximation for the arithmetic average instead. In Section 5, we derive the explicit relation between the observed futures prices and the factor processes, together with the likelihood functional. In Section 6, we estimate model parameters by maximizing the derived likelihood functional which involves the Kalman filter of the system states. The last two sections contain the simulation work and conclusion, respectively.

2. THE SCHWARTZ-SMITH MODEL

Let $S(t)$ represent the spot price of a commodity (electricity) at time $t$. Following Schwartz and Smith [2000], we decompose the logarithm of the spot price into two stochastic factors as
\[
\ln(S(t)) = \chi(t) + \xi(t) + h(t) \quad (1)
\]

where \(\chi(t)\) represents the short-term deviation in the price, \(\xi(t)\) is the equilibrium price level and \(h(t)\) is a deterministic seasonality function. Assume that the risk-neutral stochastic process for the two factors are of the form

\[
\begin{align*}
\{d\chi(t)\} &= (-\kappa\chi(t) - \lambda\chi)dt + \sigma_{\chi}dW^\omega_{\chi}(t) \\
\{d\xi(t)\} &= \lambda_{\xi}dt + \sigma_{\xi}dW^\omega_{\xi}(t)
\end{align*}
\]  
(2)

where \(W^\omega_{\chi}\) and \(W^\omega_{\xi}\) are correlated standard Brownian motions, where \(dW^\omega_{\chi}dW^\omega_{\xi} = \rho dt\). We denote the current time by \(t\), the maturity of the futures by \(T\), the time to maturity by \(\tau\) where \(\tau = T - t\), and by \(T^*\) a fixed time horizon where \(t \leq T < T^*\). The futures price \(F(t, T - t)\) is given by

\[
F(t, T - t) = \exp(B(t, T - t) + C(T - t)\xi(t)) + A(t, T - t) \quad (3)
\]

where

\[
B(t, T - t) = e^{-\kappa(T-t)}, \quad C(T - t) = 1 \quad (4)
\]

\[
A(t, T - t) = \frac{\lambda_{\xi}}{\kappa}\left(e^{-\kappa(T - t)} - 1\right) + \lambda_{\xi}(T - t) + \frac{1}{2}\sigma^2_{\chi}(T - t) + h(T) \quad (5)
\]

and

\[
\sigma^2_{\chi}(T - t) = \frac{\sigma^2_{\chi}}{2\kappa}\left(1 - e^{-2\kappa(T - t)}\right) + \sigma^2_{\xi}(T - t) + \frac{2\rho\sigma_{\chi}\sigma_{\xi}}{\kappa}\left(1 - e^{-\kappa(T - t)}\right) \quad (6)
\]

3. A NEW MODEL FOR THE ELECTRICITY PRICES

We assume that the correct model for the spot price is not exactly the same as in (1), but is close to it. We then expect the futures price also to be somewhat perturbed from the formula given in (3). Hence, suppose that the correct futures price at time \(t\) where \(t \leq T\) is given by

\[
F^{corr}(t, T - t) = \exp[B(t, T - t)\chi(t) + \tilde{C}(T - t)\xi(t)] + \tilde{A}(t, T - t) + \int_0^t \sigma dw(s, T - s) \quad (7)
\]

where

\[
\int_0^t \sigma dw(s, T - s) = \sum_{k=1}^{\infty} \int_0^t \frac{1}{\lambda_k} e_k(T - s) d\beta_k(s)
\]

(7)

and where \(e_k\) is a sequence of differentiable functions forming an orthonormal basis in \(L^2((0, T^*), \{\beta_k(0)\})\) mutually independent Brownian motions processes. Let \(q(x, y)\) represents the correlation of \(w(t, x)\) and \(w(t, y)\). The extra stochastic integral term \(\int_0^t \sigma dw(s, T - s)\) which appears in (6), represents the modeling error between the futures price given by (3) and the correct futures price. When \(T = t\), the correct spot price process is given by

\[
S^{corr}(t) = F^{corr}(t, 0) \quad (8)
\]

To get the corresponding (correct) dynamics for the spot, we need the dynamics of the futures taking into account that this dynamics under the risk-neutral measure is a martingale. Applying Ito's formula to (6), we get

\[
\frac{dF^{corr}(t, T - t)}{F^{corr}(t, T - t)} = \frac{d\tilde{A}(t, T - t)}{\tilde{A}(t, T - t)} + \frac{d\tilde{B}(T - t)}{\tilde{B}(T - t)} + \frac{d\tilde{C}(T - t)}{\tilde{C}(T - t)} + \frac{d\psi(t)}{\psi(t)} \quad (9)
\]

For the futures price to be a martingale, the dt-part of (9) has to be zero. For that, we get \(\tilde{B}(T, T - t) = B(t, T - t)\), \(\tilde{C}(t, T - t) = C(t, T - t)\) given by (4) and \(\tilde{A}\) satisfies

\[
\frac{d\tilde{A}(t, T - t)}{\tilde{A}(t, T - t)} = \lambda_{\chi}e^{-\kappa(T - t)} + \frac{1}{2}\sigma^2_{\chi}e^{-\kappa(T - t)} + \frac{1}{2}\sigma^2_{\xi}e^{-\kappa(T - t)} + \rho\sigma_{\chi}\sigma_{\xi}e^{-\kappa(T - t)} + \frac{1}{2}\sigma^2 q(T - t, T - t) = 0, \quad \tilde{A}(T, 0) = h(T) \quad (10)
\]

The solution of this is given by

\[
\tilde{A}(t, T - t) = A(t, T - t) + \frac{1}{2}\sigma^2 \int_0^{T-t} q(x, x) dx \quad (11)
\]

where \(A(t, T - t)\) is given by (5). Substituting \(\xi(t)\) in (6), we obtain

\[
F^{corr}(t, T - t) = \exp(B(t, T - t)\chi(t) + \tilde{A}(t, T - t) + \int_0^t \sigma dw(s, T - s) + \sigma_{\xi}dW^\omega_{\xi}(s)) \quad (12)
\]

Using (8), the correct spot price process is given by

\[
S^{corr}(t) = \frac{F^{corr}(t, 0)}{F^{corr}(t, 0)} = \exp(\chi(t) + h(t) + \lambda_{\xi}t + \frac{1}{2}\sigma^2 \int_0^t \sigma dw(s, T - s) + \sigma_{\xi}dW^\omega_{\xi}(s)) \quad (13)
\]

From here onwards, we omit writing the expression "corr" for \(S(t)\) and \(F(t, T)\) processes.

4. PRACTICAL MODEL FOR THE ELECTRICITY PRICES

The market prices of electricity futures are different from the standard futures traded in other markets, such as futures on commodities or futures on bonds. The electricity futures prices are based on the arithmetic averages of the spot prices over a delivery period \([T_0, T]\), given by
\[ \frac{1}{T - T_0} \int_{T_0}^{T} S(\tau) d\tau. \] (13)

Now, for \( t < T \), we can calculate the futures price by
\[
F(t, T_0, T) = E\{ \frac{1}{T - T_0} \int_{T_0}^{T} S(\tau) d\tau | \mathcal{F}_t \},
\]
(14)
where \( \mathcal{F}_t = \sigma\{S(\tau); 0 \leq \tau \leq t\} \). This price satisfies
\[
F(t, T_0, T) = \frac{1}{T - T_0} \int_{T_0}^{T} \exp \left[ B(\eta - t) \chi(t) + \tilde{A}(t, \eta - t) + \int_{0}^{t} \{ \sigma dW_s(\eta - s) + \sigma_{\xi} dW_{\xi}(s) \} \right] d\eta,
\]
(15)
where \( B \) and \( \tilde{A} \) satisfy the same equations (4) and (10).

4.1 Geometric Average approximation

Energy futures has a payoff that depends on the arithmetic average of the spot in a specific period. Since the sum of lognormal random variables is not lognormal, we adopt the geometric average as an approximation;
\[
\exp \left\{ \frac{1}{T - T_0} \int_{T_0}^{T} \log S(\tau) d\tau \right\}
\]
(16)
The futures price for this average satisfies
\[
\hat{F}(t, T_0, T) = E\{ \exp \left\{ \frac{1}{T - T_0} \int_{T_0}^{T} \log S(\tau) d\tau \right\} | \mathcal{F}_t \},
\]
(17)
and for \( t < T_0 \)
\[
\hat{F}(t, T_0, T) = \exp \left\{ \frac{1}{T - T_0} \int_{T_0}^{T} \left[ B(\eta - t) \chi(t) + \tilde{A}(t, \eta - t) + \int_{0}^{t} \{ \sigma dW_s(\eta - s) + \sigma_{\xi} dW_{\xi}(s) \} \right] d\eta \right\}
\]
(18)
where it is obvious that \( B \) and \( \tilde{A} \) satisfy the same equations (4) and (10).

5. OBSERVATION MECHANISM

In practice, the observed real data for the futures are on daily basis and transformed such that the time-to-delivery \( \tau = T_0 - t \) is fixed as a constant. For each \( t \), \( T_0 - t \) is set as a constant time period \( \tau_i \) for \( i = 1, 2, \cdots, m \) and \( T - T_0 \) is set as a constant \( \theta \) (1 month). Hence our observation data becomes
\[
y(t, \tau_i) = \log \hat{F}(t, \tau_i + t, \tau_i + t + \theta).
\]
(19)
Setting \( z = \eta - t \) in (18), \( y(t, \tau_i) \) satisfies
\[
y(t, \tau_i) = \frac{1}{\theta} \left\{ \int_{\tau_i}^{\theta + \tau_i} B(\eta) d\eta \chi(t) + \int_{\tau_i}^{\theta + \tau_i} \tilde{A}(t, \eta) d\eta + \int_{0}^{t} \left[ \sigma dW_s(\eta) + \sigma_{\xi} dW_{\xi}(s) \right] \right\}
\]
(20)
In differential form, this observation mechanism becomes
\[
y(t, \tau_i) = \left\{ -\kappa H(\tau_i) \chi(t) + \frac{1}{\theta} \left( f_w(t, \theta + \tau_i) - f_w(t, \tau_i) \right) + \frac{\theta + \tau_i}{\theta} \int_{\tau_i}^{\theta + \tau_i} \frac{\partial A(t, \eta)}{\partial t} d\eta \right\}
\]
(21)
where
\[
f_w(t, x) = \frac{\partial f_w(t, x)}{\partial x} dt + \sigma dW(t, x) + \sigma_{\xi} dW_{\xi}(t)
\]
(22)
f_w(t, 0) = 0
(23)
and
\[
H(t) = \frac{1}{\theta} \int_{\tau_i}^{\theta + \tau_i} B(\eta) d\eta = \frac{1}{\kappa \theta} \left[ e^{-\kappa \tau_i} - e^{-\kappa(\tau_i + \theta)} \right]
\]
We set the observation state as
\[ Y(t) = [y(t, \tau_1), y(t, \tau_2), \cdots, y(t, \tau_m)]. \]

6. PARAMETER ESTIMATION PROBLEM

Our objective now is to estimate the unknown system parameters. Our first difficulty is the covariance function \( q(x, y) \). If we can parametrize it with one or more parameters, say \( \kappa, \chi, \sigma, \lambda, \lambda_{\xi}, \sigma_{\xi}, \) and the seasonality function \( h(t) \). The standard approach is to use the method of maximum likelihood, for which we need to calculate the likelihood function from the observation data \( \{Y(t); 0 \leq t \leq t_f\} \), where \( t_f \) denotes a final time. However, since the observation noise covariance
\[
\Phi = \sigma^2 \delta H(\tau_i) H(\tau_j) + \rho \sigma \sigma_{\xi} H(\tau_i) H(\tau_j) + 2 \sigma^2 \]
\[
+ \frac{\sigma_{\xi}^2}{\sigma^2} \int_{\tau_i}^{\theta + \tau_i} \int_{\tau_j}^{\theta + \tau_j} q(x, z) dz dx \int_{1/2}^{1/2} \int_{1/2}^{1/2} \right\}
\]
(24)
is unknown, we do not have an obvious likelihood functional. Since our model is linear and Gaussian, we may circumvent this problem by working with a quasi likelihood functional as proposed in Bagchi [1975].

The quasi likelihood functional for our problem is
\[
QL(t_f; Y, I) = \int_{0}^{t_f} \left( \mathcal{H} \hat{f}_w(s) \right) + \hat{G}^* dY(s)
\]
where \( \hat{x}(s) \) and \( \hat{f}_w(s) \) are the “best” estimates of the states \( x(s) \) and \( f_w(s) \) given by the observation data \( \sigma \{ Y(t); 0 \leq t \leq s \} \),

\[
\mathcal{H} = [-\kappa H(\tau_i), \frac{1}{\theta} \int_0^{\tau_i} \{ \delta(x - \theta - \tau_i) - \delta(x - \tau_i) \} \, dx] (26)
\]

and

\[
\hat{G} = \left[ \frac{\partial \hat{A}(t, \tau_i)}{\partial t} - \lambda \hat{H}(\tau_i) \right]_i = [(h(\tau_i + \theta + t) - h(\tau_i + t))/\theta + \lambda \hat{H}(\tau_i)]_i (27)
\]

The MLE of the unknown parameters are then given by

\[
\begin{bmatrix} \hat{a}, \hat{b} \end{bmatrix} = \text{argmin}_{a, b} \int_0^1 \left| S(t) - (a + bt) \right|^2 dt.
\]

We get \( \hat{a} = 14.2521, \hat{b} = 32.0052. \)

\[
\mathcal{P}(t) = \begin{pmatrix} P_x(t) & P_{xw}(t) \\ P_{wx}(t) & P_w(t) \end{pmatrix},
\]

\[
\mathcal{P} = \mathcal{P}^* \text{ and } P_x(t) = p_x(t), P_{xw}(t) = p_{xw}(t, x),
\]

where \( p_{xw}(0, x) = p_{w}(0, x, y) = 0. \)

### 7. STATE ESTIMATION PROBLEM

Now we summarize the system and observation mechanism in the usual vector notation;

\[
d \begin{bmatrix} \chi(t) \\ f_w(t, x) \end{bmatrix} = \begin{bmatrix} -\kappa \chi(t) - \lambda \chi \\ \frac{\partial \hat{f}_w(t, x)}{\partial x} \end{bmatrix} dt + \begin{bmatrix} \sigma dW^x(t) \\ \hat{G} dt + \sigma \hat{H} dW^x(t) + K d\hat{w}(t, \cdot) \end{bmatrix},
\]

where \( d\hat{w}(t, x) = \sigma dt(\hat{w}(t, x)) + \sigma dW^x(t) \) and \( dY(t) = \mathcal{H} \begin{bmatrix} \chi(t) \\ f_w(t, \cdot) \end{bmatrix} dt + \hat{G} dt + \sigma \hat{H} dW^x(t) + K d\hat{w}(t, \cdot), \)

where \( \mathcal{H} \) is defined by (26),

\[
\hat{H} = [H(\tau_i)]_{m \times 1}, \text{ and } K(\cdot) = \left[ \frac{1}{\theta} \int_{\tau_i}^{\theta + \tau_i} \cdot \, dz \right]_{m \times 1}.
\]

Under the assumption \( \Phi > 0 \), we can derive the optimal filtering equations from Kallianpur [1980] in p.269.

The optimal estimates for \( x(t) \) and \( f_w(t, x) \) are given by

\[
d \begin{bmatrix} \hat{x}(t) \\ \hat{f}_w(t, x) \end{bmatrix} = \begin{bmatrix} -\kappa \hat{x}(t) - \lambda \hat{x} \\ \frac{\partial \hat{f}_w(t, x)}{\partial x} \end{bmatrix} dt + \begin{bmatrix} \sigma_{\chi} dW^x(t) \\ \hat{G} dt + \sigma_{\hat{H}} dW^x(t) + K d\hat{w}(t, \cdot) \end{bmatrix},
\]

where \( \hat{f}_w(t, \cdot) = \begin{bmatrix} \hat{f}_w(t, x) \end{bmatrix} \),

\[
QK^* = \left[ \frac{1}{\theta} \int_{\tau_i}^{\theta + \tau_i} q(x, y) dy, \frac{1}{\theta} \int_{\tau_m}^{\theta + \tau_m} q(x, y) dy \right],
\]

\[
P_x(t) = p_x(t), P_{xw}(t) = p_{xw}(t, x),
\]

where \( p_{xw}(0, x) = p_{w}(0, x, y) = 0. \)

### 8. SIMULATION STUDIES

First, we simulate the observation data such that it will be similar to the real data. We used a real data set which includes a historical time-series of UK-Gas-NBP spot prices quoted daily from 2-Jan-2007 to 28-Dec-2008.

From this data we identify the parameters as follows:

\[
[a, b] = \text{argmin}_{a, b} \int_0^1 \left| (S(t) - (a + bt)) \right|^2 dt.
\]

We get \( \hat{a} = 14.2521, \hat{b} = 32.0052. \)
• By using FFT, we picked up the first 2 frequencies \( \omega_1, \omega_2 \) from the biggest magnitude:

\[
S(t) - (\hat{a} + \hat{b}t) \sim \sum_{k=1}^{2} \left[ m s_k \sin(2\pi f_k \tau) + m c_k \cos(2\pi f_k \tau) \right]
\]

Table 1 shows the obtained estimates of these parameters. The real data and the fitted curve are shown in (Fig. 1).

Table 1. Estimated parameters

<table>
<thead>
<tr>
<th>( k )</th>
<th>( f_k )</th>
<th>( m s_k )</th>
<th>( m c_k )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.0040</td>
<td>7.9838</td>
<td>1.4593</td>
</tr>
<tr>
<td>2</td>
<td>2.0080</td>
<td>3.1169</td>
<td>0.4337</td>
</tr>
</tbody>
</table>

Fig. 1. Real and estimated curves

The periodic part \( h_p(t) \) of \( h(t) \) is set as

\[
h_p(t) = \sum_{k=1}^{2} \left[ m s_k \sin(2\pi f_k \tau) + m c_k \cos(2\pi f_k \tau) \right]
\]

and this function is shown in (Fig. 2).

From the above initial estimates, we set the system parameters as follows:

\[
\kappa = 1.321, \quad \lambda_\chi = 0.623, \quad \sigma_\chi = 1.2, \\
\lambda_\xi = 0.04, \quad \sigma_\xi = 1.2, \quad \rho = 0.6
\]

The seasonality function is set as

\[
h(t) = 14.2521 + 4.0052t + h_p(t).
\]

The initial conditions for \( \chi, \xi \) are set as

\[
\chi(0) = 0.8, \quad \xi(0) = 20.
\]

We assume that the covariance kernel of \( \sigma w(t, x) \) is given by

\[
\sigma q(x, y) = \sum_{k=1}^{100} \sigma \sin(k\pi x/5) \sin(k\pi y/5),
\]

with \( \sigma = 1 \).

The simulated observation data is shown in (Fig. 3), for \( T - T_0 = 1 \) month. The factor process \( \log \hat{F}(t, x) \) is also demonstrated in (Fig. 4).

8.1 MLE results

We assume that the unknown parameters are \( \kappa, \lambda_\chi, \sigma_\chi, \lambda_\xi, \sigma_\xi, \rho \) and \( \sigma \). For finding MLE, we used the GA-algorithm in MATLAB. The initial values are set as

\[
\kappa = 1.5, \quad \lambda_\chi = 0.5, \quad \sigma_\chi = 0.15, \quad \lambda_\xi = 0.05, \\
\sigma_\xi = 0.1, \quad \rho = 0.5, \quad \sigma = 0.5
\]

with the upper and lower bounds as

\[
1 \leq \kappa \leq 2, \quad 0.1 \leq \lambda_\chi \leq 1.0, \quad 0.1 \leq \sigma_\chi \leq 0.2, \quad 0.01 \leq \lambda_\xi \leq 0.1, \\
0.05 \leq \sigma_\xi \leq 0.3, \quad 0.1 \leq \rho \leq 1.0, \quad 0.1 \leq \sigma \leq 2
\]
The estimates of the parameters are listed in Table 2.

The estimated log $\tilde{F}(t, x)$ with MLE parameters is shown in (Fig. 5).

![Fig. 5. Estimated log $\tilde{F}(t, x)$ process](image)

The true and estimated spot and log $\tilde{F}(t, 1.2698)$ processes are also shown in (Fig. 6) and (Fig. 7).

![Fig. 6. True and estimated $S(t)$-processes](image)

![Fig. 7. True and estimated log $\tilde{F}(t, 1.2698)$-processes](image)

9. CONCLUSION

In this article, we propose a new arbitrage free model for the futures prices of energy. The new model can be used in a mathematically sound way when estimating the parameters of the model using the method of maximum likelihood. The estimation procedures are all performed under the risk neutral measure. By using the estimated parameters, we can also get the estimate of $\chi(t)$ process in the real world. For $\xi(t)$ process, we only identify $\lambda_\xi$ and it is not possible to separate the value of the risk premium term. However, it may be enough to estimate the unknown parameters obtained here to be used in pricing other derivatives.

ACKNOWLEDGEMENTS

We are grateful to Dr. Mahmoud Hamada of Essent Trading for helpful comments.

REFERENCES


