Calibrating mean-reverting jump diffusion models: an application to the NSW electricity market

Jamie Alcock
University of Queensland

Joanne Goard
University of Wollongong

Tony Vassallo
The University of Sydney

1 Introduction

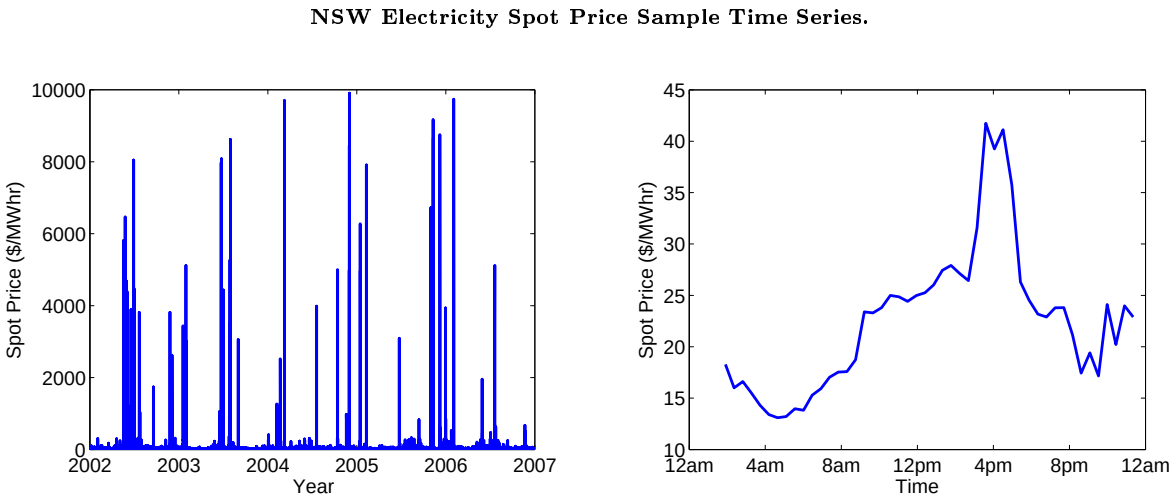
In recent years the Australian National Electricity Market (the NEM) has been undergoing the transition to a fully deregulated marketplace. In December 1998, for the first time, the wholesale price of electricity was subject to market forces. The NEM includes the Queensland, New South Wales, Victorian, South Australian, Tasmanian and Australian Capital Territory electricity markets. Almost 8 million end users are supplied by the world's longest interconnected power system, and the NEM trades up to $7 billion of electricity annually [20].

NEMMCO Ltd (The National Electricity Market Management Company) was established in 1996 to manage the NEM, a role which carries with it the responsibility for setting the spot price. The spot price is determined via a sellers' dutch auction. Each day, each generator submits a complex bid of prices and volumes. The demand fluctuates throughout the day, and every 5 minutes short-term supply and demand are realigned by NEMMCO. The 5 minute dispatch price for all bidders is set to the winning bid of the marginal supplier, and six sequential dispatch prices are averaged to determine the half-hour spot price. All successfully bidding generators receive the spot price for their product. Currently spot prices are artificially bound by NEMMCO to remain between $-\$1000$ and $\$10,000$ per MWhr.

The introduction of market forces to the NEM has provided a plethora of challenges for mathematicians, economists and financial economists[1] amongst others. As electricity is not storable the spot price process is extremely volatile. However the trend underlying the spot price process is highly predictable and highly periodic. The supervolatility of spot price stems from a combination of unexpected events. Unexpectedly high demand (possibly because of an unexpected temperature change) or unexpectedly low supply (because of an unscheduled generator outage or distributional failures /constraints) results in a rapid change in price. There is a strong positive relationship between the spot price spikes and large changes in demand (load).

Retail customers typically pay a fixed price for electricity. Hence the electricity retailer manages a portfolio of floating-for-fixed swap instruments. The retailer receives a fixed pay-

[1] For example see recent and significant works by [5, 7, 21]



(a) NSW Electricity Spot Price Series

(b) Sample Daily Spot Price Series

Figure 1: Time series of NSW Electricity spot prices. Figure (a) provides a 5-year sample of spot prices; (b) provides a daily sample series.

ment for electricity and pays a floating rate to the generator. Managing financial risk in such a volatile market is a formidable challenge. The magnitude of this challenge is highlighted when one considers that the retailers financial risk is the product of load and spot price. Insofar as there is a positive correlation between load and price, high demand implies a doubly heavy loss for the electricity retailer. Understanding the dynamics of the electricity spot-price/demand process is a fundamental step in developing adequate risk-management strategies.

Much of the existing literature on modelling commodity prices deals with the assumption of geometric Brownian motion (GBM) and methods by which the assumption of Gaussian returns can be improved. As such, many papers deal with stochastic volatility [14, 15] and fat tail distributions [11] and related issues. With the emergence of electricity as a traded commodity in the 1990's came the challenge to model the spot price behaviour. As electricity is largely non-storable, market supply and demand must be kept in constant, continuous balance. This results in electricity prices not following a 'smooth' process as other commodities do. As a result, a GBM alone cannot successfully describe the evolution of its price.

Clearly an equilibrium spot price model should include information on supply, or potential supply, in addition to demand. A rapid shift in available supply, via an unscheduled outage, is likely to have a similar impact on price as a rapid shift in demand. However data availability remains an issue. While load data is readily available, potential supply data is more difficult to obtain.

The aim of this article is to develop and calibrate a model for the spot price of electricity as a function of market demand (load), and to calibrate this model using historical load and spot prices from the NSW market. In particular, this model will exploit two, quite apparent, features of historical spot price data: (i) mean-reversion, and (ii) sudden price spikes, or jumps. Hence we shall explore the suitability of fitting a mean-reversion jump-diffusion model to NSW spot price and load data. The justification for, and development of, these models is given in Section 2. The exposition of a calibration methodology is given in Section 3. The

model is fitted to the given data using this calibration technology in Section 4. We conclude in Section 5.

2 Model development

Two notable characteristics of electricity spot prices are mean-reversion and price-spikes (see Figure 1 for a sample time series of the NSW spot price). Electricity prices can also be highly volatile, with warmer seasons exhibiting significantly higher volatility than the colder seasons. This indicates a higher stability of the mean price for cold seasons as compared to the warmer seasons. In order to capture the characteristics of the electricity spot prices, [16] indicates the need to introduce jumps and stochastic volatility into their models. In [3] tested the effectiveness of Brownian motion (BM), Mean Reversion, GBM and Geometric Mean Reversion (GMR) and found that the GMR Model outperformed the others, and that including jumps into the models further improved their performance. Other papers related to electricity pricing include those by [4, 9, 10, 17]. As noted by many of these authors, general diffusion models of the form

$$dS_t = a(S_t, t)dt + b(S_t, t)dZ_t, \tag{1}$$

where $dZ_t \sim N(0, dt)$ is an increment in a Wiener process, fail to capture the large, non-negligible observed spikes in the market.

An obvious extension to the models is to include a jump component modelled by a Poisson distribution. This was first introduced by [18] and then extended by [7, 8] to incorporate the mean-reversion feature. An example of a simple mean reversion jump diffusion model (MRJD) is

$$dS_t = \alpha(S^* - \ln S_t)S_t dt + S_t \sigma dZ_t + S_t K\ dq_t, \tag{2}$$

where α is the mean-reversion rate, S^* is the mean-reversion level, σ the volatility of the spot price, K is the jump size which may, for example, be taken to follow a normal distribution $N(\mu, v^2)$ and dq is a Poisson process so that

$$dq_t = 1 \ \text{ with probability } \ \lambda \text{ dt}, \ \ = 0 \ \text{ with probability } \ 1 - \lambda \text{ dt}, \tag{3}$$

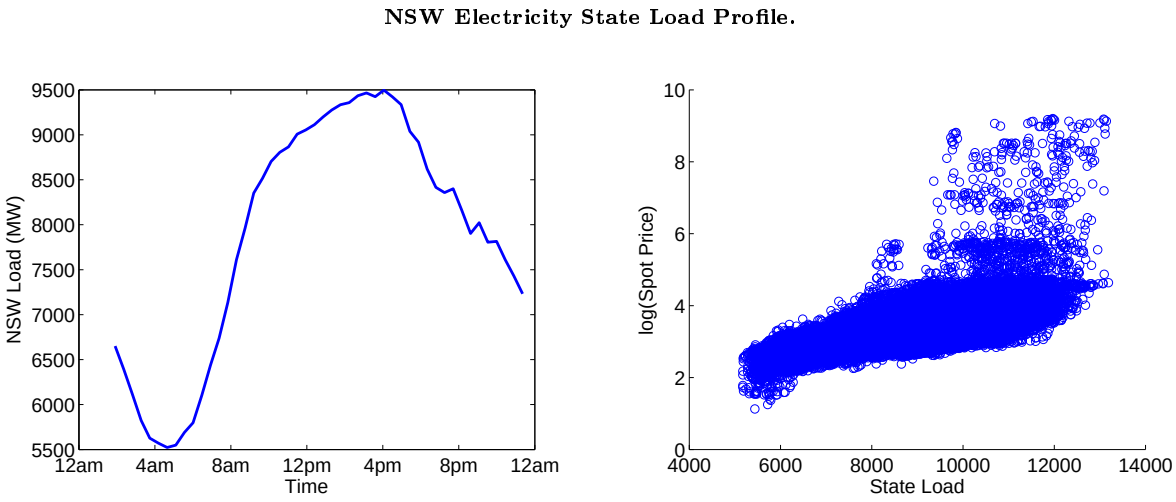
(where λ is the average number of jumps per year). Many variations of (2) exist that might include multiple jump processes, doubly-stochastic jump processes and stochastic volatility (see e.g. [9]). In this article we examine two classes of MRJD and within both classes examine the performance with various forms of volatility functions.

Model 1

As mentioned, spot price is a function of both demand (state load) and available supply. To partially verify this a scatter plot of spot price vs state load is given in Figure 2 (b). Clearly a monotonic relationship between spot price and demand exists. To capture this relationship, we assume that the log price $Q_t (= \ln p_t)$, where p_t is the current spot price of electricity, depends on the load via

$$Q_t = f(L_t) + X_t, \ \ \text{where} \ \ dX_t = -\alpha X_t + \sigma\ dZ_t + K\ dq_t \tag{4}$$

and L_t is the load at time t.



(a) NSW Electricity State Load Sample Daily Time Series

(b) Spot Price vs State Load

Figure 2: Profile of the State Load (Demand) for Electricity. Figure (a) presents a sample day state load - the same time period for Figure 1(b). Figure (b) presents a scatter plot of Spot Price vs State Load.

From (5), X_t follows a mean-reverting process (to a mean of zero), where α is the reversion rate, K is the jump size, dq_t is as in (3) and $dZ_t \sim N(0, dt)$ denotes throughout this paper an increment in a Wiener process. We made no prior assumptions about the distribution of the jump size K.

The dependence of Q_t on load in (4), will be reflected in the stochastic process followed by Q_t. The following forms for the volatility σ_t will be considered:

$$\sigma_t = \sigma_0, \tag{6a}$$
$$\sigma_t = \sigma_0 L_t, \tag{6b}$$
$$\sigma_t = \sigma_0 (dL_t), \tag{6c}$$
$$\sigma_t = \sigma_0 e^{Q_t/2}, \tag{6d}$$
$$\sigma_t = \sigma_0 e^{nQ_t}, \tag{6e}$$
$$\sigma_t = \sigma_0 + \sigma_1 \cos(2\pi t) + \sigma_2 \cos(4\pi t), \tag{6f}$$

where σ_0, σ_1, σ_2 are constants and dL_t denotes the change in the load, L_t, in one time-step.

In volatility model (6a) we assume a constant volatility throughout the period, while in volatility models (6b) and (6c) we assume that volatility depends on the load or the change in the load respectively. As will be seen shortly, models (6d) and (6e) imply a volatility for the stochastic differential equation (SDE) for p_t of the form $\sigma_0 p_t^{3/2}$ and $\sigma_0 p_t^{n+1}$ respectively. These were chosen for comparison with mean-reversion models of the type $dr = (\alpha + \beta r)dt + \sigma r^\gamma dZ$, for which $\gamma = 1.5$ is often the unconstrained estimate [6]. A similar dependence could possibly occur with electricity prices. The volatility model (6f) was chosen in an effort to capture the yearly and half-yearly cycles in volatility.

From (4) and (5) the process followed by Q_t is

$$dQ_t = \alpha \left\{ \frac{1}{\alpha} f'(L_t) L_t' + f(L_t) - Q_t \right\} dt + \sigma_t \, dZ_t + K \, dq_t. \tag{7}$$

Equation (7) represents Model 1. From it we can also derive the process followed by p_t, namely

$$dp_t = \alpha p_t \left\{ \frac{1}{\alpha} f'(L_t) L_t' + \frac{\sigma_t^2}{2\alpha} + f(L_t) - \ln p_t \right\} dt + \sigma_t p_t \, dZ_t + p_t (e^K - 1) dq_t. \tag{8}$$

Note then that the volatility terms (6d) and (6e) imply the volatility forms $\sigma_0 p_t^{3/2} dZ$ and $\sigma_0 p_t^{n+1} dZ$ in (8) respectively.

Model 2

For Model 2 we simply assume that Q_t follows a seasonal pattern in time with

$$Q_t = \ln p_t = h(t) + X_t \tag{9}$$

where dX_t follows the process as in (5) with the volatility as in (6a) - (6f). This then implies the process for Q_t as

$$dQ_t = \alpha \left[\frac{1}{\alpha} h'(t) + h(t) - Q_t \right] dt + \sigma_t dZ_t + K \, dq_t \tag{10}$$

which we call our Model 2. From (10) the process followed by the price p_t itself is

$$dp_t = \alpha p_t \left\{ \frac{1}{\alpha} h'(t) + \frac{\sigma_t^2}{2\alpha} + h(t) - \ln p_t \right\} dt + \sigma_t p_t \, dZ_t + (e^K - 1) p_t \, dq_t. \tag{11}$$

3 Data

Integral Energy Australia provided half-hourly data for the electricity spot prices and load for the years 2002-2006. We generated a new time series for the price by calculating the arithmetic average of the 48 data values for each day and then taking logs of these values. This is done rather than taking the average of the 48 log prices for each day (which would have been equivalent to a geometric average) since we needed to analyse the jumps in the data rather than attempt to mask them. This series is plotted in Figure 3. The daily series is initially formed in order to test the methodology on a smaller dataset.

At first glance the price spikes in Figure 3 are prominent and seem relatively frequent. Statistics for the $Q_t = \ln p_t$ series as well as for the return $d \ln p_t = \ln p_t - \ln p_{t-1}$ are provided with the figure. Notice in particular the high standard deviations in the two series and the very high kurtosis coefficients. The kurtosis estimates are significantly different from 3 (the kurtosis for a normal distribution) under the null hypothesis of normality and imply leptokurtic distributions with high peaks, thin midrange and fat tails. This is supported by the histogram of the returns in Figure 4.

In Figure 5, the average daily load is plotted from which a yearly seasonal pattern is obvious as well as an increasing trend. Statistics for the average daily load and for $d \text{ load}_t = \text{load}_t - \text{load}_{t-1}$ are provided with the figure.

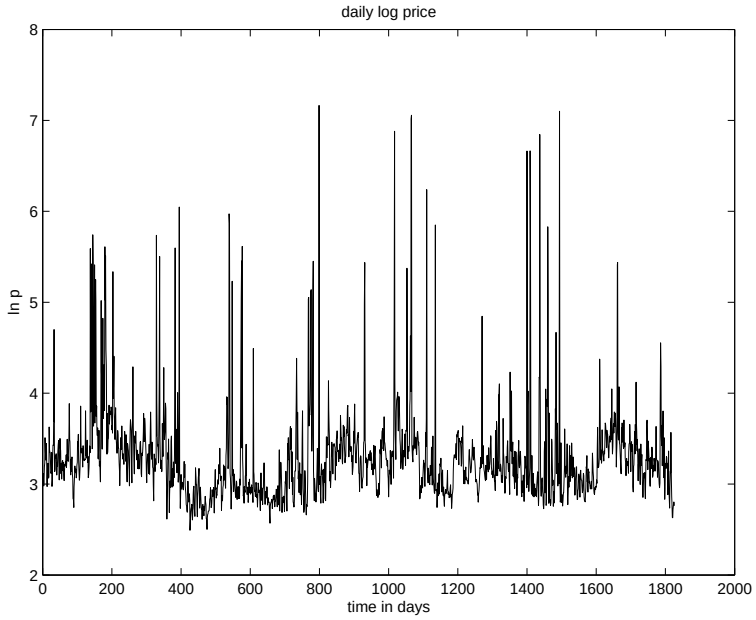


Figure 3: Daily log price series.

Max(ln p)= 7.16472	Max(d ln p)=3.7024
Min(ln p)= 2.4915	Min(d ln p)=-4.0328
Mean(ln p)= 3.2848	Mean(d ln p)=-1.3744
Std(ln p)= 0.518044	Std(d ln p)=0.48745
Kurtosis (ln p) = 18.715	Kurtosis (d ln p)= 25.57

Jumps

The discontinuous jump component in (7) and in (10) controls the nature of the jumps - their frequency and size. In order to estimate the parameters for the jumps (see Section 4.4) we first extract them from the data.

There are various approaches by which this can be achieved depending on what is defined as a 'jump'. For example price returns beyond three standard deviations of the mean could be classified as a jump. We choose a particular threshold jump size js and on finding a return $\ln p_{t+1} - \ln p_t$ larger than js, remove that return and subsequent ones until $\ln p_{t+k} - \ln p_t < js$. The jump size js is chosen so that upon filtering in such a way, the remainder is a series that can be approximated by a normal distribution. A higher js is preferable, as it would imply fewer 'random' jumps.

From an examination of percentage returns within 1, 2 and 3 standard deviations of their mean in the filtered data we chose $js = 0.7$ and $js = 0.6$ (see Table 1). Kurtosis estimates also given in Table 1, are higher than 3, but are significant improvements on the kurtosis estimate of the unfiltered data namely 25.57. We note that a double exponential distribution or Levy process would seem a better choice to model the filtered data. However, here our aim is to model the data using a MRJD model where a jump component is added to a GBM.

Using $js = 0.7$ and 0.6 respectively, the filtered data series and the histogram of their returns are plotted in Figures 6 and 7.

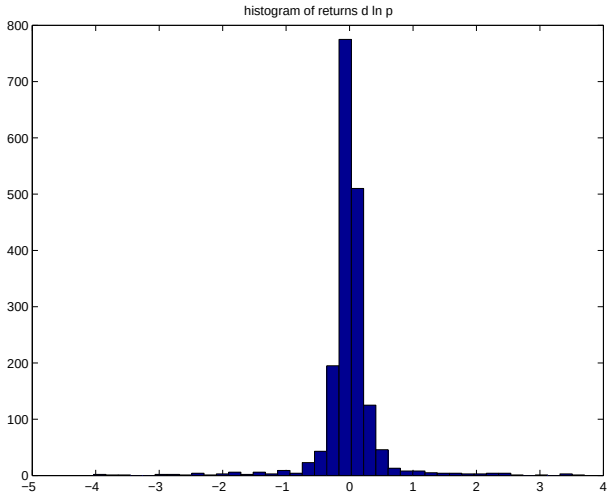


Figure 4: Histogram of returns $d\ln p_t = \ln p_t - \ln p_{t-1}$

js	1 std from mean	2 std from mean	3 std from mean	Kurtosis
0.6	74.4%	93.2%	98.3%	4.7211
0.65	75.9%	93.7%	98.4%	5.3833
0.7	77%	93.66%	98.7%	5.3927
0.75	77.9%	94%	98.9%	5.7651
0.8	78.15%	94.3%	99%	5.8209
0.85	78.6%	94.6%	99.13%	5.8963

Table 1: Percentage returns using various threshold jump sizes js.

4 Calibration

We estimate the functions and the parameters in Models 1 and 2 (equations (7) and (10)) respectively. We do this by first estimating the values in the 'non-jump' part of the process (i.e. the diffusion process) using the filtered data, and then analysing the jump parameters from the jump data set that was extracted from the raw data.

4.1 First approximations for the seasonal function and mean reversion rate

From a plot of the log price versus the load (see Figure 8), a linear relationship is indicated so that we take $f(L_t) = a_0 + a_1 L_t$ in equation (4).

The least squares fit approximations of $a_0 = 1.7565$ and $a_1 = 0.000171$ for $js = 0.7$ and $a_0 = 1.7946$, $a_1 = 0.000166$ for $js = 0.6$ are used as first approximations for the final fit.

For Model 2, we approximate $h(t)$ with a truncated Fourier series

$$h(t) = \sum_{n=0}^{12} (b_n \sin n\pi t + c_n \cos n\pi t) \tag{12}$$

and again find least squares estimates for the coefficients. These are listed in Table 2. We

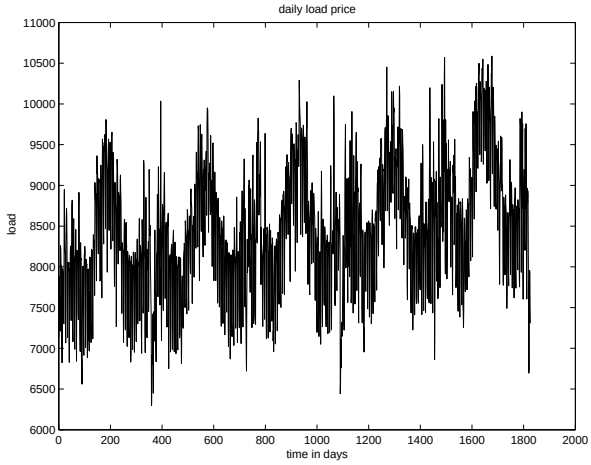


Figure 5: Daily load series.

Mean (daily load) = 8487	mean(d load)=0.0927
Std (daily load) = 767.6803	std(d load) = 542.7508
Max (daily load) = 10586	max(d load) = 1892.9
Min (daily load) = 6297.6	min(d load) = -1894.1

can also estimate the mean reversion rate α in the filtered data using a least squares fit of $(X_{t+1} - X_t, X_t)$ data. For Model 1 we get $\alpha \approx 75.58$ for $js = 0.7$ and $\alpha \approx 69.9$ for $js = 0.6$. For Model 2, $\alpha \approx 119.7$ for $js = 0.7$ and $\alpha \approx 115.34$ for $js = 0.6$.

4.2 The estimation technique for the diffusion process

Paper [13] Generalised Method of Moments (GMM) is used to estimate the parameters in the diffusion processes corresponding to Models 1 and 2. The GMM approach has the advantage that it does not require that the distribution of log price changes be normal. As well the GMM estimators and their standard errors are consistent even if their disturbances ϵ_{t+1} are conditionally heteroscedastic. GMM is used by [6, 1, 12], amongst others, to empirically test interest rate models.

We use a discrete time specification. For Model 1:

$$Q_{t+1} - Q_t = \alpha \left(\frac{1}{\alpha} a_1 \left(\frac{L_t - L_{t-1}}{dt_t} \right) + (a_0 + a_1 L_t) - Q_t \right) dt_{t+1} + \epsilon_{t+1} \tag{13}$$

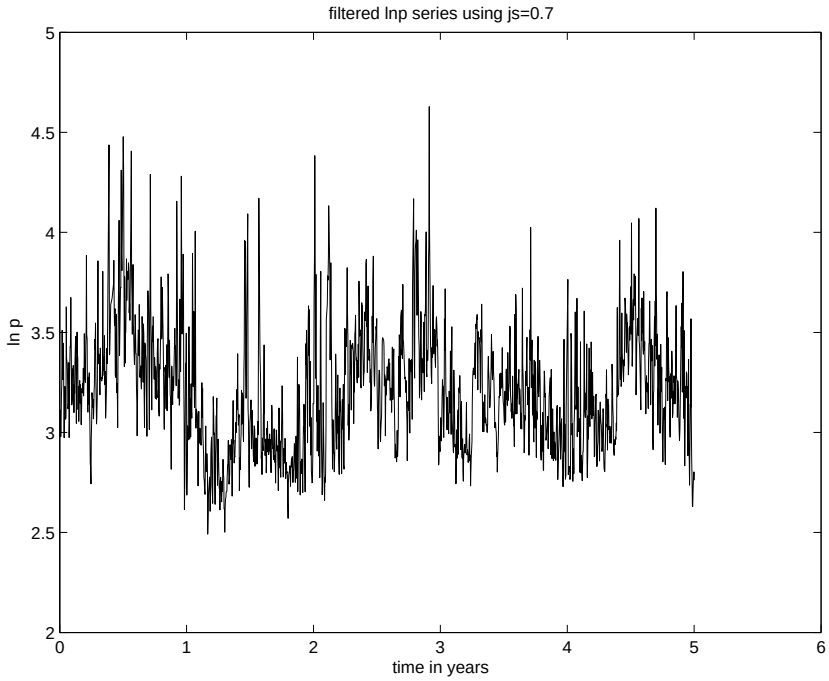
$$E(\epsilon_{t+1}) = 0, \qquad E(\epsilon_{t+1}^2) = \sigma_t^2 dt, \tag{14}$$

where we use a backward difference (BD) approximation rather than a forward difference (FD) approximation for the derivative $L'(t)$, as the load in the next time-step is assumed not to be known.

For Model 2:

$$\begin{aligned} Q_{t+1} - Q_t &= \alpha \left[\frac{1}{\alpha} \sum_{n=1}^{12} (b_n n\pi \cos n\pi t - c_n n\pi \sin n\pi t) \right. \\ &\quad \left. + \sum_{n=0}^{12} (b_n \sin n\pi t + c_n \cos n\pi t) - Q_t \right] dt_{t+1} + \epsilon_{t+1} \end{aligned} \tag{15}$$

(a)



(b)

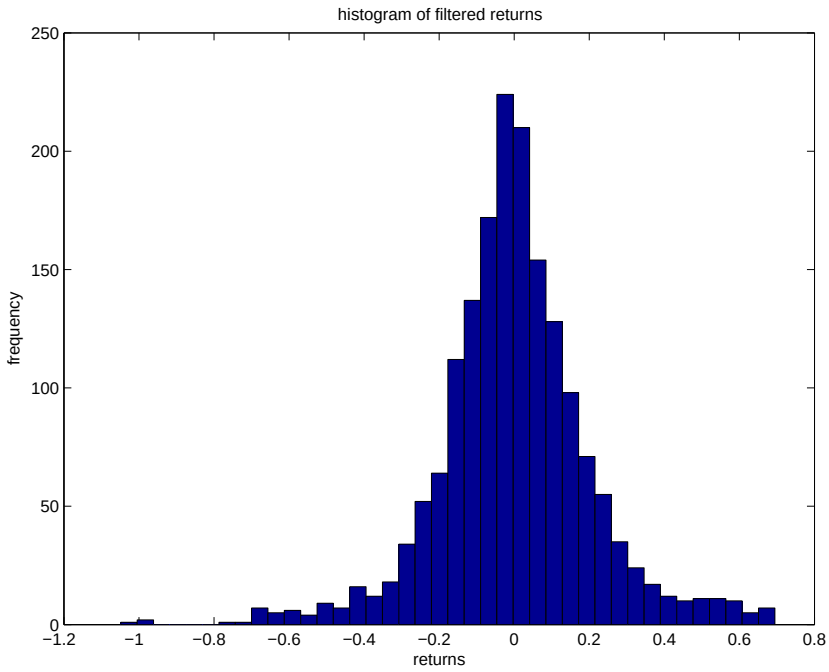
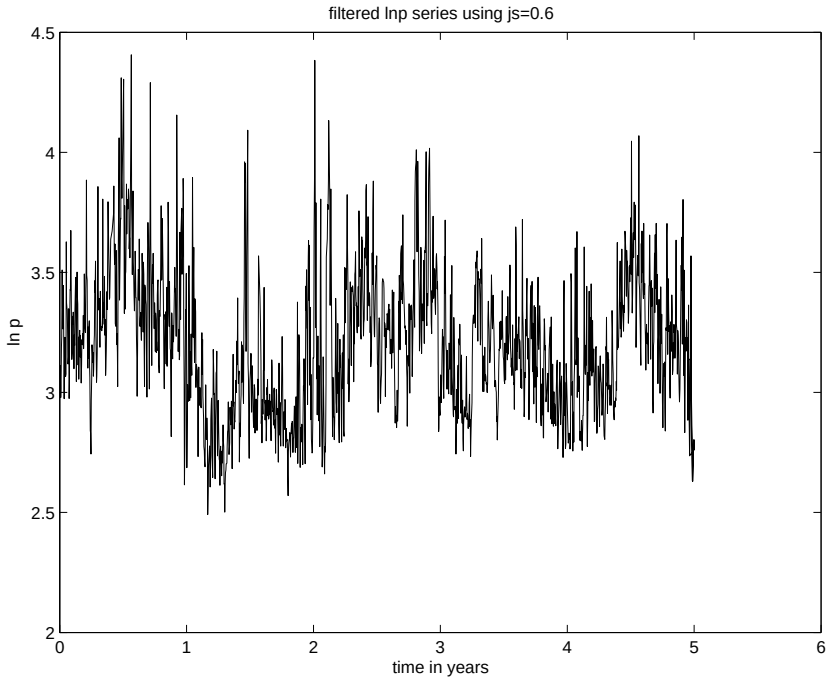


Figure 6: Using $js = 0.7$: (a) The filtered $\ln p$ series; (b) Histogram of filtered returns.

(a)



(b)

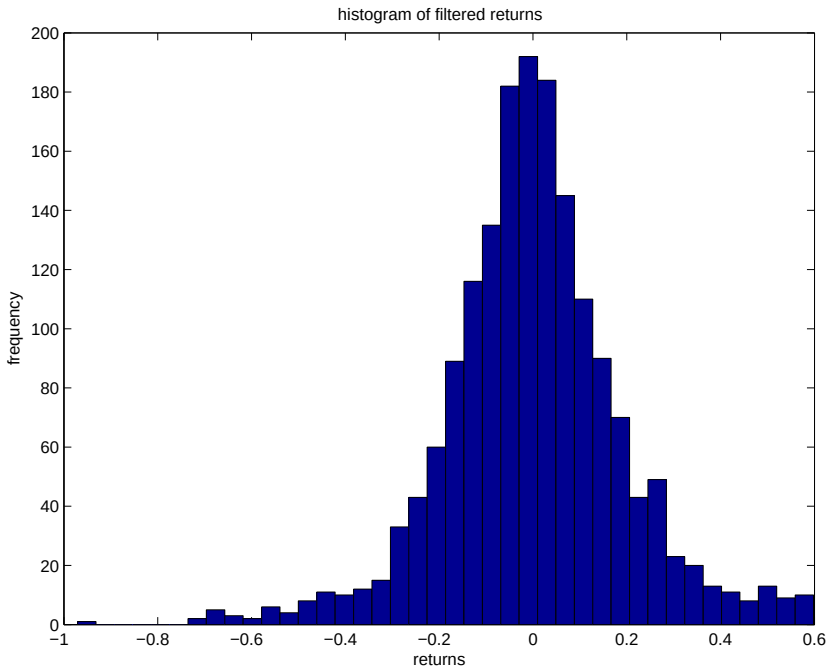


Figure 7: Using $js = 0.6$: (a) The filtered $\ln p$ series; (b) Histogram of filtered returns.

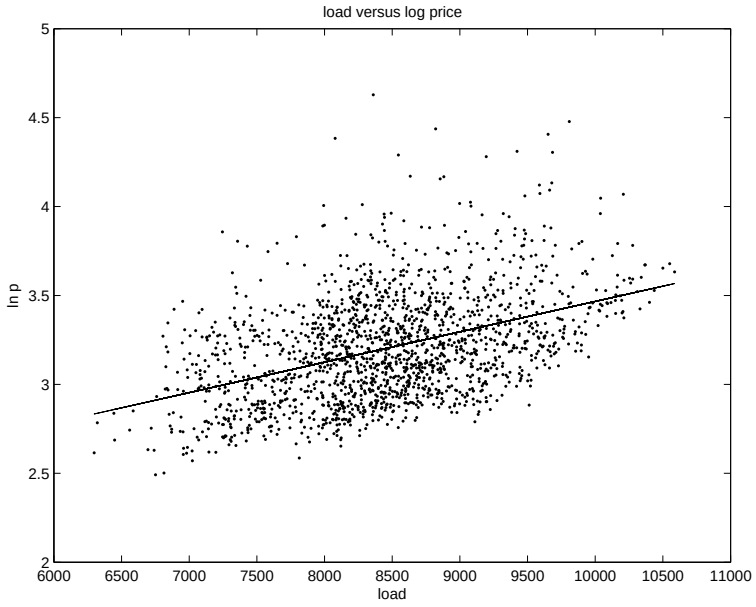


Figure 8: Log price vs. load.

$$E(\epsilon_{t+1}) = 0, \quad E(\epsilon_{t+1}^2) = \sigma_t^2 dt. \tag{16}$$

For a given model we let θ_1 and θ_2 be the parameter vectors for the drift and diffusion components respectively. For example for Model 1, $\theta_1 = (\alpha, a_0, a_1)^T$, and for volatility (6f), $\theta_2 = (\sigma_0, \sigma_1, \sigma_2)^T$. Define the vectors

$$(f_1)_t(\theta_1) = \left(\epsilon_{t+1} \otimes [I_1, \ldots, I_{m_1}]^T\right) \tag{17}$$

and

$$(f_2)_t(\theta_2) = \left((\epsilon_{t+1}^2 - \sigma^2 dt) \otimes [\bar{I}_1, \ldots, \bar{I}_{m_2}]^T\right) \tag{18}$$

where I_i and $\bar{I}_i$, $i = 1 \ldots M$ denote instrumental variables (i.e. predetermined variables that are independent of the errors) and $M = m_1$ or m_2 denote the number of parameters in the corresponding parameter vector.

Under the null hypothesis that (13) and (14) or ((15) and (16)) are true, the orthogonality conditions $E(f_t(\theta)) = 0$ hold where $f = f_1$ or f_2. The GMM technique replaces $E(f_t(\theta))$ with its sample counterpart $g_T(\theta)$, using T observations where

$$g_T(\theta) = \frac{1}{T} \sum_{t=1}^{T} f_t(\theta),$$

and then estimates the parameters in the vector θ that minimises the quadratic form

$$J_T(\theta) = g_T(\theta)^T W g_T(\theta),$$

where W is a positive definite, symmetric, weighting matrix with the sample estimate adjusted for serial correlation and heteroscedasticity using the method of [19] with Bartlett weights.

	$js = 0.7$	p-val	$js = 0.6$	p-val
c_0	3.179222	0.000	3.172496	0.000
b_1	0.173213	0.000	0.172806	0.000
c_1	-0.047912	0.022	-0.046492	0.024
b_2	-0.049416	0.010	-0.045357	0.024
c_2	-0.075053	0.000	-0.074299	0.000
b_3	0.046542	0.007	0.045453	0.010
c_3	-0.012851	0.412	-0.007405	0.511
b_4	-0.018237	0.402	-0.018950	0.349
c_4	0.072677	0.001	0.070253	0.000
b_5	0.049531	0.013	0.047362	0.012
c_5	0.006828	0.620	0.003013	0.872
b_6	0.016508	0.295	0.014970	0.303
c_6	-0.003616	0.654	-0.005839	0.608
b_7	-0.012291	0.611	-0.012048	0.736
c_7	0.008393	0.929	0.010194	0.981
b_8	0.005076	0.649	0.006002	0.719
c_8	0.011422	0.816	0.010589	0.880
b_9	-0.014460	0.485	-0.013943	0.511
c_9	0.007826	0.732	0.005344	0.740
b_{10}	-0.003157	0.786	0.000214	0.953
c_{10}	-0.021833	0.298	-0.021451	0.220
b_{11}	-0.012577	0.429	-0.012940	0.462
c_{11}	0.027130	0.141	0.026998	0.123
b_{12}	-0.005397	0.589	-0.007792	0.635
c_{12}	-0.015081	0.274	-0.011100	0.364

Table 2: Estimates of Fourier series coefficients in equation (12).

Paper [13] shows that setting

$$W = \left(E(f_t(\theta) f_t(\theta)^T)\right)^{-1},$$

delivers an estimate for the vector θ with the smallest asymptotic covariance matrix for the GMM estimates of θ. With the number of unknowns exactly equal to the number of orthogonality conditions, the model is exactly identified and so $J_T(\theta) = 0$.

4.3 Results

Results for Model 1 using $js = 0.7$ are listed in Table 3. For Model 2 using $js = 0.7$ the Fourier coefficients in Table 2 are used as no significant improvements in SSE were found with GMM. The other parameters estimated are listed in Table 4. Tables 5 and 6 list the results for Models 1 and 2 respectively when the data is filtered using $js = 0.6$.

For both threshold jump sizes $js = 0.6$ and 0.7, Model 2 outperforms Model 1. The performance of the volatility is also consistent. From an examination of the SSE (sums of squares of errors) and the MSE (mean square errors), the volatility $\sigma_t = \sigma_0 e^{nQ_t}$ performs best in all models, followed by $\sigma_t = \sigma_0 e^{Q_t/2}$ then $\sigma_t = \sigma_0 + \sigma_1 \cos 2\pi t + \sigma_2 \cos 4\pi t$. Volatility with a dependence on load does not perform significantly better than a constant volatility $\sigma_t = \sigma_0$. The volatility with a dependence on the change in the load performs least well, perhaps not surprisingly as the change in load related to the previous time-step.

		parameter value	t value	p value	SSE (MSE) Eq (1)	Eq (2)
	α	67.51146	10.28	<.0001		
	a_0	3.593127	24.72	<.0001	67.682	
	a_1	-0.00005	-2.73	0.0064	(.0389)	
$\sigma_t = \sigma_0$	σ_0	3.791098	31.28	<.0001		12.0717 (.00694)
$\sigma_t = \sigma_0 L_t$	σ_0	0.000445	31.28	<.0001		12.1124 (.00696)
$\sigma_t = \sigma_0 dL_t$	σ_0	0.003843	4.08	<.0001		13.9356 (.00891)
$\sigma_t = \sigma_0 e^{Q_t/2}$	σ_0	0.745511	30.23	<.0001		11.106 (.00683)
$\sigma_t = \sigma_0 e^{nQ_t}$	σ_0	0.376201	3.88	0.0001		
	n	0.696733	8.89	<.0001		11.0441 (.00635)
$\sigma_t = \sigma_0 +$	σ_0	3.672192	34.15	<.0001		
$\sigma_1 \cos 2\pi t +$	σ_1	0.183404	1.11	0.2681		11.819
$\sigma_2 \cos 4\pi t$	σ_2	0.669647	4.05	<.0001		(.0068)

Table 3: Parameter estimates for Model 1 with $js = 0.7$.

In Model 1 we notice that a_1, the coefficient of L in the function $f(L_t)$ is negative for both thresholds. As the log price is an increasing function of the load, this indicates that the BD approximation for the derivative $L'(t)$ is not a good estimate for the slope. Re-estimating the parameters for Model 1 using a FD approximation for $L'(t)$ using volatilities (6a), (6d) and (6e) yield the values in Tables 7 and 8.

The estimates for a_1 are positive. As well, comparing SSE in Tables 4 and 7 and again in Tables 6 and 8, Model 1 outperforms Model 2. However, for the purpose of forecasting, a FD approximation could not be used as it is anticipatory and assumes knowledge of load in the next time-step.

As an alternative measure to gauge relative performance of the models, we test their forecast powers for (filtered) price changes and squared (filtered) price changes using the coefficient of determination R^2 metric. The R^2 metrics of the drift and diffusion measure the proportion of the total variation in the ex-post price changes (R_1^2) and squared price changes (R_2^2) that can be explained by the conditional expected price changes and conditional volatility measures. The metric provides information about how well each model is able to forecast the future level and volatility of the filtered price. The order of performance in the R^2 statistics from best to worst is naturally the same as the order of performance of the models by examination of the SSE. Model 1 using the FD approximation for the derivative $L'(t)$ provides the best measures with values of $R_1^2 = 21.36\%$ and 31.5% and $R_2^2 = 44.2\%$ and 47% for $js = 0.7$ and 0.6 respectively. The forecasting ability of this SDE is very good compared to other stochastic models.

The best values for Model 2 were $R_1^2 = 15.21\%$ with $js = 0.7$ and $R_1^2 = 15.1\%$ with $js = 0.6$ and $R_2^2 = 35.7\%$ and 33% with $js = 0.6$ and $js = 0.7$ respectively and volatility

		parameter value	t value	p value	SSE (MSE) Eq (1)	Eq (2)
	α	85.13754		0.1141	61.5419 (.0353)	
$\sigma_t = \sigma_0$	σ_0	3.598054	30.97	<.0001		9.7136 (.0055)
$\sigma_t = \sigma_0 L_t$	σ_0	0.000422	30.93	<.0001		9.7736 (0.0055)
$\sigma_t = \sigma_0 dL_t$	σ_0	0.00357	4.24	<.0001		11.2403 (0.0064)
$\sigma_t = \sigma_0 e^{Q_t/2}$	σ_0	0.71121	29.34	<.0001		8.9661 (0.00515)
$\sigma_t = \sigma_0 e^{nQ_t}$	σ_0	0.303368	3.74	.0002		8.9114
	n	0.744553	9.16	<.0001		(0.00512)
$\sigma_t = \sigma_0 +$	σ_0	3.477864	33.54	<.0001		9.5688
$\sigma_1 \cos 2\pi t +$	σ_1	0.236683	1.50	0.1333		(.0055)
$\sigma_2 \cos 4\pi t$	σ_2	0.626515	4.15	<.0001		

Table 4: Parameter estimates for Model 2 with $js = 0.7$.

(6e). Model 1 with a BD approximation for $L'(t)$ does not perform as well.

4.4 Estimating the jump parameters

As described in Section 2, the jumps are filtered from the data using a threshold jump size of $js = 0.7$ and also using $js = 0.6$. From the extracted jump data we estimate the parameters in the jump component in Models 1 and 2 (equations (7) and (10) respectively).

We assume that the jumps were instantaneous to the first approximation. In effect we collapse jumps that may have had a duration of 2 days (which occurred 34.78% of the time) into an instantaneous jump on a single day. Hence in the filtering process, when a return is found above the threshold, the jump size is then taken as the sum of this return and if applicable, the subsequent positive return after this. This is necessary for the model to attain extreme jumps. The mean-reversion characteristic of the models allow the prices to return quickly to the pre-jump level without the need to impose any conditions on the jumps.

Statistics for the jumps are provided in Table 9 and their histograms given in Figure 9.

With the distribution of the jumps difficult to determine with so few data points, and the frequency of jumps inconsistent across the months, we use for the term $K\ dq_t$ in (7) and (10)

$$\sum_{j=1}^{12} K_j\ dq_j\ M_{tj}$$

where $M_{tj} = 1$ if the date t belongs to the jth calendar month and 0 otherwise, while $dq_j = 1$ with probability λ_j, and $dq_j = 0$ with probability $1 - \lambda_j$. Each K_j is a random number from a discrete set ψ_j of equally possible elements. These are listed in Table 10.

We investigate the relationship between jumps and load. For the 5 years of data provided, jumps occurred when the load in the previous time period was above the average load for

		parameter value	t value	p value	SSE (MSE) Eq (1)	Eq (2)
	α	64.88192	10.38	<.0001	58.1593	
	a_0	3.541309	26.61	<.0001	(0.0338)	
	a_1	-0.00004	-2.62	.0089		
$\sigma_t = \sigma_0$	σ_0	3.477373	34.96	<.0001		7.2578 (.00421)
$\sigma_t = \sigma_0 L_t$	σ_0	0.000408	35.22	<.0001		7.2001 (.00417)
$\sigma_t = \sigma_0 dL_t$	σ_0	0.002767	2.73	.0064		8.887 (.00515)
$\sigma_t = \sigma_0 e^{Q_t/2}$	σ_0	0.683717	33.52	<.0001		6.7307 (.0039)
$\sigma_t = \sigma_0 e^{nQ_t}$	σ_0	0.400374	3.78	.0002		6.6874
	n	0.656174	8.03	<.0001		(0.00388)
$\sigma_t = \sigma_0 +$						
$\sigma_1 \cos 2\pi t +$	σ_0	3.38278	39.48	<.0001		
$\sigma_2 \cos 4\pi t$	σ_1	0.172996	1.19	0.2337		7.0723
	σ_2	0.623552	4.41	<.0001		(0.0041)

Table 5: Parameter estimates for Model 1 with $js = 0.6$.

		parameter value	t value	p value	SSE (MSE) Eq (1)	Eq (2)
	α	87.81233	0.88	0.3769	52.9991 (.0307)	
$\sigma_t = \sigma_0$	σ_0	3.264612	36.04	<.0001		5.8113 (.00337)
$\sigma_t = \sigma_0 L_t$	σ_0	0.000384	37.01	<.0001		5.7977 (.00336)
$\sigma_t = \sigma_0 dL_t$	σ_0	0.00274	3.33	0.0009		7.1084 (.00412)
$\sigma_t = \sigma_0 e^{Q_t/2}$	σ_0	0.644729	38.67	<.0001		5.4292 (.00315)
$\sigma_t = \sigma_0 e^{nQ_t}$	σ_0	0.373755	4.16	<.0001		5.4036
	n	0.662956	8.9	<.0001		(0.00313)
$\sigma_t = \sigma_0 +$	σ_0	3.226232	39.18	<0.0001		5.6859
$\sigma_1 \cos 2\pi t +$	σ_1	0.222965	1.63	0.1033		(0.0033)
$\sigma_2 \cos 4\pi t$	σ_2	0.594682	4.45	<0.0001		

Table 6: Parameter estimates for Model 2 with $js = 0.6$.

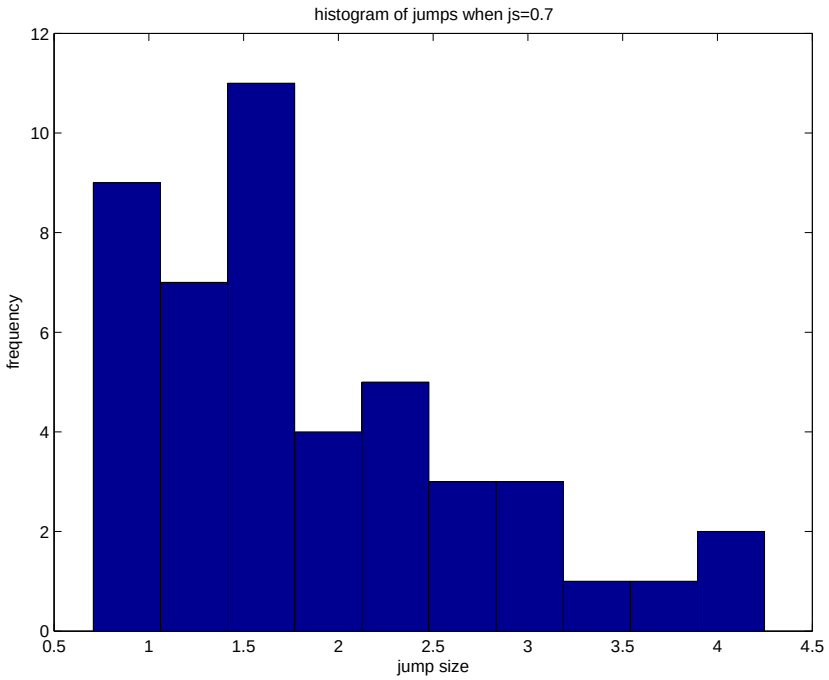
		parameter value	t value	p value	SSE (MSE) Eq (1)	Eq (2)
	α	67.68492	11.52	<0.0001		
	a_0	0.223646	0.26	0.7963		
	a_1	0.000353	3.46	0.0006	57.075 (.0328)	
$\sigma_t = \sigma_0$	σ_0	3.380626	34.27	<.0001		8.153 (.0046)
$\sigma_t = \sigma_0 e^{Q_t/2}$	σ_0	0.683908	27.87	<.0001		7.5619 (.00433)
$\sigma_t = \sigma_0 e^{nQ_t}$	σ_0	0.304197	3.28	0.0011		7.4219
	n	0.732653	7.87	<.0001		(0.0042)

Table 7: Parameter estimates for Model 1 using $js = 0.7$ and FD for $L'(t)$.

		parameter value	t value	p value	SSE (MSE) Eq (1)	Eq (2)
	α	62.50735	12.06	<0.0001	42.7666	
	a_0	0.935759	1.55	.1210	(.0248)	
	a_1	0.000268	3.77	.0002		
$\sigma_t = \sigma_0$	σ_0	2.925856	34.88	<.0001		4.737 (.0027)
$\sigma_t = \sigma_0 e^{Q_t/2}$	σ_0	0.580077	35.91	<.0001		4.484 (.0026)
$\sigma_t = \sigma_0 e^{nQ_t}$	σ_0	0.288463	3.31	0.001		4.453
	n	0.708476	7.62	<.0001		(.0025)

Table 8: Parameter estimates for Model 1 using $js = 0.6$ and FD for $L'(t)$.

(a)



(b)

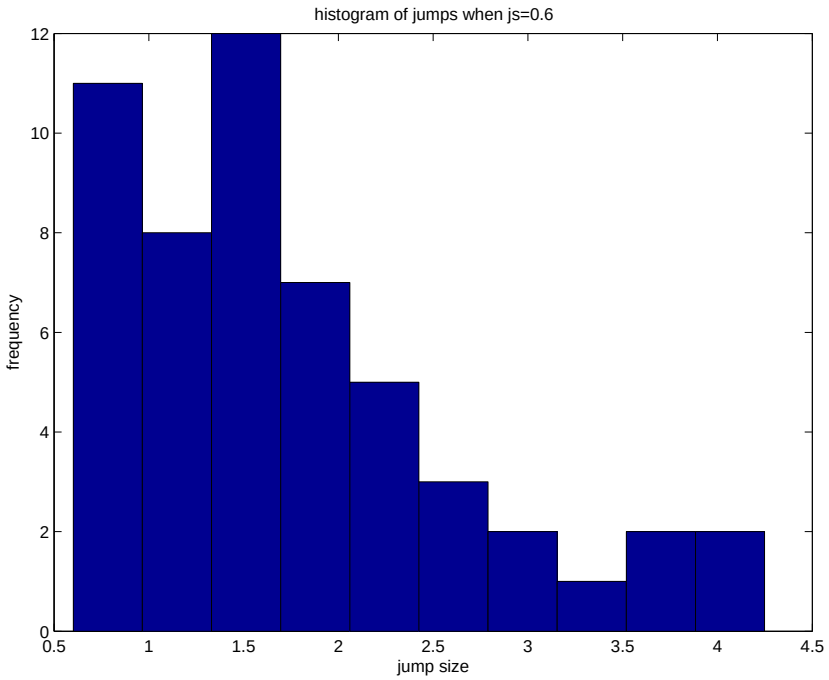


Figure 9: Histogram of jumps using (a) $js = 0.7$ and (b) $js = 0.6$.

	$js = 0.7$	$js = 0.6$
frequency of jumps	46	53
mean of jumps	1.8621	1.7255
std of jumps	0.8818	0.9418
max jump	4.2484	4.2484
min jump	0.7074	0.6015

Table 9: Statistics for 'jumps'.

that month in the year. This occurred 9 out of 12 times in the first year, 4 out of 7 times in the second year, 6 out of 9 times in the third year, 7 out of 11 times in the fourth year and 3 out of 7 times in the fifth year. 85% of jumps occurred when the current load was above the average load for its month/year and all jumps occurred when either the load in the previous time period or the load in the current time period to the jump were above the average load for the month/year when the jump occurs.

This information is only be useful for predicting jumps if we can predict the monthly averages for the loads and the load for the next time period. Regression estimates for the average monthly loads can be used to predict future monthly averages. Using the five year data provided, approximate estimates are

$$\begin{aligned}
y_{\text{jan}} &= 7.1029 + 1.0271t - 0.3343t^2 + 0.0392t^3 \\
y_{\text{feb}} &= 7.0833 + 1.0903t - 0.2816t^2 + 0.0276t^3 \\
y_{\text{march}} &= 7.7564 - 0.0026t + 0.0354t^2 \\
y_{\text{april}} &= 7.5167 + 0.1342t + 0.0013t^2 \\
y_{\text{may}} &= 8.3784 - 0.1209t + 0.0478t^2 \\
y_{\text{june}} &= 9.0757 - 0.3430t + 0.1021t^2 \\
y_{\text{july}} &= 8.7571 + 0.5941t - 0.2405t^2 + 9.1434t^3 \\
y_{\text{aug}} &= 8.5218 + 0.1136t + 0.0164t^2 \\
y_{\text{sep}} &= 7.3469 + 0.8225t - 0.2285t^2 + 0.0239t^3 \\
y_{\text{oct}} &= 7.7909 - 0.0047t + 0.0262t^2 \\
y_{\text{nov}} &= 8.0474 - 0.1060t + 0.047t^2 \\
y_{\text{dec}} &= 7.6714 + 0.1737t
\end{aligned}$$

where t represents the time (in years). With more data these estimates could be improved.

4.5 Simulations

Simulations of Model 2 with volatilities of the form (6a) and (6e) and using parameter values as in Tables 4 and 6 are provided in Figures 10 – 13 with jump size thresholds $js = 0.7$ and 0.6 respectively.

As the volatility (6e) depends on the price level, in the simulation, if a jump occurred at one time, t, then dZ_t, the random term in the volatility in the next time step $t + dt$ is set to its expected value of 0. Also, jumps were not allowed if the price was above 4.2. From a comparison of the trajectories and histogram of returns in Figures 10 and 12 with the

Calendar Month	$js = 0.7$	$js = 0.6$
1	$\lambda_1 = 6/155$ $\psi_1 = \{0.7, 1, 1, 1.7, 2.5, 2.9\}$	$\lambda_1 = 7/155$ $\psi_1 = \{.62, .72, 1, 1.1, 1.7, 2.5, 2.9\}$
2	$\lambda_2 = 6/141$ $\psi_2 = \{1, 1, 1.5, 1.6, 2.5, 4.1\}$	$\lambda_2 = 7/141$ $\psi_2 = \{.62, 1.1, 1.5, 1.6, 2.5, 4.1\}$
3	$\lambda_3 = 1/155$ $\psi_3 = \{4.25\}$	$\lambda_3 = 1/155$ $\psi_3 = \{4.25\}$
4	$\lambda_4 = 1/150$ $\psi_4 = \{0.71\}$	$\lambda_4 = 1/150$ $\psi_4 = \{0.71\}$
5	$\lambda_5 = 5/155$ $\psi_5 = \{1.1, 1.3, 1.6, 2.24, 2.3\}$	$\lambda_5 = 5/155$ $\psi_5 = \{1.1, 1.3, 1.6, 2.24, 2.3\}$
6	$\lambda_6 = 6/150$ $\psi_6 = \{1.4, 1.5, 1.5, 1.6, 1.8, 2.4\}$	$\lambda_6 = 6/150$ $\psi_6 = \{1.4, 1.5, 1.5, 1.6, 1.8, 2.4\}$
7	$\lambda_7 = 6/155$ $\psi_7 = \{1.3, 1.77, 2, 2, 2.2, 2.2\}$	$\lambda_7 = 6/150$ $\psi_7 = \{1.77, 1.9, 2, 2, 2.2, 2.2\}$
8	$\lambda_8 = 1/155$ $\psi_8 = \{0.93\}$	$\lambda_8 = 2/155$ $\psi_8 = \{0.66, 0.93\}$
9	$\lambda_9 = 2/150$ $\psi_9 = \{1.3, 1.7\}$	$\lambda_9 = 4/150$ $\psi_9 = \{0.66, 0.69, 1.3, 1.7\}$
10	$\lambda_{10} = 2/155$ $\psi_{10} = \{3.2, 3.7\}$	$\lambda_{10} = 2/155$ $\psi_{10} = \{3.7, 3.8\}$
11	$\lambda_{11} = 6/150$ $\psi_{11} = \{1, 1.5, 1.8, 2.5, 3, 3.5\}$	$\lambda_{11} = 7/150$ $\psi_{11} = \{0.69, 1, 1.5, 1.8, 2.5, 3, 3.5\}$
12	$\lambda_{12} = 4/150$ $\psi_{12} = \{0.81, 1, 1.3, 1.3\}$	$\lambda_{12} = 5/150$ $\psi_{12} = \{0.6, 0.81, 1, 1.3, 1.3\}$

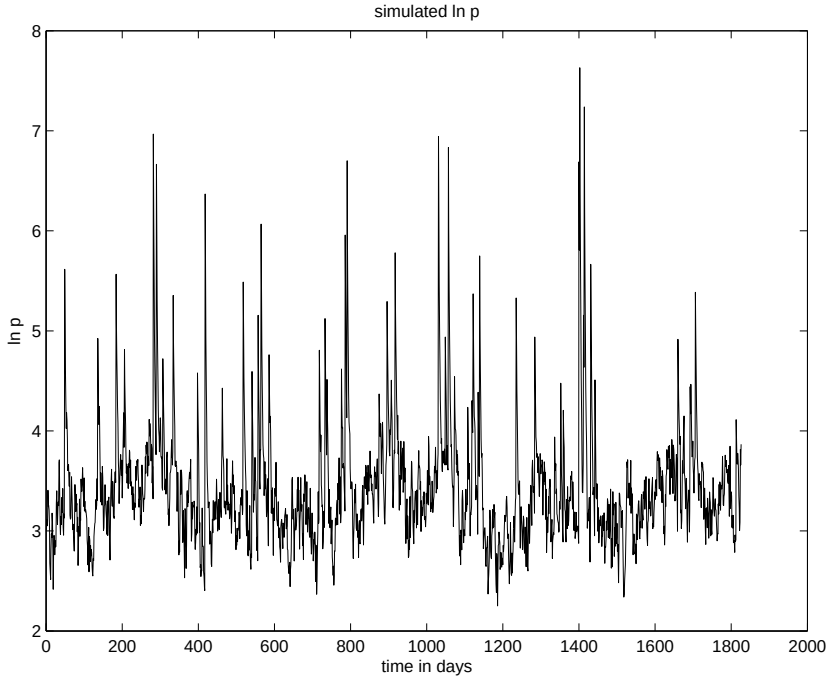
Table 10: Data values for the jump component $k\ dq$.

trajectory in Figures 3 and 4, our models seem to capture the critical features observed in the spot electricity market, and reproduce the price levels observed in the data. Further, the volatility of the form $\sigma_t = \sigma_0 e^{nQ_t}$ in particular, matches the observed magnitude of price fluctuations.

Acknowledgements

Thanks to Philip Broadbridge, Pam Davy, David Griffiths and an anonymous referee for their helpful comments and suggestions.

(a)



(b)

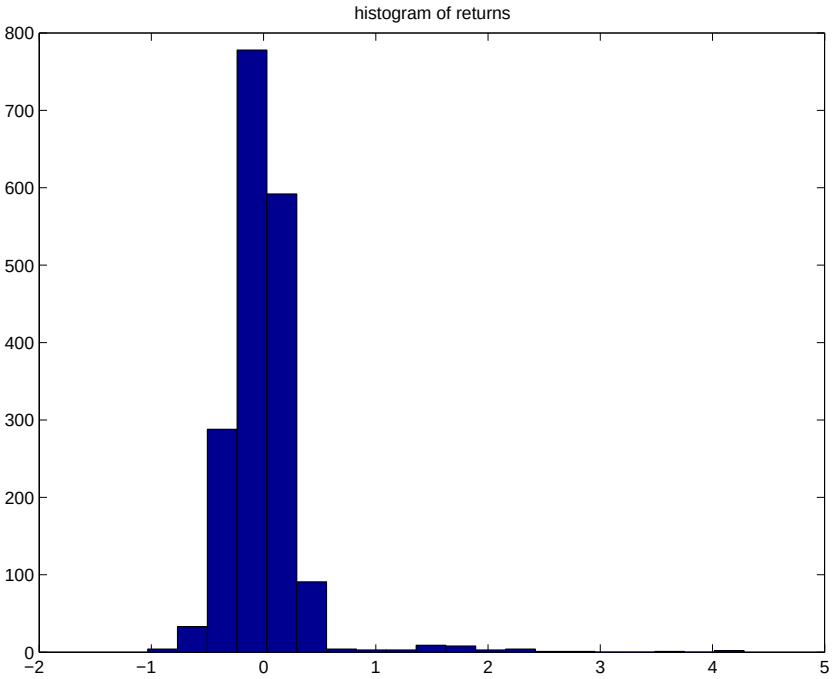
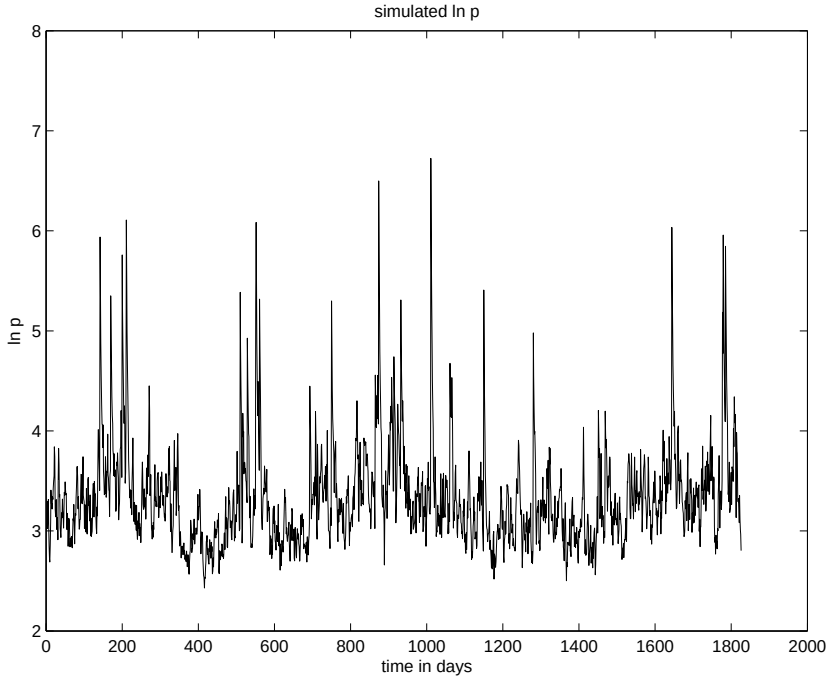


Figure 10: Simulated time series of prices and corresponding histograms of returns, simulated using Model 2 and parameter estimates in Table 4 when $\sigma_t = \sigma_0$.

(a)



(b)

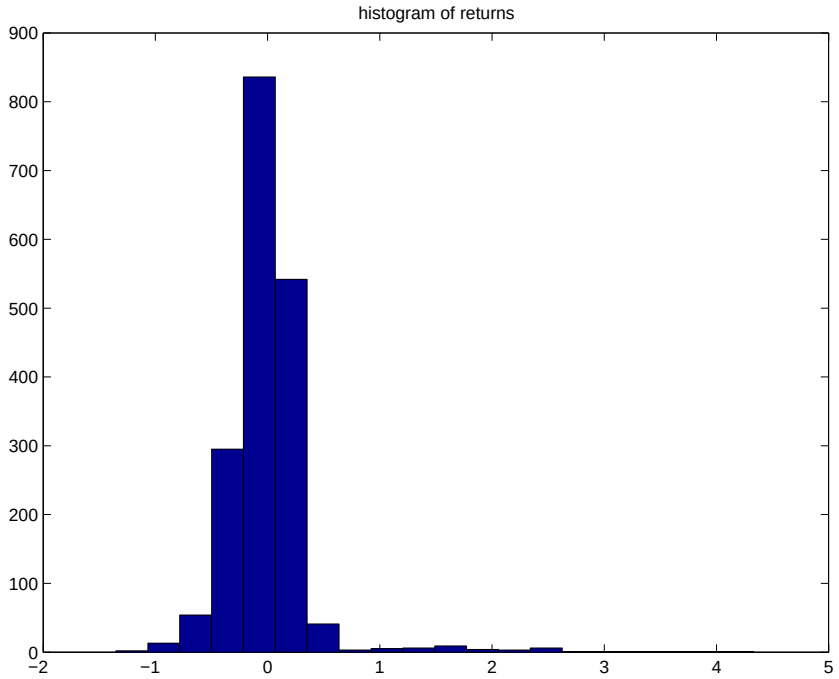
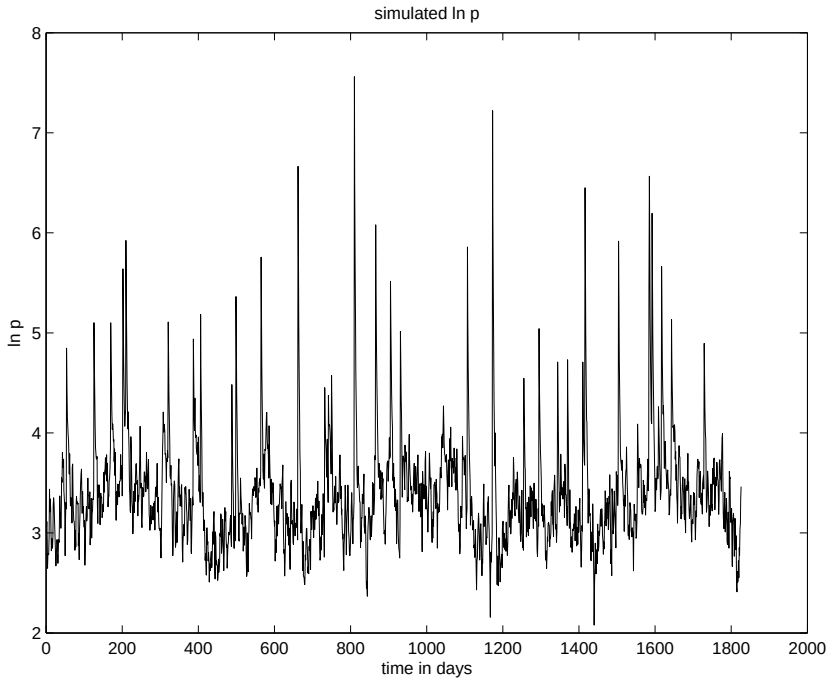


Figure 11: Simulated time series of prices and corresponding histograms of returns, simulated using Model 2 and parameter estimates in Table 4 when $\sigma_t = \sigma_0 e^{nQ_t}$.

(a)



(b)

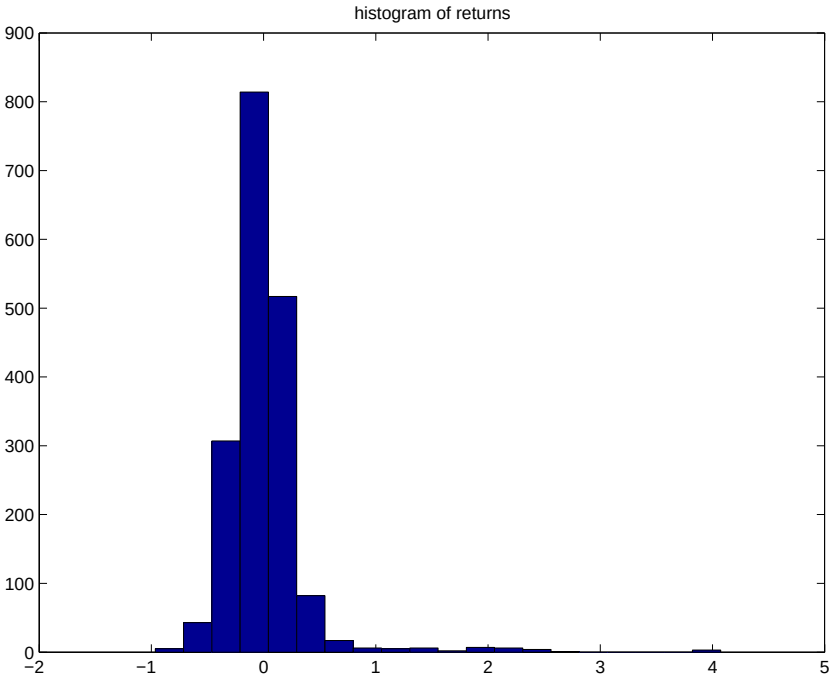
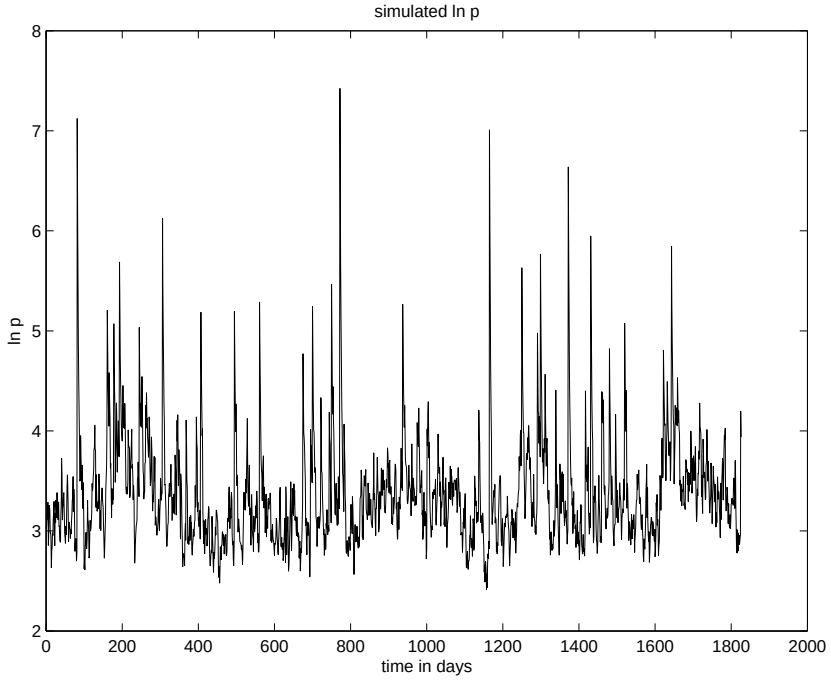


Figure 12: Simulated time series of prices and corresponding histograms of returns, simulated using Model 2 and parameter estimates in Table 6 when (a) $\sigma_t = \sigma_0$.

(a)



(b)

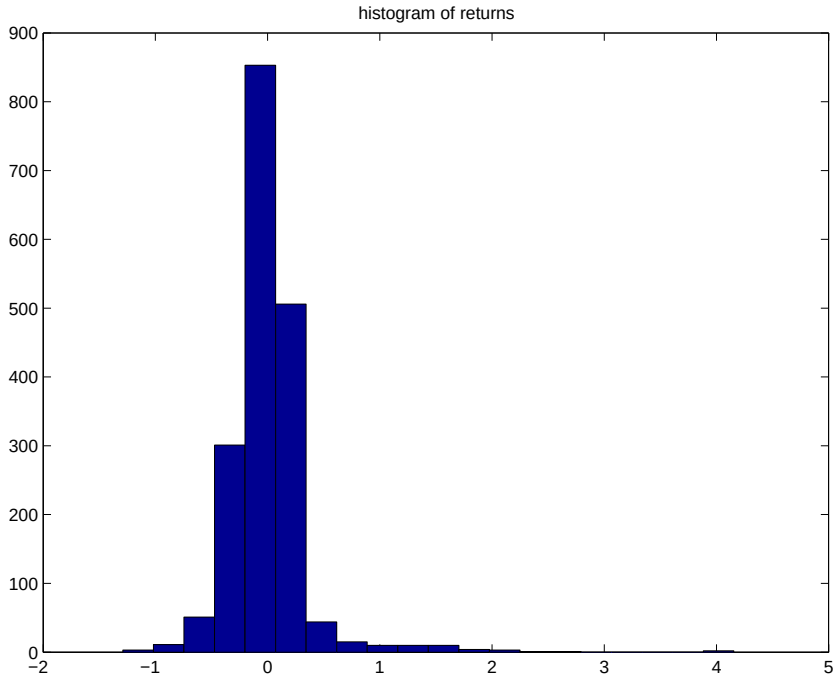


Figure 13: Simulated time series of prices and corresponding histograms of returns, simulated using Model 2 and parameter estimates in Table 6 when $\sigma_t = \sigma_0 e^{nQ_t}$.

References

[1] Ahn, D. & Gao, B. (1999) A parametric nonlinear model of the term structure dynamics, *Review of Fin. Stud.*, **12** (4), 721-762.

[2] Alcock, J. & Burrage, K. (2004) A genetic estimation algorithm for parameters of stochastic ordinary differential equations, *Comp. Stat. and Data Anal.*, **47** (2), 255–275.

[3] Barz, G. & Johnson, B. (1999) *Energy modelling and the management of uncertainty*, RISK Books.

[4] Blanco, C. & Soronow, D. (2001) Jump diffusion processes - energy price processes used for derivatives pricing and risk management, *Commodities Now*, 83-87.

[5] Chan, K. (2006) *Modelling Short-Term Interest Rates and Electricity Spot Prices*, PhD Thesis, UQ Business School, The University of Queensland.

[6] Chan, K.C., Karolyi, A., Longstaff, F.A. & Sanders, A.B. (1992) An empirical comparison of alternative models of the short-term interest rate, *J. of Fin.*, **47** (3), 1209-1227.

[7] Clewlow, L. & Strickland, C. (2000) *Energy derivatives, pricing and risk management*, Lacima Publications.

[8] Clewlow, L., Strickland, C. & Kamski, V. (2001) (Feb) *Extending mean reversion jump diffusion*, Energy Power Risk Management, Risk Waters Group.

[9] Deng, S. (1999) *Stochastic models of energy commodity prices and their applications, mean reversion with jumps and spikes*, Working paper, University of California, Berkeley.

[10] Eydeland, A. & Geman, H. (1998) (Oct) Pricing power derivatives, Risk, 71-73.

[11] Focardi, S.M. & Fabozzi, F.J., (2003) (Fall) Fat tails, scaling and stable laws: a critical look at modeling extremal events in financial phenomena, *J. of Risk Fin.*, London, **5** (1), 51.

[12] Goard, J. & Hansen, N. (2004) Comparison of the performance of a time-dependent short-interest rate model with time-independent models, *Appl. Math. Fin.*, **11**, 147-164.

[13] Hansen, L.P. (1982) Large sample properties of generalised method of moments estimators, *Econometrica*, **50** (4), 1029-1054.

[14] Heston, S. (1993) A closed form solution for options with stochastic volatility with applications to bond and currency options, *Review of Fin. Stud.*, **6** (2), 237-343.

[15] Hull, J.C. & White, A. (1987) The pricing of options on assets with stochastic volatility, *J. of Fin.*, **42** (2), 281-300.

[16] Kamski, V. (1997) The challenge of pricing and risk managing electricity derivatives, Chapter 10 in *The US Power Market*, London: Risk Publications.

[17] Lucia, J. & Schwartz, E. (2002) Electricity prices and power derivatives: Evidence from the Nordic Power Exchange, *Review of Der. Research*, **5**, 5-50.

[18] Merton, R.C. (1976) Option pricing when underlying stock returns are discontinuous, *J. of Fin. Econ.*, **3** (1), 125-144.

[19] Newey, W. & West, K. (1987) A simple, positive semi-definite heteroscedasticity and autocorrelated consistent covariance matrix, *Econometrica*, **55** (2), 703-708.

[20] NEMMCO Ltd, Melbourne (2004) *An Introduction to Australia's National Electricity Market.*

[21] Simshauser, P. (2001) *Microeconomic Reform of Wholesale Power Markets: A Dynamic Partial Equilibrium Analysis of Negative Externalities* PhD thesis, The University of Queensland.