Electricity Spot Price Modelling and Derivatives Pricing

Barry Thornton

5th February 2010

Submitted to the

University of Zürich

and the

Swiss Federal Institute of Technology (ETH), Zürich

for the degree of

Master of Advanced Studies in Finance

with specialisation in Quantitative Finance and Risk Management.

Supervised by

Prof. Dr. Christoph Schwab

Seminar for Applied Mathematics,
ETH Zürich.

This thesis work was carried out during an internship at

EGL Trading AG

13th July 2009 - 12th January 2010.

Acknowledgments

I would like to sincerely thank my supervisor, Prof. Christoph Schwab, for his support and advice throughout my thesis work, as well as for the two courses on Computational Finance taught by he and his colleagues at the Seminar for Applied Mathematics (SAM). Special thanks also go to Oleg Reichmann at SAM for many interesting and helpful discussions, for proof-reading various drafts of this thesis and for always being available to provide sound advice and support. Thanks are additionally due to Dr. Andrea Barth for also proof-reading the script, as well as Sohrab Kehtari for providing some useful references.

This thesis would not have been possible without the sponsorship of Roberto Frassanito and Simone Faragalli at EGL Trading and I am deeply grateful to them for the opportunity they provided me. Thanks are also due to the other members of the Business Management team, particularily Frédéric Dumas and Kai Schönle, who were always willing to help in any way they could.

Most importantly, I would like to thank my wife, Deirdre, for her loving support throughout my studies, and Maebh and Luke for providing many welcome distractions.

Abstract

We introduce an additive three factor model for the Day Ahead baseload price of electricity. The model uses two Lévy-driven Ornstein-Uhlenbeck processes to capture the day-to-day variation in spot prices and a diffusion to model the stochastic mean level. The model is calibrated to historical EEX price data. A detailed risk-neutral calibration is carried out and properties of the market prices of risk described. Monthly futures contracts are priced in closed form and options on the spot price and on futures contracts are valued using the Streamline Diffusion Finite Element Method.

Contents

1 Introduction

Besides its obvious unique social importance, electricity is different to other commodities such as gold and oil in two fundamental respects. Firstly, it is very difficult to store in any great quantity, so much so that it is often referred to as *non-storable*. In some countries, relatively small amounts can be implicitly stored in hydro-reservoirs but, by and large, it is not possible to buy, hold and sell a quantity of electricity to exploit a perceived arbitrage opportunity. Therefore, electricity is not considered *tradable* in the way that other commodities are and care must be taken when applying results from arbitrage theory to electricity markets. Furthermore, non-storability can cause temporary over- and under-supply to increase or decrease prices dramatically for a short period of time, a phenomenon which must be taken into account when modelling prices.

Secondly, electricity must be delivered over a period of time and not at a single instant and is thus called a *flow* commodity. For this reason, *futures* contracts must always specify the time period over which the electricity will be delivered and consumed. This period typically ranges from one hour to a whole year. Consequently, electricity futures are more akin to swap contracts in that they are settled against a reference price over a period of time. Additionally, due to the varying delivery periods of futures contracts, it is a non-trivial task to find the 'forward curve' of market implied future electricity prices - something that is taken for granted in other commodity and fixed-income markets.

In Europe, several markets exist that manage electricity supply on a national or regional level. The focus of this analysis is the german market where products are mainly traded at the European Energy Exchange (EEX), but the results can readily be applied to other liquid markets. In general, two types of products are traded on European exchanges: those with physical delivery in the near-immediate future ($\leq$ 1 day), which we call the spot market, and those which do not require market participants to be capable of supply or consumption and with delivery over a longer period of time further in the future, which we call the futures market.

The principal spot market is known as the *Day Ahead* (DA) market. Each morning, market participants submit their bids for buying or selling electricity in each hour of the following day. The market closes at noon and the prices for each hour are set based on the bids of suppliers and consumers. Thus, the atomic delivery period is a single hour in a day and the *baseload* price is the average of all hours in each day.

The futures market consists mainly of monthly, quarterly and yearly contracts, although daily, weekly, weekend and other products are also traded. This thesis will be mainly concerned with monthly contracts which are delivered over each calendar month and at any given time contracts for the next six months are traded with varying degrees of liquidity. The trading period closes when the delivery period begins and the monthly price is settled on a daily basis against the DA hourly price.

All the above futures contracts are also traded on an Over-The-Counter (OTC) basis along with options and other derivatives. However, monthly OTC

contracts are settled after the end of the delivery period rather than on a daily basis during delivery. Furthermore, OTC prices only exist when an actual trade has been agreed between two parties, whereas exchange prices are published regardless of whether any trades have taken place or not.

It will be useful to also consider contracts that deliver a unit of power at a single specific time, even though these are not explicitly traded. We will refer to these as *forward* contracts.

It is spot price modelling that provides the focus of this analysis. We wish to price monthly futures contracts and vanilla options by specifying the dynamics of the DA price. As all hourly prices are set at auction simultaneously and therefore the baseload price impounds all information that is captured in each of the 24 hourly prices, from a pricing point of view it is sufficient to focus on the daily baseload, rather than on the hourly price (see [32]).

In addition, the calibration of a spot price model to the forward market will involve the discovery of the market price of risk as a function of time and time-to-delivery. The form of the market price of risk is a worthy goal in itself as it tells us much about the evolution of the risk preferences of market participants over time and over the tradable life of a contract. Finally, it is hoped that an effective spot price model will provide a starting point for pricing more complicated contracts such as tolling agreements or spark-spread contracts, which in turn can be used to model the operation of gas-fired power plants.

Note that the markets for natural gas have many similarities with the above mentioned features of electricity markets: gas is a flow commodity with limited storage capacity and a balanced physical spot market exists alongside a financial futures market. The model described in this thesis was developed specifically with gas markets in mind and the results given here for electricity could also be applied to gas.

2 Spot price modelling

2.1 Spot versus forward price modelling

One of the main motivations for the stochastic modelling of electricity markets is the pricing and hedging of futures contracts. The modelling of these products is analagous to the modelling of bond prices and swap and forward rates in the fixed-income world, which generally falls into two categories (see, for example, [10]). On the one hand, spot models focus on the direct modelling of the *non-tradable, exogenously-given* instantaneous short rate of interest. Within this paradigm, traded instruments such as bonds and forward contracts are priced as derivatives of the short rate. On the other hand, with the HJM framework and LIBOR and Swap Market models, the whole term structure of forward rates is modelled simultaneously.

The same dichotomy exists in the modelling of electricity markets. One class of models focuses on specifying and calibrating the dynamics of forward and futures prices of electricity and explicitly uses concepts from interest rate

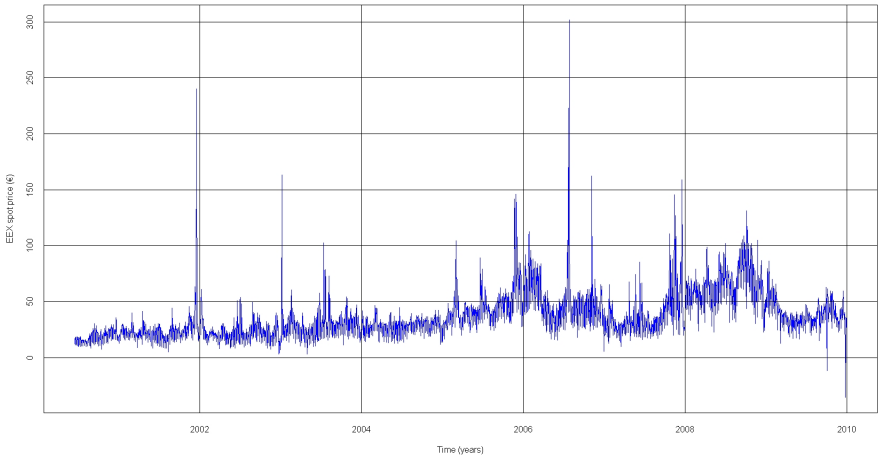

Figure 1: EEX Day Ahead baseload prices 2000 - 2009.

theory (see, for example, [7, Chapters 6 and 8] or [28]). On the other hand, spot price models specifying the dynamics of the *exogenously-given* DA price of electricity, which is considered *non-tradable*, are used to derive the prices of futures and more exotic contracts.

Although the direct modelling of forward and futures prices can better account for the complicated dynamics of the forward curve, they have one main disadvantage: forward price models, in general, use as an initial condition the observed forward price curve. Therefore, if the market contains any mis-pricing, it will not be detected by these models, which will automatically calibrate the mis-priced data. From this point of view spot models are very useful for giving an alternative perspective of the forward market.

2.2 Stylized facts about spot prices

Before discussing price modelling in depth, we summarise some well-known stylized facts concerning the dynamics of the spot price.

- Prices exhibit large 'spikes', where the price can increase by several magnitudes over a single day and return to 'normal' levels just as quickly. Downward spikes also occur, but less frequently and are of smaller magnitude although they can lead to negative prices. Both positive and negative spikes are clearly visible in Figure 1.

- Not only do large price spikes revert to a mean price, but also day-to-day changes revert to a long-term mean level.

- Prices exhibit a regular seasonal and daily variation that is caused by regular variations in demand. For example, demand is usually higher on work-

ing days than at the weekend or on public holidays. Similarily, heating and cooling requirements cause high demand in winter and midsummer, respectively.

It will be seen in the next section how these three observations have greatly influenced the modelling of spot prices to date. However, in this thesis we wish to focus on the following further two aspects of price modelling.

Non-Gaussian Increments: Given the above list of observations, it is tempting to think that once seasonal factors and price spikes have been filtered out, the time series could be well-modelled by a Brownian-driven mean-reverting process. However, as we will see in the following sections, the distribution of the increments of deseasonalised prices without spikes is far from normal.

Stochastic Mean: It is possible to observe in Figure 1 that the long-term mean price, to which day-to-day price variations revert, can vary quite considerably itself. This is even more evident in the top left-hand graph of Figure 6, where the 31-day moving average (MA) of deseasonalised prices is highlighted in red. Here we see that throughout 2008, the MA was consisently above the seasonal level, whereas for most of 2009 the MA persisted below the seasonal level. This stochastic mean price is of critical importance in risk neutral modelling; if it is ignored, the market prices of risk will have to account for much more of the price dynamics and will become highly irregular.

Moreover, we can see in Figure 1 that in the period 2000-2005, the day-to-day price variations reverted to a more or less constant level. In contrast, in the period 2006-2009, we can see an increased volatility of the stochastic mean. This could be due to the gradual evolution of a still immature market and means the price time series is unlikely to be stationary. Therefore, all 10 years of prices should not be employed for model calibration and instead, relatively short periods of time (e.g. 3 or 4 years) should be used. However, this has the disadvantage that statistically calibrated parameters may then be unreliable and, therefore, frequent recalibration is required.

2.3 Review of spot models

In order to fully explain our choice of model, we briefly survey some of the types of spot price models that are common in the literature. Most models focus on taking care of the three well-known properties of price time series listed above and try to separate seasonal factors and mean-reverting spikes from day-to-day price variations. Thus, typical models are of the following form (see [7, Section 3.2.2]).

$$\begin{aligned} S_t &= \Lambda(t) + \sum_{i=1}^{n} X_t^i,\ i = 1, \ldots, n \\ dX_t^i &= a_i(m_i - X_i)dt + \sigma_i(t)dL_t^i. \end{aligned} \tag{2.1}$$

$\Lambda : \mathbb{R}_+ \mapsto \mathbb{R}$ is a deterministic function capturing seasonal and daily variations and $X^i = \{X_t^i : t \geq 0\}, i = 1, \ldots, n$ are known as Ornstein-Uhlenbeck (OU) processes and are described in detail in Section 4. $L^i = \{L_t^i : t \geq 0\}, i = 1, \ldots, n$ are Lévy processes, whose definition and properties are given in Section 3. $S = \{S_t : t \geq 0\}$ can be taken as either the spot price, in which case the model is *additive* or the log spot price, where the model is *geometric*. Obviously, geometric models do not allow negative prices, which is a disadvantage that has become more pronounced recently[1]. Examples of models that fit into the above framework include [4, 9, 13, 32]. Such models typically capture spikes in one factor with a mean-reverting compound Poisson (CP) process and day-to-day stochastic variations in a second factor. Historically, most models have used a Brownian motion driven OU process to model the day-to-day variations, although recently, [9] used a jump-diffusion OU process, while [4] used gamma OU processes for both spikes and day-to-day variations, which again has the disadvantage of precluding negative prices.

When two factors are used, typical rates of mean reversion are about 1 for the spike factor and about .1 for the other factor (see [9, 32]), resulting in a spike half-life of about 1 day and a half-life for other price movements of about 10 days. Therefore, these models have the distinct drawback that the prices of futures contracts derived from the spot model have almost zero volatility if the time to delivery is more than a few weeks. Consequently, the market price of risk, which is a result of the risk neutral calibration of the spot model to the forward market, becomes stochastic, making it almost impossible to price futures contracts with the spot model. Furthermore, when all factors are mean-reverting, the deseasonalised mean-level is constant, which can cause problems during calibration as prices are assumed to revert to an unrealistic level.

A similar class of models can be obtained by allowing one factor to be non-mean-reverting. For example, [2, 29, 35] suggest using a (drifted) Brownian motion to capture the long-term stochastic mean level of spot prices. However, all three models also assume that the mean-reverting day-to-day price variations can be well-modelled by a Brownian motion driven OU process - an assumption which we will show to be unrealistic.

An alternative way of circumventing the issue of the stochastic mean is obtained by letting spot prices revert to the forward price. The forward price curve, $F(t, T), 0 \leq t \leq T$, where T is the start of the delivery period, can be modelled separately and the spot prices derived by setting $F(t, t) = S_t$ (see, for example, [14]). However, forward price dynamics are qualitatively different to spot price dynamics and do not contain spikes or strong mean reversion. Therefore, the resulting spot prices will not have realistic dynamics.

Finally, we mention in passing regime-switching models such as [17, 23]. These models typically use a single OU factor, but the specification of the factor changes to the so-called 'spike regime' during periods of high volatility when spikes can occur. While these models can provide a finer structure for the spike dynamics, they come with the additional cost of having to calibrate the

[1] Strongly negative spot prices were twice recorded on EEX in the last quarter of 2009.

hidden Markov chain that manages the regime switching. In any case, we will see that the intricacies of price spikes are not very important when it comes to pricing instruments such as monthly futures, as only the overall expected value of the spikes is included in the futures price.

2.4 Three-factor model

We aim to specify a model that incorporates the best aspects of the models descibed above and also specifically addresses the issues of Non-Gaussian Increments and the Stochastic Mean. Therefore, we use a two-factor model of the form (2.1), where one factor is driven by a CP process and specifies price spikes, while the other factor is driven by a variance gamma (VG) process (see Section 3.2) and captures day-to-day price variations. We will see in Section 6.5 that the non-Gaussian increments of prices are well-modelled by a VG process. Furthermore, we add to the two factor model a non-mean-reverting factor which is intended to capture the long-term stochastic mean level of the prices. Thus, we have the following specification for our spot price model.

Definition 2.1 (Three factor model)**.** *The spot price of electricity can be modelled by the stochastic process, $S = \{S_t : t \geq 0\}$, which is defined as*

$$\begin{aligned} S_t &:= \Lambda(t) + X_t^1 + X_t^2 + X_t^3 \\ dX_t^1 &= \sigma_1 dW_t \\ dX_t^2 &= a_2(m - X_t^2)dt + dL_t \\ dX_t^3 &= -a_3 X_t^3 dt + dP_t \end{aligned} \tag{2.2}$$

where $\Lambda : \mathbb{R}_+ \mapsto \mathbb{R}$, $W = \{W_t : t \geq 0\}$ is a Brownian Motion, $L = \{L_t : t \geq 0\}$ and $P = \{P_t : t \geq 0\}$ are VG and CP processes respectively, σ_1, a_2 and a_3 are positive constants and $m \in \mathbb{R}$.

Note that Lévy processes other than VG, such as the normal inverse Gaussian (NIG) process, could also be used to the same effect. We choose VG as the form of its Lévy measure is relatively straightforward (unlike the NIG process) and therefore amenable to finite element pricing[2].

We remark that there is probably very little difference in the economic factors driving the processes X^2 and X^3. Both are a result of the day-to-day variations in supply and demand and it could be argued that the realisations of X^3 are simply extreme values of X^2. If we assume $a_2 = a_3$, one stochastic factor could be used in place of X^2 and X^3, but then it becomes almost impossible to calibrate a single Lévy driven OU process to all day-to-day variations (including spikes).

On the other hand, we assert that the economic factors driving X^1 are not caused by day-to-day changes in supply and demand, but rather longer term

[2]FEM for Lévy processes requires the calculation of anti-derivatives of the Lévy density. This task is significantly more complicated for the NIG process as its density contains a modified Bessel function of the second kind and order one.

influences such as the price of commodities used as hedging instruments and fuels as well as changes in GDP or the total number of functioning power stations.

Before calibrating this model and using it for pricing, we first recall some basic properties of Lévy processes and then compile some mathematical results that will be needed later.

3 Lévy processes

3.1 Definitions, properties and stochastic calculus

Unless otherwise stated, all stochastic processes and random variables are $\mathbb{R}^d$-valued and defined on a filtered probability space $(\Omega, \mathcal{F}_\infty, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, where the filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfies the usual conditions and expectations are taken under the measure $\mathbb{P}$.

Definition 3.1. *An adapted, càdlàg stochastic process $X = \{X_t : t \geq 0\}$, with $X_0 = 0$, is called a* Lévy process *if it satisfies the following conditions:*

- *X has independent increments: for all $0 \leq s < t < \infty$, $X_t - X_s$ is independent of $\mathcal{F}_s$.*
- *X has stationary increments: for all $0 \leq s < t < \infty$, $X_t - X_s \sim X_{t-s}$.*
- *X is stochastically continuous: given $\epsilon > 0$, for all $s, t \geq 0$,*

$$\lim_{t\to s} \mathbb{P}(|X_t - X_s| > \epsilon) = 0.$$

The stochastic continuity property implies that a Lévy process can contain jumps and we associate with $X = \{X(t,\omega) : t \in [0,T], \omega \in \Omega\}$, where T is some finite time horizon, a random measure J_X on $[0,T] \times \mathbb{R}^d$ such that $\forall \omega \in \Omega$

$$J_X(\omega, \cdot\,, \cdot\,) = \sum_{t\in[0,T]}^{\Delta X(t,\omega)\neq 0} \mathbb{1}_{(t,\Delta X(t,\omega))}(\cdot\,, \cdot\,)$$

where $\Delta X(t,\omega) = X(t,\omega) - X(t-,\omega)$. Thus, for any $B \in \mathcal{B}(\mathbb{R}^d)$, $J_X([s,t], B)$ counts the number of jumps with magnitude in B occurring between times s and t. Furthermore, we define a measure, ν, on $\mathbb{R}^d$, called the *Lévy measure*, which gives the expected number of jumps per unit time with magnitude in B as

$$\nu(B) = \mathbb{E}[\#\{t \in [0,1] : \Delta X_t \neq 0, \Delta X_t \in B\}].$$

If a Lévy measure is absolutely continuous with respect to the Lebesgue measure, it admits a density and we write $\nu(dz) = k(z)dz$, where $k(z)$ is the Lévy density.

By examining the defining properties of Lévy processes, we can easily see that a Brownian motion, which is continuous and has independent and stationary increments, is a Lévy process with a Lévy measure that is 0 everywhere. Moreover, compound Poisson (CP) processes, whose paths are not continuous, are also Lévy processes and have a Lévy measure of the form $\nu(dz) = \lambda f_Z(z)dz$ where $\lambda > 0$ is the jump intensity and f_Z is the probability density of the jump sizes Z, a $\mathbb{P}$-integrable random variable. In this case, the Lévy measure is integrable everywhere, but the definition of Lévy processes allows much weaker integrability conditions and, consequently, a much broader class of processes.

Theorem 3.2. *Every Lévy measure satisfies*

$$\int_{\mathbb{R}^d} 1 \wedge z^2 \nu(dz) < \infty$$

Proof. See [34, Section 8]. □

The condition $\int_{|z|>1} \nu(dz) < \infty$ of the above theorem means that if we only consider the jumps of magnitude greater than 1 then the process is a CP process and thus has finite intensity. For jumps less than or equal to 1, if $\int_{|z|\leq 1} |z|\nu(dz) < \infty$ is satisfied, the process can be thought of as a CP process where the intensity of occurrence of small jumps is infinite, although in this case, the sum of the jump sizes is finite. Hence, the process is said to be of *infinite activity* and *finite variation*. If only $\int_{|z|\leq 1} z^2\nu(dz) < \infty$ holds, the sum of jumps less than 1 does not converge and the process must be compensated with the expected value of the small jumps. In this case the process is said to be of *infinite activity* and *infinite variation*. This leads us to the general form of all Lévy processes.

Theorem 3.3 (Lévy-Itô decomposition)**.** *Every Lévy process X can be written as*

$$X_t = \gamma t + W_t^A + \int_0^t \int_{|z|>1} z J_X(ds, dz) + \underbrace{\lim_{\epsilon \downarrow 0} \int_0^t \int_{\epsilon \leq |z| \leq 1} z \tilde{J}_X(ds, dz)}_{(a)} \tag{3.1}$$

where $\gamma \in \mathbb{R}^d$, $W^A = \{W_t^A : t \geq 0\}$ is a d-dimensional Brownian motion with covariance matrix A and $\tilde{J}_X(ds, dz) = J_X(ds, dz) - \nu(dz)ds$ is the compensated jump measure.

Proof. See [34, Section 20]. □

Remark 3.4. *Due to the compensation of small jumps,* (3.1)(a) *is a square integrable martingale.*

Remark 3.5. *If $\int_{|z|>1} |z|\nu(dz) < \infty$, $\int_0^t \int_{|z|>1} |z|\nu(dz)ds$ can be added to and subtracted from* (3.1) *to obtain*

$$X_t = \left(\gamma + \int_{|z|>1} |z|\nu(dz)\right) t + \underbrace{W_t^A + \lim_{\epsilon \downarrow 0} \int_0^t \int_{\epsilon \leq |z| \leq 1} z \tilde{J}_X(ds, dz)}_{(a)}. \tag{3.2}$$

As the terms in (3.2)(a) *are martingales, this form of the Lévy-Itô decomposition can be useful when computing expectations.*

Given the form of the Lévy-Itô decomposition, every Lévy process can be defined by its *characteristic triplet* (A, ν, γ). These parameters can also be used to define the characteristic function, φ, of a Lévy process, which can be shown to have the following general form.

Theorem 3.6 (Lévy-Khinchin representation)**.** *The characteristic function, φ_X, of a Lévy process, X, can be written as*

$$\varphi_X(\xi) = \mathbb{E}[e^{i\langle \xi, X_t\rangle}] = e^{-t\psi(\xi)}, \xi \in \mathbb{R}^d,$$

where ψ is called the characteristic exponent and has the representation

$$\psi(\xi) = -i\langle \gamma, \xi\rangle + \frac{1}{2}\langle \xi, A\xi\rangle + \int_{\mathbb{R}^d} (\underbrace{1 - e^{i\langle \xi, z\rangle}}_{(a)} + i\langle \xi, z\rangle \mathbb{1}_{|z|\leq 1})\nu(dz). \tag{3.3}$$

Proof. See [34, Theorem 8.1] □

Remark 3.7. *By rewriting* (3.3)(a) *as* $(1 - e^{i\langle \xi, z\rangle})(\mathbb{1}_{|z|\leq 1} + \mathbb{1}_{|z|>1})$ *and setting $\gamma, A = \mathbf{0}$, we recover the characteristic exponents of a CP process with jumps greater than one and a compensated CP process with jumps less than 1, corresponding to the Lévy-Itô decomposition.*

Remark 3.8 (Truncation function)**.** *The separation of jumps into those greater than and less than or equal to 1 in* (3.1) *and* (3.3) *is achieved with the truncation function $\mathbb{1}_{|z|\leq 1}$. However, this choice is arbitrary and other functions with the same asymptotics could be employed. If the process is of finite variation,* 0 *can be used as the truncation function and if $\int_{|z|>1} |z|\nu(dz) < \infty$,* 1 *can be used. Note that when the truncation function is changed, the value of the 'drift', γ, also changes.*

If the stationary increments property of Lévy processes is dropped, the characteristic triplet becomes time-dependent. In this case, the Lévy measure must satisfy for $0 \leq s \leq t$

$$\int_s^t \int_{\mathbb{R}^d} 1 \wedge z^2 \nu(dz, du) \leq \infty.$$

Such processes are called time-inhomogeneous or *additive* Lévy processes and are dealt with in detail in [34, Sections 9 and 11]. In this thesis, we only consider time-homogeneous processes, although time-inhomogeneous processes can arise through an Esscher transform of a time-homogeneous process (see Section 5).

The jumps of a Lévy process also need to be taken into account in the Itô Formula for functions of a Lévy process.

Theorem 3.9 (Itô's Formula for Lévy processes)**.** *Let $X = (X^1, \dots, X^d)^\top$ be a Lévy process with characteristic triplet (A, ν, γ). Then, for $f \in C^2(\mathbb{R}^d)$,*

$$\begin{aligned} f(X_t) =& f(\mathbf{0}) + \int_0^t \sum_{i=1}^d \partial_i f(X_{s-}) dX_s^i + \int_0^t \sum_{i,j=1}^d \frac{1}{2} A_{ij} \partial_{ij} f(X_{s-}) ds \\ &+ \sum_{0 \le s \le t} \left[f(X_{s-} + \Delta X_s) - f(X_{s-}) - \sum_{i=1}^d \Delta X_s^i \partial_i f(X_{s-}) \right]. \end{aligned} \tag{3.4}$$

Proof. See [1, Theorem 4.4.10]. □

Heuristically, the two integral terms in the above equation are due to the continuous, Brownian component of the Lévy process and $f(X_{s-} + \Delta X_s) - f(X_{s-})$ counts the contribution to the change in f from the jumps at each instant. As the jump component may be of infinite variation the sum of these jump contributions may not converge and therefore $\sum_i \Delta X_s^i \partial_i f$ is subtracted at each instant. Hence, if the Lévy process is of finite variation, the following form of Itô's Formula holds for functions that are also explicitly dependent on time.

Corollary 3.10 (Itô's Formula for Lévy processes of finite variation)**.** *Let $X = (X^1, \dots, X^d)^\top$ be a Lévy process with characteristic triplet $(0, \nu, \gamma)$ such that $\int_{|z| \le 1} |z| \nu(dz) < \infty$. Then, for $f \in C^{1,1}([0,T] \times \mathbb{R}^d)$,*

$$\begin{aligned} f(t, X_t) =& f(0, \mathbf{0}) + \int_0^t \partial_s f(s, X_s) ds + \int_0^t \sum_{i=1}^d \partial_i f(X_{s-}) dX_s^i \\ &+ \sum_{0 \le s \le t} \left[f(X_{s-} + \Delta X_s) - f(X_{s-}) \right]. \end{aligned} \tag{3.5}$$

Proof. See [1, Lemma 4.4.6]. □

3.2 Variance gamma process

The use of the variance gamma (VG) process was originally described in [31] and further developed in [30] (see also [16, Section 4.4] for a concise summary). It is a Lévy process and can be obtained by subordination of Brownian motion, but we begin by considering the VG distribution.

A normal variance-mean mixture is a random variable, X, obtained by mixing a standard Gaussian random variable, Z, with an independent, non-negative, absolutely continuous random variable, V, as follows:

$$X = \alpha + \beta G + \sigma \sqrt{G} Z$$

where $\alpha, \beta \in \mathbb{R}$ and $\sigma > 0$. Depending on the distribution of G, the density of the resulting random variable can have heavier tails and more kurtosis than a Gaussian density and can also be asymmetric, properties that can be useful in

modelling financial returns. If $\alpha, \beta = 0$ and $G \sim \Gamma(\frac{1}{\kappa}, \frac{1}{\kappa})$, $\kappa > 0$ the resulting variance-mean mixture is a VG random variable (see [36] for its distribution and properties).

Similarily, leptokurtic, asymmetric, (semi-)heavy tailed distributions for increments of stochastic processes can be derived from the Gaussian Brownian motion by means of *subordination*. Consider a drifted Brownian motion with volatility σ and drift θ:

$$b(t, \theta, \sigma) = \theta t + \sigma W_t.$$

Using an increasing process or subordinator, $G = \{G_t : t \geq 0\}$, we can change the time of the drifted Brownian motion to obtain a new process $b(G_t, \theta, \sigma)$. This has an intuitive economic interpretation that when price volatility fluctuates, trading or business time can be thought of as randomly speeding up or slowing down. For this reason the subordinator is often referred to as a stochastic clock or business time and the resulting subordinated Brownian motion could be thought of as a standard Brownian motion with stochastic volatility.

When G is a gamma process with $G_1 \sim \Gamma(\frac{1}{\kappa}, \frac{1}{\kappa})$, $b(G_t)$ is known as a VG process. As G is an increasing Lévy process, b inherits the independent and identically distributed properties of the increments of the Brownian motion, W. Furthermore, as G is stochastically continuous, b is also stochastically continuous. Hence, is a Lévy process. Moreover, as gamma processes are pure jump processes, so are VG processes.

A VG process can also be expressed as the difference between two independent gamma processes. Specifically, any VG process, $X = \{X_t : t \geq 0\}$, can be written as

$$X_t = G_t^1 - G_t^2,$$

where G^1, G^2 are independent gamma processes. In electricity markets, G^1 can be interpreted as the demand for power and G^2 the supply, which provides an intuitive rationale for using a VG process to model day-to-day price variations in (2.2).

The Lévy measure of a VG process can be derived from that of a gamma process and is given by

$$\nu(dz) = \frac{1}{\kappa|z|} e^{Az - B|z|} dz, \tag{3.6}$$

where $A = \frac{\theta}{\sigma^2}$, $B = \frac{\sqrt{\theta^2 + \frac{2\sigma^2}{\kappa}}}{\sigma^2}$ and $B > |A|$ must be satisfied to ensure that the exponential part of the measure is decreasing. The volatility of the process is specified by σ, while the kurtosis of the distribution of the increments is given by ν. θ specifies the skew and for this reason, if $\theta = 0$, the process is known as the symmetric VG (see [31]).

Given (3.6), we can clearly see that $\int_{|z| \leq 1} z\nu(dz) < \infty$ and $\int_{\mathbb{R}} \nu(dz) = \infty$ and therefore the process is of finite variation and infinite activity. Therefore, by Remark 3.8 we can use the 0 truncation function in the Lévy-Kintchine

representation (3.3) to write

$$\psi(\xi) = -i\gamma\xi + \int_{\mathbb{R}} (1 - e^{i\xi z})\nu(dz). \tag{3.7}$$

By conditioning on the gamma time change, the first four moments of the VG process are calculated in [30] and the expectation of the process at time t is shown to be θt. Using this result and the well-known connection between characteristic functions and moments, we write $i\psi'(0)t = \theta t$. Computing the derivative of the characteristic exponent at 0 we find

$$\begin{aligned}
\psi'(0) &= -i\gamma + \int_{\mathbb{R}} -ize^{i\xi z}\nu(dz)\Big|_{\xi=0} \\
&= -i\left(\gamma + \int_{\mathbb{R}} z\nu(dz)\right) \\
&= -i(\gamma + \theta)
\end{aligned}$$

where the last line was obtained by explicit calculation of the integral using the Lévy density. Hence, with the 0 truncation function, $\gamma = 0$ and the characteristic triplet of the VG process can be written as $(0, \nu, 0)$.

Similarily, we can use the 0 truncation function in the Lévy-Itô decomposition (3.1) to write

$$X_t = \int_0^t \int_{\mathbb{R}} zJ_X(ds, dz)$$

where $\gamma, A = 0$. By adding and subtracting $\int_0^t \int_{\mathbb{R}} z\nu(dz)ds = \theta t$ we can write

$$X_t = \theta t + \underbrace{\int_0^t \int_{\mathbb{R}} z\tilde{J}_X(ds, dz)}_{(a)} \tag{3.8}$$

where (a) above is a martingale.

Note that by rewriting (3.6) as

$$\nu(dz) = \frac{1}{\kappa|z|^{1+Y}} e^{Az - B|z|} dz, \; Y < 2,$$

we see that a VG process is a special case of a CGMY process with $C = \kappa^{-1}$, $G = B + A$, $M = B - A$ and $Y = 0$. For $Y \in [1, 2)$, $\int_{|z|\leq 1} z\nu(dz)$ is infinite and the process is of infinite variation (see [12] for more details).

The increments of a VG process over $(t, t+1)$ are distributed as independent VG random variables and we can therefore calibrate a VG process to a time series of asset prices using maximum likelihood (see [36, Section 3] and [6] for a detailed analogous example using the NIG process). Note that when calibrating the VG distribution a non-zero location parameter, $C \in \mathbb{R}$, can be produced. This location contributes to the drift so that it becomes $\theta + C$ (see [36, equation (18)]).

4 Ornstein-Uhlenbeck process

4.1 Definition and properties

Ornstein-Uhlenbeck (OU) processes are used to model mean-reverting behaviour in asset price dynamics, where prices fluctuate around a (possibly moving) long-term mean or fundamental level. Such processes are widely used in interest rate theory (e.g. the Vasicek model) and in commodity markets and are usually Brownian motion driven. In electricity markets, however, the stylized facts of spot prices naturally lead to the use of non-Gaussian OU processes, which may be defined as follows:

Definition 4.1. *A càdlàg process $X = \{X_t : 0 \leq s < t\}$ is called a Lévy-driven OU process if it is the unique solution of the stochastic differential equation*

$$dX_t = a(m - X_t)dt + \sigma dL_t,\ X_s = x \in \mathbb{R} \tag{4.1}$$

where L is a Lévy process and $m \in \mathbb{R}$ and $a, \sigma \geq 0$ are constant parameters.

a is known as the speed of mean reversion and m is called the mean and for simplicity, we set $\sigma = 1$. The above definition only makes sense if a unique solution to the SDE exists and this is given by the next proposition.

Proposition 4.2. *Given $\mathcal{F}_s$ for $0 \leq s < t$, the unique solution X_t of (4.1) exists and is given by*

$$X_t = \underbrace{xe^{a(t-s)} + m(1 - e^{-a(t-s)})}_{(a)} + \underbrace{\int_s^t e^{-a(t-u)} dL_u}_{(b)} \tag{4.2}$$

Proof. See [7, Proposition 3.1] □

We now use the above solution to find the conditional expectation.

Proposition 4.3. *If the Lévy measure of X satisfies $\int_{|z|>1} |z|\nu(dz) < \infty$, then the conditional expectation of X_t given $\mathcal{F}_s$ for $0 \leq s < t$ is*

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s e^{a(t-s)} + (1 - e^{-a(t-s)}) \left[m + \frac{1}{a} \underbrace{\left(\gamma + \int_{|z|>1} |z|\nu(dz) \right)}_{(a)} \right] \tag{4.3}$$

Proof. Given $\mathcal{F}_s$, the terms in (4.2)(a) are deterministic and hence $\mathcal{F}_s$-measurable. We rewrite (4.2)(b) using the form of the Lévy-Itô decomposition in (3.2) to get

$$\int_s^t e^{-a(t-u)} dL_u = \underbrace{\left(\gamma + \int_{|z|>1} |z|\nu(dz)\right) \int_s^t e^{-a(t-u)} du}_{(a)}$$
$$+ \underbrace{\int_s^t \sigma e^{-a(t-u)} dW_u + \lim_{\epsilon \downarrow 0} \int_s^t \int_{\epsilon \leq |z| \leq 1} z e^{-a(t-u)} \tilde{J}_X(du, dz)}_{(b)} \tag{4.4}$$

Both terms in (4.4)(b) are martingales and, given $\mathcal{F}_s$, have expectation equal to 0. The result follows by combining (4.4)(a) with (4.2)(a). □

Remark 4.4. *Given the* (3.2) *and* (3.8)*, for a VG process* (4.3)(a) *is equal to* θ*, the drift of the process at* $t = 1$*. A similar result holds for all processes of finite variation where* $\int_{|z|>1} |z|\nu(dz) < \infty$*.*

An OU process is the continuous-time analogue of a discrete-time autoregressive AR(1) process, Y, which for $t = 1, 2, 3, \ldots$ can be defined recursively as

$$Y_t := c + \mu Y_{t-1} + \epsilon_t$$

where $c \in \mathbb{R}$ and $\epsilon = \{\epsilon_t : t = 0, 1, 2 \ldots\}$ are the realisations of a discrete set of random variables indexed by t and Y_0 is some initial value. By approximating the OU dynamics for $t = 1, 2, 3, \ldots$ with

$$X_{t+1} \approx X_t e^{-a} + m(1 - e^{-a}) + \Delta L_t$$

we can conveniently calibrate an OU model as follows (see also [7, Section 5.1.1]).

Given a set of price data $\{S_t : t = 0, 1, 2, \ldots, N\}$, $\forall t = 1, \ldots, N$ plot S_t against S_{t-1} and fit a straight line to the data with linear regression. The slope, μ, and the intercept c of the fitted line can be used to determine the parameters of an OU process by solving for a and m in the pair of equations

$$\mu = e^{-a}, \; c = m(1 - e^{-a}).$$

The residuals from the regression can then be used to calibrate the increments of the Lévy process L.

5 Girsanov and Esscher Transforms

5.1 The Girsanov Transform

The Girsanov theorem provides a receipe for transforming a $\mathbb{P}$-Brownian motion to a $\mathbb{Q}$-Brownian motion, where $\mathbb{Q}$ is second probability measure equivalent to

$\mathbb{P}$. It is used in finance to change the dynamics of a model from calibration under the historical probability measure to the risk-neutral specification. We being by recalling the statement of the theorem in one dimension.

Theorem 5.1 (Girsanov)**.** *Let $W = \{W_t : 0 \leq t \leq T\}$ be a $\mathbb{P}$-Brownian motion and $\lambda = \{\lambda_t : 0 \leq t \leq T\}$ an adapted process that satisfies the* Novikov condition

$$\mathbb{E}\left[e^{\frac{1}{2}\int_0^T \lambda_t^2 dt}\right] < \infty.$$

For $0 \leq t \leq T$, define the density process

$$Z(t) = \exp\left(\int_0^t \lambda_s dW_s - \frac{1}{2}\int_0^t \lambda_s^2 ds\right) \tag{5.1}$$

Then, a new probability measure, $\mathbb{Q} \sim \mathbb{P}$, can be defined such that

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = Z(t)$$

and

$$W_t^{\mathbb{Q}} = W_t - \int_0^t \lambda(s)ds \tag{5.2}$$

is a $\mathbb{Q}$-Brownian motion

Proof. See [37, Theorem 5.2.3]. □

Remark 5.2. *The Novikov condition ensures that Z is a positive martingale and, as it has expectation equal to 1, it is a Radon-Nikodym derivative.*

Remark 5.3. *Usually, λ is chosen as a uniformly bounded, continuous function $\lambda : [0, T] \mapsto \mathbb{R}$, in which case the Novikov condition trivially holds.*

In financial applications, λ is known as the market price of risk and induces a risk premium in the price of derivatives.

5.2 The Esscher Transform

Whereas the Girsanov transform applies to Brownian motion, the Esscher transform is a generalisation that applies to all Lévy processes. In this case, the Radon-Nikodym derivative and the specification of the $\mathbb{Q}$-dynamics of the Lévy process have more general forms, but the basic mechanics of the transform remains the same.

Theorem 5.4 (Esscher)**.** *Let $L = \{L_t : 0 \leq t \leq T\}$ be a Lévy process with characteristic triplet (σ^2, ν, γ) and characteristic exponent ψ and let $\lambda : [0, T] \mapsto \mathbb{R}$ be a uniformly bounded, continuous function. For $0 \leq t \leq T$, define the martingale*

$$Z(t) = \exp\left(\int_0^t \lambda(s)dL_s - \int_0^t \psi(i\lambda(s))ds\right). \tag{5.3}$$

Then, a new probability measure, $\mathbb{Q} \sim \mathbb{P}$, can be defined such that

$$\left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = Z(t) \tag{5.4}$$

and L is a time-inhomogeneous Lévy process with characteristic triplet $(\sigma^2, \nu^\lambda, \gamma^\lambda)$ where

$$\nu^\lambda(dz, dt) = e^{\lambda(t)z}\nu(dz)dt$$

and

$$\gamma^\lambda = \gamma t + \int_0^t \int_{\mathbb{R}} (e^{\lambda(s)z} - 1)z\mathbb{1}_{|z|\leq 1} d\nu ds$$

Proof. See [7, Proposition 4.4]. □

Remark 5.5. *When L is a Brownian motion,* (5.3) *simplifies to* (5.1) *and the characteristic triplet reduces to* $(1, 0, \int_0^t \lambda(s)ds)$*, corresponding to the transform* (5.2)*.*

Remark 5.6. *It is because ν^λ depends implicitly on t, thus causing the increments of L to be time-inhomogeneous, that L becomes a time-inhomogeneous Lévy process after the Esscher transform.*

In a model with n independent factors, the measure change should be carried out with a vector of uniformly bounded, continuous functions

$$\lambda(t) = (\lambda_1(t), \ldots, \lambda_n(t)), t \in [0, T]$$

and the Radon-Nikodym derivative should be composed of a product of exponential factors of the form (5.3)

$$Z(t) = \prod_{i=1}^{n} \exp\left(\int_0^t \lambda_i(s)dL_s - \frac{1}{2}\int_0^t \psi(\lambda_i(s))ds\right).$$

5.3 Esscher transform for Variance Gamma

Applying the Esscher transform to a VG process with characteristic triplet $(0, \nu, 0)$, we get the time-inhomogeneous process $(0, \nu^\lambda, 0)$, where

$$\nu^\lambda(dz, dt) = \frac{1}{\kappa|z|} e^{(A+\lambda(t))z - B|x|} dzdt$$

and $-(B + A) < \lambda < B - A$ must hold to ensure that the $\nu^\lambda(dz, dt) \to 0$ as $z \to \pm\infty$.

Proposition 5.7. *The expectation of a VG process $X = \{X_t : t \geq 0\}$ under a measure $\mathbb{Q}$ defined by* (5.4) *is*

$$\mathbb{E}^\lambda(X_t) = \int_0^t \frac{2(A + \lambda(s))}{\kappa(B^2 - (A + \lambda(s))^2)} ds. \tag{5.5}$$

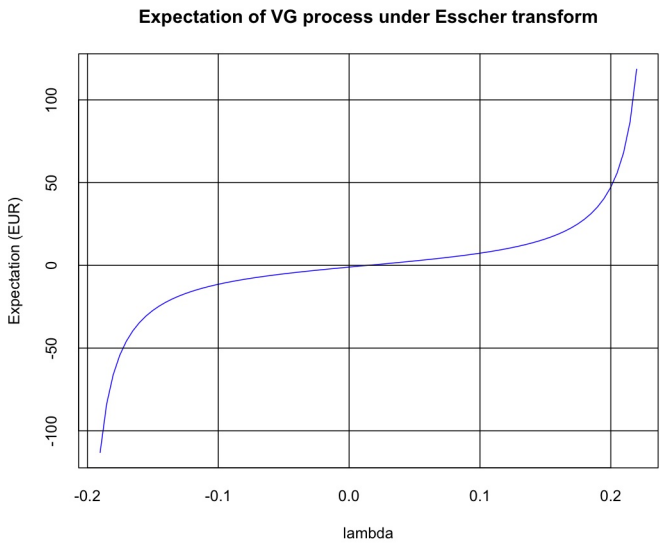

Figure 2: Expectation of VG process at $t = 1$, with $A = -.015$, $B = .22$ and $\nu = .57$.

Proof. Letting $G(s) := B + (A + \lambda(s))$ and $M(s) := B - (A + \lambda(s))$, we find the expectation via the characteristic function:

$$\begin{aligned} i\psi'(0,t) &= i \int_0^t \int_{\mathbb{R}} -ize^{i\xi z} \nu^{\lambda}(dz, ds) \Big|_{\xi=0} \\ &= \int_0^t \frac{1}{\kappa} \left(-\int_{-\infty}^0 e^{G(s)z} dz + \int_0^{\infty} e^{-M(s)z} dz \right) ds \\ &= \int_0^t \frac{1}{\kappa} \left(-\frac{1}{G(s)} + \frac{1}{M(s)} \right) ds \end{aligned}$$

and the result follows upon addition of the terms inside the brackets. □

Remark 5.8. *If $\lambda = 0$, (5.5) reduces to $\mathbb{E}^{\lambda}(X_t)\big|_{\lambda=0} = \theta t$, as before.*

As with the Girsanov transform of a Brownian motion, the Esscher transform of a VG process results in a change in the 'drift' of the process. However, assuming λ to be independent of t, we can see that the expectation of the drifted Brownian motion is linear in λ, whereas the expectation of the transformed VG process is non-linear. This is illustrated in Figure 2, where the expectation of the VG process is plotted for all possible values of λ. We can see that the expectation goes to $\pm\infty$ as $\lambda \to -(B + A), (B - A)$, which implies that all possible expected values can be achieved by means of an Esscher transform. However, there is no economic reason why small changes in the market price of risk near $-(B+A)$ or $(B-A)$ should produce very large changes in the expected

value of an asset modelled by a VG process. For this reason, it is desirable that λ only varies in the near-linear region of the above graph.

6 Calibration to historical spot prices

6.1 Preliminaries

We now describe how the three factor model (2.2) can be calibrated to historical daily baseload prices. We use spot prices from 2007-2009 as this period coincides approximately with the delivery period of the monthly futures prices that will be used for the risk neutral calibration in Section 7.

Before proceeding with the calibration, we calculate, using the results of Section 4, the conditional expectation of the three factor spot price model.

Proposition 6.1. *The conditional expectation of $S(T)$, as defined in* (2.2)*, at time t is given by*

$$\begin{aligned}\mathbb{E}\left[S(T)|\mathcal{F}_t\right] =& \Lambda(T) + \sigma_1 W_t \\ &+ X_t^2 e^{-a_2(T-t)} + (1 - e^{-a_2(T-t)})(m + \frac{\theta}{a_2}) \\ &+ X_t^3 e^{-a_3(T-t)} + (1 - e^{-a_3(T-t)})\frac{\mu_P}{a_3}\end{aligned} \tag{6.1}$$

where μ_P is the expected value of the CP process at $t = 1$.

Proof. The result follows from the conditional expectation of OU processes given in Proposition 4.3 and Remark 4.4. □

The above proposition highlights the fact that, without the non-mean-reverting factor X^1, the long-term dynamics of the conditional expectation of S is almost deterministic due to the exponential decay of the coefficients of X^2 and X^3.

6.2 Seasonality

The first task that must be carried out in calibrating spot prices is filtering out the seasonality of the time series. We do this by fitting the following function to the price data using linear regression.

$$\begin{aligned}\Lambda(t) =& \underbrace{a_0 + a_1 t}_{(a)} + \underbrace{\sum_{n\in N}\left(b_n \sin\left(\frac{2\pi n t}{365.25}\right) + c_n \cos\left(\frac{2\pi n t}{365.25}\right)\right)}_{(b)} \\ &+ \underbrace{\sum_{i=1}^{7} \mathbb{1}_i(d(t))d_i + \mathbb{1}_h(t)e_h + \mathbb{1}_p(t)e_p + \mathbb{1}_c(t)e_c}_{(c)}\end{aligned} \tag{6.2}$$

where $a_0, a_1, b_n, c_n \in \mathbb{R}, n \in \mathbb{N}$, are constants to be determined by the regression and

- $N = \{1, 2, 4, 12\}$, corresponding to annual, semi-annual, quarterly and monthly harmonics.
- $d(t)$ assigns a number to each day (e.g. Monday = 1, Tuesday = 2 etc.) and $d_i, i = 1, \ldots, 7$ is a constant for each day of the week.
- h, p and c are sets containing public holidays, partial public holidays and days during the Christmas period, respectively, and e_p, e_h and e_c are corresponding constants.

(6.2)(a) captures the overall drift of the spot price (due to inflation, for example), the truncated Fourier series (6.2)(b) specifies the seasonal and monthly variations and (6.2)(c) encapsulate variations on particular days. The results of the fit are displayed in Figure 6.2. A clear overall drift (blue line) and regular monthly and seasonal variations (black curve) are visible. Variations due to holidays, Christmas and the days of the week are also prominent (green oscillations).

As only three years of data was used for the above regression, the resulting parameters may be unreliable and should not be applied to another time period. However, the goal of fitting Λ to the data is not to find *the* seasonality function that will hold for all time periods, but rather to remove the seasonal trend from the chosen period of our price time series.

In order to ensure that spikes do not influence the seasonal function, the regression was carried out only after initially removing spikes from the seasonal price series (see next section). However, removing spikes from seasonal prices can be imprecise due to the seasonal variations and not all spikes can be removed. Therefore, after the price time series has been deseasonalised, the spikes removed from the seasonal prices must be added back to the deseasonalised prices, so that all spikes can filtered out together.

6.3 Spikes

The removal of spikes from deseasonalised prices is a non-trivial task that has attracted a lot of attention in the literature. Kalman filtering is often suggested but as it assumes that prices are normally distributed, it will not identify spikes as expected (see [7, Section 5.3.1]). Particle Filters [20] and Markov Chain Monte Carlo methods [24] are also suggested, although such implementations are beyond the scope of this thesis. [32] suggests a non-parametric approach where for a fixed speed of mean reversion a, N spikes of the form $s_i \mathbb{1}_{t \geq \tau_i} e^{-a(t-\tau_i)}$, $i = 1, \ldots, N$ where $s_i \in \mathbb{R}$ are the spike magnitudes and $\tau_i \in \mathbb{R}$ are the starting times of the spikes, are fitted one-by-one to the price series using Maximum Likelihood. We implemented this algorithm and the results are displayed in Figure 4.

As the speed of mean reversion can only be chosen in advance, it is necessarily an average value - individual spikes will have faster or slower decay. When a spike decays more quickly than a, subtraction of the fitted spike from the price series will induce an artifical spike in the opposite direction. Therefore, the

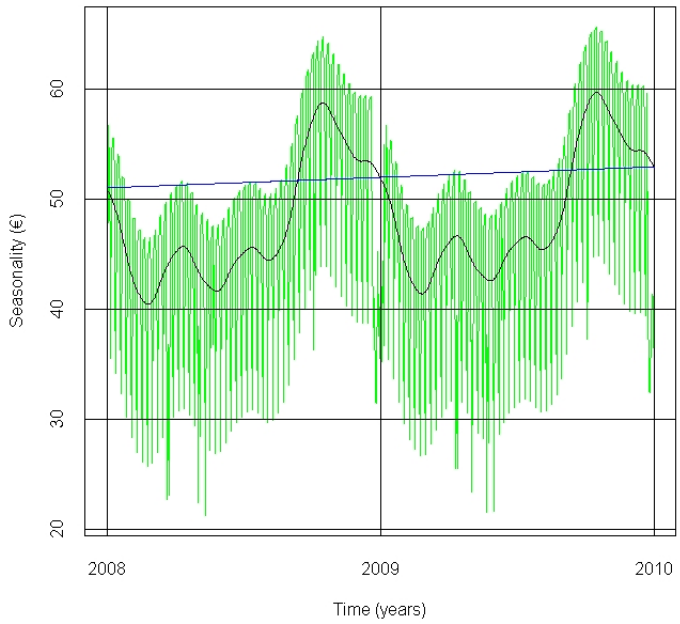

Figure 3: Seasonality function Λ for 2008 and 2009.

removal of some spikes will actually increase the standard deviation of the price signal, as is evidenced in the lower right-hand graph in Figure 4. If a spike decays more slowly than a, the remainder could be counted as second spike, thus overestimating the total number of spikes.

Instead, we use a simpler algorithm akin to the Peaks-Over-Threshold method of Extreme Value Theory (EVT), as outlined in [15]. A threshold is chosen as a certain number of standard deviations and all prices that have an absolute value greater than the threshold are removed from the price time series and replaced with the mean of the two nearest prices. The standard deviation is recalculated and the process is repeated until no more spikes are identified for the given threshold. The resulting spike time series is shown in the lower right-hand graph of Figure 6. The threshold is generally between 2 and 3 standard deviations and we choose the value that gives the best fit not only to the spike distributions, but also to the remaining X^1 and X^2 factors. This corresponds to 2.6 standard deviations.

This algorithm gives the extreme price data but does not tell us how many spikes are present i.e. two extreme prices may correspond to the same spike. To determine the number of spikes, the daily increment in the spike time series must be examined. If the increment is greater than the threshold, then a new

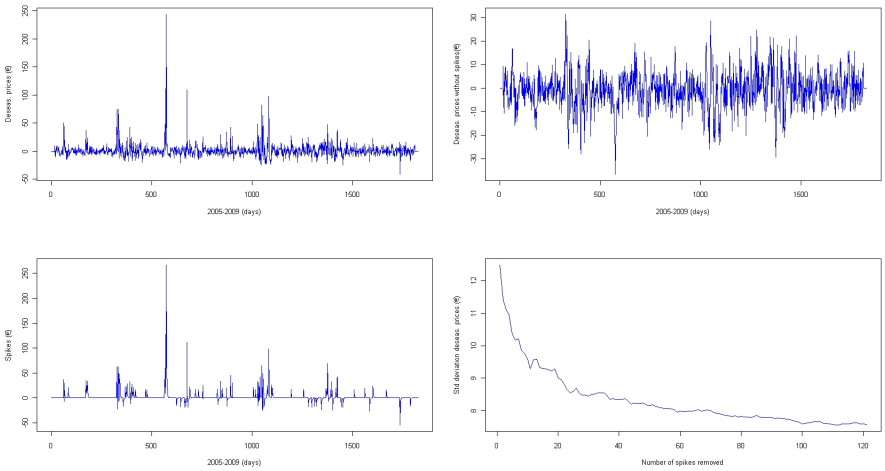

Figure 4: Filtering of price spikes using maximum likelihood (2005-2009).

spike is counted on that day. We assume a time-homogeneous jump intensity and this can be calculated by dividing the number of spikes by the total number of days with separate intensities for positive and negative spikes. A seasonally varying intensity could be calibrated (see [23]), but given the small sample of spike data and the relatively small impact a time varying intensity would have on our pricing, we only consider a constant intensity.

Positive and negative spikes should also be separated when fitting their sizes to distributions and given the small number of spikes, only distributions with two parameters should be considered. Furthermore, as our filtering technique is based on EVT, generalized Pareto distributions should be used (see [33]). Good candidates in this class include the Pareto and exponential distributions. As the VG distribution has exponential tails, in this work we choose the exponential distribution for the CP process. The fit of positive and negative spikes is shown in Figure 5 and the resulting parameters are given in Table 1, where λ^{-1} is the mean of the exponential distribution and s_0 is the minimum of the absolute value of the jump sizes. The speed of mean reversion can be deduced using the procedure outlined in Section 4. Due to the small amount of spike data collected, we omit the statistical hypothesis testing of the fitted distributions.

Note that prior to filtering the spikes from the deseasonalised prices, it is

Table 1: Compound Poisson parameters for X^3.

	Intensity	s_0	λ
Positive spikes	.031	6.52	.0653
Negative spikes	.018	-8.30	.0377

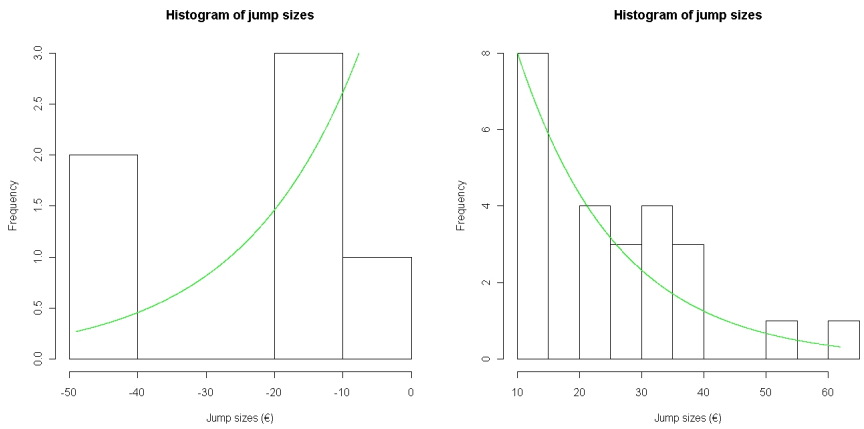

Figure 5: Fit of positive (left) and negative spikes to the exponential distribution.

necessary to 'flatten' the time series to remove changes in the long-term stochastic mean level. This is done by extracting a 31-day moving average from the deseasonalised prices before running the algorithm to filter the spikes. The moving average should then be added back to the price signal after the spikes have been removed.

6.4 Stochastic mean process

After the spikes have been removed, the time series must be separated into the factors X^1 and X^2. X^1 is extracted by filtering out a 31-day moving average (MA) which can be modelled as a Brownian motion, where the volatility is given by the standard deviation of the increments. The left hand panel of Figure 7 shows that the MA increments are well-modelled by a normal distribution and Table 2 gives the corresponding Kolmogorov-Smirnoff (KS) and Anderson-Darling (AD) test results - even the AD test, which places emphasis on the tails of the distribution, shows that the Normal hypothesis should not be rejected at the 2% level.

Table 2: Test statistics for X^1 fit.

	Test statistic	p-value
Kolmogorov-Smirnoff (Normal)	.031	.226
Anderson-Darling (Normal)	.911	.020

A period of 31 days is chosen for the MA as this corresponds to the median delivery period of the monthly contracts which we eventually wish to price. Note that this MA is distinct from the MA used to flatten the deseasonalised prices in the previous section.

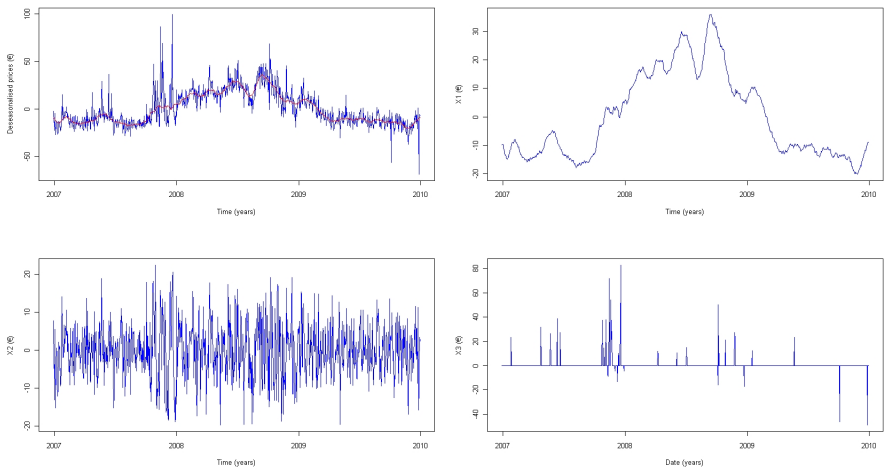

Figure 6: Decomposition of deseasonalised spot prices into X^1, X^2 and X^3.

6.5 Day-to-day variations

The remaining time series can finally be calibrated to a VG driven OU process by the procedure outlined in Section 4. The empirical density is plotted on the right of Figure 7 where leptokurtosis and the inability of the normal density to fit the data are clearly visible. In contrast, the VG distribution appears to provide a good fit and this is confirmed by the test results given in Table 3.

Table 3: Test statistics for X^2 fit.

	Test statistic	p-value
Kolmogorov-Smirnoff (Variance Gamma)	.016	.938
Kolmogorov-Smirnoff (Normal)	.049	.011
Anderson-Darling (Normal)	4.71	$< 10^{-10}$

The calibrated values for the for all factors are summarised in Table 4. We note that the rate of mean reversion, a_2, is larger than usually calibrated in 2 factors models (.74 versus .1 for a typical two factor model). This is to be expected as our model contains a non-mean-reverting process and a_2 is measured for reversion to a more realistic level.

Table 4: Calibrated parameters of spot price model.

Parameter	σ_1	m	a_2	C	σ	θ	κ	a_3
Value	.388	-.056	.784	0.207	7.74	-0.192	0.727	1.06

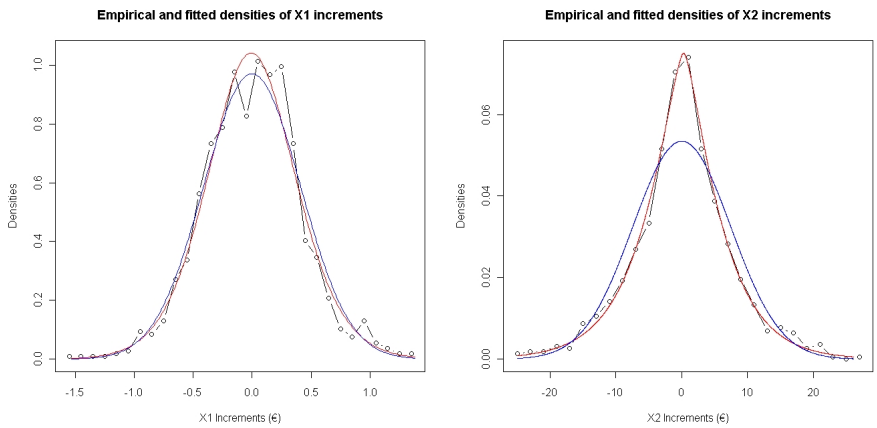

Figure 7: Empirical densities of X^1 (left) and X^2 increments with fitted VG (red) and normal (blue) densities.

7 Risk neutral calibration

7.1 Preliminaries

Henceforth, we assume that the filtration $(\mathcal{F}_t)_{t\geq 0}$ is generated by the DA spot price i.e. $\mathcal{F}_t = \sigma(S_t), t \geq 0$.

In the preceding chapter, a spot price model was calibrated to historical market prices. From a probabilistic point of view, this means that the model was calibrated under the 'physical' or 'historical' probability measure, $\mathbb{P}$. However, as the future spot price is uncertain, the fair value of a derivative written on the spot price will *not* be the conditional expectation of the discounted spot price under this measure i.e. the following holds:

$$v(t, S_t) \neq \mathbb{E}[g(S_T)|\mathcal{F}_t]$$

where $g : \mathbb{R} \mapsto \mathbb{R}$, a $\mathbb{P}$-integrable function of the spot price, is the derivative payoff and v is its current value.

Rather, rational market participants will charge a premium for entering into such a risky contract. Therefore, the fair value of a derivative will only equal the expected future spot price under an alternative 'risk-neutral' measure, $\mathbb{Q}$, that is equivalent to $\mathbb{P}$. The Fundamental Theorem of Asset Pricing guarantees, in the absense of arbitrage, the existence of a unique 'equivalent martingale measure' in complete markets where derivative contracts can be perfectly hedged. In order to use the spot price model (2.2) for derivatives pricing, we need to adjust the dynamics so that it is calibrated under a risk-neutral measure. As the market is incomplete a continuum of equivalent martingale measure will exist and one such measure must be chosen for pricing.

Few studies in the literature go into any great detail in the risk neutral calibration of electricity spot price models. [7] shows how a spot model can be adjusted so that the expected values coincide with futures prices, but only a handful of futures contracts are used with a relatively simple model based on a one factor mean-reverting jump diffusion. [9] goes into far greater detail with its two factor model and obtains risk premia properties similar to those described here in Section 7.3.2. We propose a new approach to specifying the market prices of risk, which is made possible by the addition of the non-mean-reverting factor X^1 in (2.2).

7.2 Spot-forward relationship

The essence of the following argument is taken from [7, Chapter 4]. We consider a market consisting of DA electricity with price dynamics specified by (2.2) and a risk-free bank account, B, with dynamics

$$dB(t) = rB(t)dt,$$

where r is a constant interest rate. As argued in Section 1, electricity is not considered a tradable asset and therefore, the only tradable asset in our market is the risk-free bank account. According to arbitrage theory, a risk neutral measure, $\mathbb{Q}$, is a probability measure equivalent to the measure $\mathbb{P}$, under which all tradable assets in the market are martingales after discounting (see, for example, [8, Theorem 7.8]). As B is trivially a martingale under any probability measure, we find that all meaures equivalent to $\mathbb{P}$ are risk neutral. For simplicity, we only consider measures equivalent to $\mathbb{P}$ given by the Radon-Nikodym derivative 5.4 of an Esscher transform. Nevertheless, there is still a continuum of risk-neutral measures from which we are free to choose, as is usual in incomplete markets.

The set of risk neutral measures can be reduced by adding to the market futures contracts. These contracts are tradable and so their discounted payoffs must be martingales under the risk neutral measure. If a futures contract, F, entered into at time t, is continuously settled and compounded over the delivery period $[\tau_1, \tau_2]$, the holder will receive the amount

$$\int_{\tau_1}^{\tau_2} e^{-r(u-t)}(S(u) - F(t, \tau_1, \tau_2))du.$$

As the futures contract is costless to enter we have the following identity:

$$\mathbb{E}^{\mathbb{Q}}[\int_{\tau_1}^{\tau_2} e^{-r(u-t)}(S(u) - F(t, \tau_1, \tau_2))du)|\mathcal{F}_t] = 0$$

Therefore, the following relationship between spot and futures prices holds:

$$F(t, \tau_1, \tau_2) = \mathbb{E}^{\mathbb{Q}}[\int_{\tau_1}^{\tau_2} w(u, \tau_1, \tau_2)S(u)du)|\mathcal{F}_t]. \tag{7.1}$$

w is a weighting function that depends on when the futures contract is settled:

$$w(u,s,t) = \frac{\hat{w}(u)}{\int_s^t \hat{w}(v)dv},$$

where the choice of $\hat{w}$ depends on the settlement i.e. $\hat{w}(u) = 1$ if settlement occurs at the end of the contract, while $\hat{w}(u) = e^{-ru}$ if settlement is continuous.

Finding a measure, $\mathbb{Q}$, that satisfies (7.1) for a range of market futures prices, F, and a given spot model price, S, can be quite complicated. It is more convenient to rewrite $F(t,\tau_1,\tau_2)$ in terms of a continuum of forward contracts, $f(t,u), u \in [\tau_1,\tau_2]$, each of which promises delivery at a single specific time, u, over $[\tau_1,\tau_2]$. Using the fact that $f(t,u) = F(t,u,u)$, (7.1) reduces to

$$f(t,u) = \mathbb{E}^{\mathbb{Q}}[S(u)|\mathcal{F}_t]. \tag{7.2}$$

By applying Fubini's theorem to (7.1) and then substituting (7.2), we have

$$F(t,\tau_1,\tau_2) = \int_{\tau_1}^{\tau_2} w(u,\tau_1,\tau_2) f(t,u)du. \tag{7.3}$$

Therefore, given market futures prices, $F(t,\tau_1,\tau_2)$, if we can solve (7.3) to find $f(t,u), u \in [\tau_1,\tau_2]$, we can then use the spot price model to find the risk neutral measure $\mathbb{Q}$ via (7.2).

Most institutions have algorithms that break a list of futures prices into a 'forward curve', thus solving (7.3). More information on this topic can be found in [7, Chapter 7]. We implement our own simple algorithm in Section 7.3.1 and then proceed to solve (7.2) in Sections 7.3.2 and 7.3.3.

Note that any tradable derivatives, and not just futures contracts could in theory be used to find $\mathbb{Q}$. However, in the case of electricity, futures contracts are the only derivatives consistently traded in a liquid manner for which data are readily available. As we are ultimately interested in pricing monthly contracts, we use monthly futures prices to calibrate our model and it is important to note that it is possible that different results could be obtained if quarterly or calendar contracts were used, even if all the contracts covered roughly the same period of delivery. The reason is not only due to the varying length of delivery, but also the differing attitudes to risk that may exist between the relatively small pool of market participants. For this reason, it is not feasible to calibrate the model with monthly futures and then price quarterly or calendar contracts with the results.

Finally we remark that (7.2) is only an approximation as it ignores deterministic information about the future that is available to market participants. For example, if a nuclear power plant is scheduled for maintenance during August, this will affect the futures price for August but will not affect the spot price prior to August. Therefore, before August, the filtration $(\mathcal{F}_t)_{t\geq 0}$ will not contain the information about the plant maintenance. Other information about the future is also not impounded in the spot price such as weather forecasts or

extraordinary future events which will cause an increase in demand. As investigated in [5], a more accurate representation of the spot-forward relationship is of the form

$$f(t,u) = \mathbb{E}^{\mathbb{Q}}[S(u)|\mathcal{F}_t \vee \mathcal{G}_t]$$

where the filtration $(\mathcal{G}_t)_{t\geq 0}$ is generated by the flow of deterministic information about the future not contained in the spot price. Fortunately, the maintenance of power plants is usually scheduled at the same time each year and so can be approximated in a seasonal variation of the risk premium (see Section 7.3.2). Moreover, weather forecasts are only accurate over a short period of time and other information, such as the scheduled opening of new power plants, will usually only have an effect on long-term calendar contracts. Therefore, for monthly contracts only five or six months in future, we assume that (7.2) will hold *on average*.

7.3 Calibration to monthly futures prices

Using market prices for monthly futures contracts with delivery periods ranging from June 2007 to June 2010, we now aim to calibrate the three factor model (2.2) under the risk neutral measure $\mathbb{Q}$. This will be carried out by applying an Esscher transform to the $\mathbb{P}$-calibrated dynamics of S so that (7.2) holds. Therefore, as argued above, we first need to solve (7.3). We use OTC monthly contracts which have a settlement date after the end of the month, so that the weighting function $\hat{w}$ has a constant value 1. As our spot price model is calibrated to daily baseload prices, we take $f(t,u)$ to be the forward price for delivery on day u so that (7.3) can be discretised as

$$F(t,\tau_1,\tau_2) = \frac{1}{\tau_2-\tau_1}\sum_{u=1}^{\tau_2-\tau_1} f(t,u). \tag{7.4}$$

7.3.1 Daily forward curve

Consider a set of months M and their corresponding monthly futures prices at time t, $F_m := F(t,\tau_m,\tau_{m+1}), \forall m \in M$. Using F_m and historical spot price data we wish to predict future daily baseload prices $f(d_m)$ for each day $d_m \in D_m$, the set of all days in month m. We do this by finding factors $a : D_m \mapsto \mathbb{R}$ such that $f(d_m) = a(d_m)F_m$. These factors should reflect the variation in the spot price of electricity over each day during the month, but ignore overall seasonal variation as this is already impounded in the futures price, F_m. One such set of daily factors can be obtained as follows:

1. Construct the seasonality function $\Lambda(t), \forall t \in [\tau_m,\tau_{m+1}], m \in M$, using the coefficients calibrated in (6.2).

2. Construct a second seasonality function $L(t), \forall t \in [\tau_m,\tau_{m+1}], m \in M$ that just uses the terms in (6.2)(a) and (b). Thus L captures seasonal price variation but ignores daily variations.

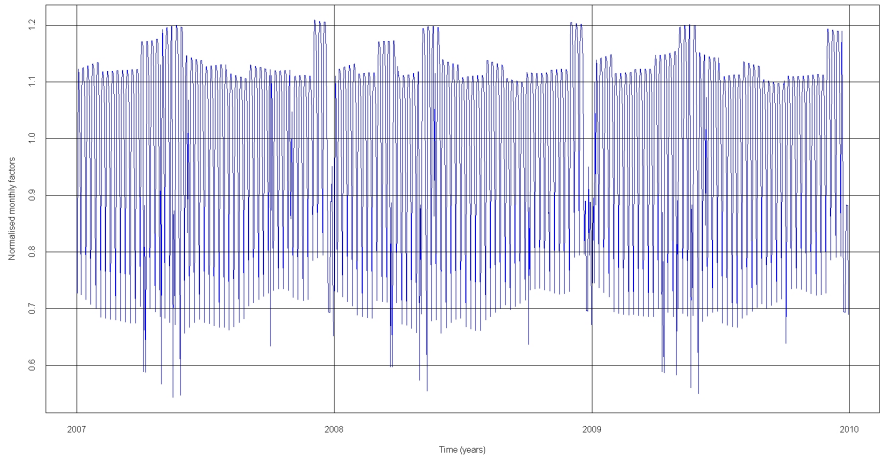

Figure 8: Daily prices factors for monthly contracts in 2007-2009.

3. Calculate
$$\hat{a}(d_m) := \frac{\Lambda(d_m)}{L(d_m)}, \quad \forall\, d_m \in D_m, m \in M.$$
(This is equivalent to dividing the values plotted in green in Figure 3 by those plotted in black).

4. For each fixed m, normalise $\hat{a}$ by setting
$$a(d_m) = \frac{\hat{a}(d_m)}{\sum_{d_m \in D_m} \hat{a}(d_m)}.$$
This ensures that
$$\sum_{d_m \in D_m} f(d_m) = F_m \times \#D_m.$$

The results of applying this algorithm to all months in the period 2007-2009 are illustrated in Figure 8. Using these values of a, we can find daily baseload prices given a set of monthly futures prices, thus finding a solution to (7.3).

7.3.2 Risk premia

Before applying an Esscher transform to the $\mathbb{P}$-dynamics of the spot price model, we wish to investigate the properties of the risk premia inherent in monthly futures prices. We calculate expected spot prices for each day in each month in the set M using Proposition 6.1 and compare the results to forward prices.

Figure 9 shows the average risk premium for each day of the week and each month of the year. A clear pattern in seasonal and daily variations can

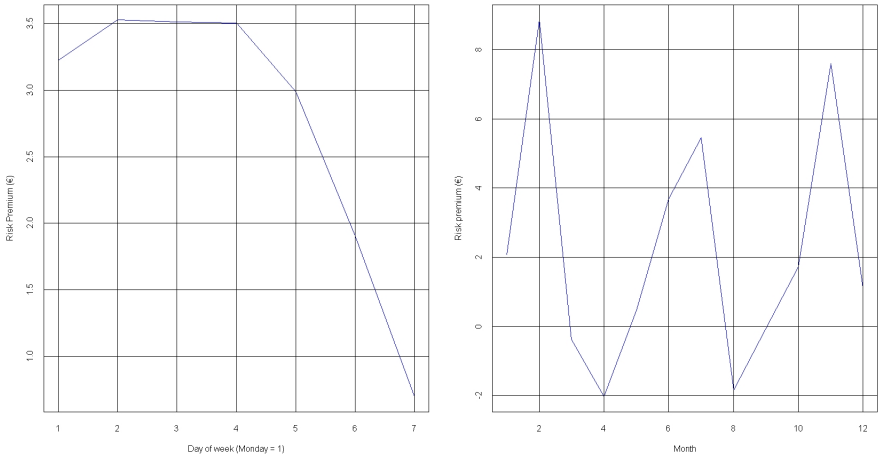

Figure 9: Average risk premia for each day of the week (left) and month of the year in monthly futures contracts from June 2007 - June 2010.

be observed with low risk premia in spring, late summer and at the weekend, corresponding to low demand and plentiful supply. Interestingly, December and January premia are lower than November and February. This is possibly due to increased liquidity and potential arbitrage opportunities involving quarterly and calendar contracts whose delivery periods end or begin at the turn of the year. Moreover, negative risk premia can be observed in April and August where producers are willing to accept low prices for futures contracts in order to hedge against over-supply.

Figure 10 shows the average risk premium per number of days to delivery. The graph is irregular for very short or long times to delivery and this can be explained by the total number of contracts traded for each number of days to delivery, which gives an indication of the overall liquidity of the prices. Between 5 and 150 days to delivery there is a pattern to the risk premia even though the graph oscillates unpredictably. Premia seem to increase until about 90 or 100 days to maturity when they decrease slightly, possibly due to arbitrage opportunities with quarterly contracts. Between 100 and 150 days to maturity risk premia increase again as time to delivery becomes further away in time.

7.3.3 Market prices of risk

An Esscher transform may be applied to any of the three factors in (2.2) so that the expected value coincides with any given forward price. However, this must be done in such a way that the resulting prices of risk form a definite pattern over time and over the time to delivery of each contract. We therefore make the following assumptions about the risk preferences of market participants.

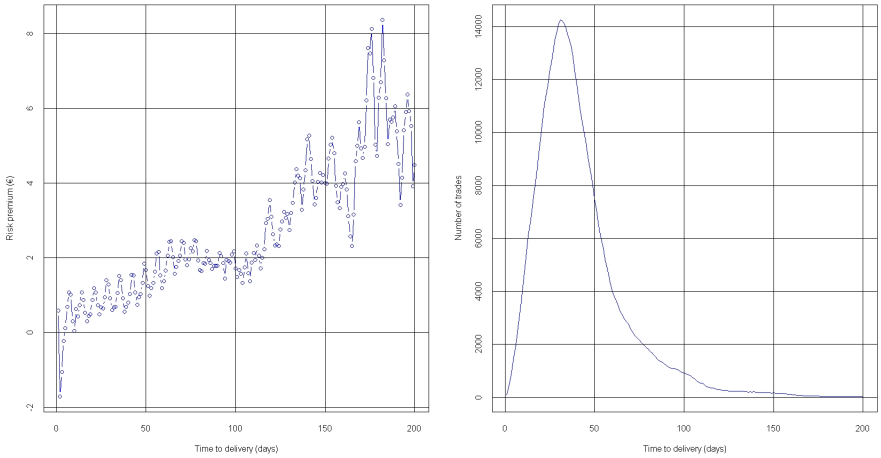

Figure 10: Average risk premia (left) and total number of contracts traded per number of days to delivery.

Assumption 7.1. *There is no risk premium associated with price spikes.*

This can be justified on the grounds that the monthly futures contracts are based on the average price of electricity over a whole month. Therefore a limited number of potential spikes that last for a day or two will have little impact on the futures price. Thus, we will not perform any adjustment of the dynamics of X^3.

Assumption 7.2. *The risk in day-to-day spot price variations depends solely on the forward delivery time.*

During periods of high demand in winter or midsummer it is possible that there will be large fluctuations in the day-to-day spot price of electricity and so futures contracts for these periods will be riskier. On the other hand, if a month contains a lot of public holidays or has low seasonal demand, the monthly futures price will be less risky. These are the patterns that were observed in Figure 9. This kind of risk is fixed for a particular delivery period and is independent of the time to delivery of the contract. Therefore, to the factor X^2, we associate a market price of risk, $\hat{\lambda}_T \in \mathbb{R}$, which only depends on the delivery date, T, of a forward contract.

Assumption 7.3. *The risk in the overall spot price level depends solely on the time to delivery of the forward contract.*

If there is a long time to the delivery of a forward contract (e.g. 6 months or more), there is a relatively large risk that spot prices will change their long-term deseasonalised level by a large amount. On the other hand, if there is only a days few to go before a futures contract starts delivering power, the risk that

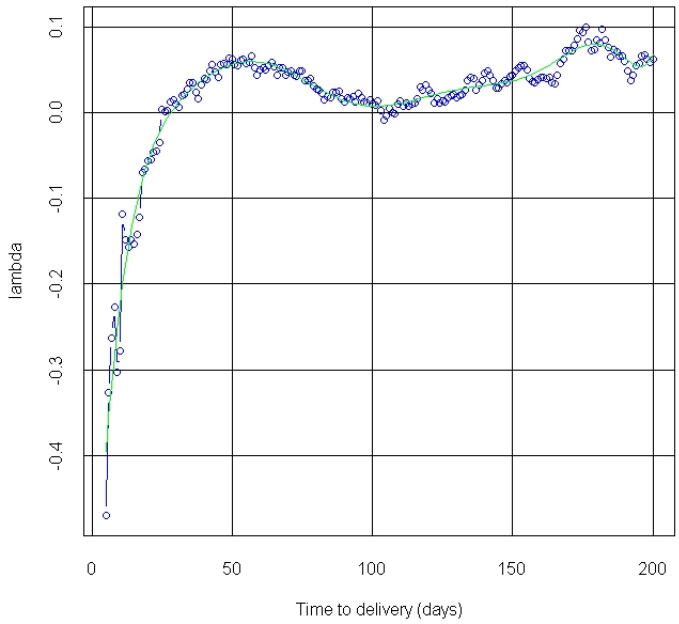

Figure 11: The market price of risk $\tilde{\lambda}$ per number of days to delivery.

the long-term spot price level will change is relatively small. This is the pattern that was just about discernible in Figure 10. Therefore, to the factor X^1, we associate a market price of risk, $\tilde{\lambda}(T-t) \in \mathbb{R}$, that is independent of the delivery date of a forward contract.

Using these assumptions, together with the results in Section 5, we define the risk neutral measure, $\mathbb{Q}$, via its Radon-Nikodym derivative with respect to $\mathbb{P}$, at the time of delivery of a forward contract, T, conditional on $\mathcal{F}_t$, as follows:

$$\begin{aligned} \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathcal{F}_T|\mathcal{F}_t} &= \exp\left(\int_t^T \tilde{\lambda}(T-s)dW_s - \frac{1}{2}\int_t^T \tilde{\lambda}^2(T-s)ds\right) \\ &\quad \times \exp\left(\hat{\lambda}_T(L(T)-L(t)) - \psi(i\hat{\lambda}_T)(T-t)\right) \end{aligned} \tag{7.5}$$

With this density process we apply an Esscher transform to the spot price dynamics and hence calculate the conditional $\mathbb{Q}$-expectation of the spot price.

Proposition 7.4. *The expectation of the spot price, S, defined in* (2.2)*, under*

the measure $\mathbb{Q}$ as defined in (7.5)*, is*

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}}\left[S(T)|\mathcal{F}_t\right] = \Lambda(T) + \sigma_1 \underbrace{\left(W_t^{\mathbb{Q}} + \int_t^T \tilde{\lambda}(s)ds\right)}_{(a)} + X_t^2 e^{-a_2(T-t)} \\ + (1-e^{-a_2(T-t)})\left(m + \frac{1}{a_2}\underbrace{\left[\frac{2(A+\hat{\lambda}_T)}{\kappa(B^2-(A+\hat{\lambda}_T)^2)}\right]}_{(b)}\right) \\ + X_t^3 e^{-a_3(T-t)} + (1-e^{-a_3(T-t)})\frac{\mu_P}{a_3}\end{aligned} \tag{7.6}$$

where $W^{\mathbb{Q}}$ is a $\mathbb{Q}$-Brownian motion and A, B and κ are the parameters of the VG process as defined in Section 3.2.

Proof. (7.6) only differs from (6.1) in terms (a) and (b) and (a) is given by (5.2), while (b) is given by (5.5). □

Remark 7.5. *As discussed in Section 5.3, applying the Esscher transform to S adjusts the Lévy measure of the process L so that it becomes:*

$$\nu_2^{\hat{\lambda}_T}(dz) = \frac{1}{\kappa|z|}e^{(A+\hat{\lambda}_T)z-B|x|}dz$$

Note that because $\hat{\lambda}_T$ is fixed for every forward contract and independent of time, t, L remains a time homogeneous process after the Esscher transform.

We can now apply (7.6) and (7.2) to a collection of forward prices to find $\tilde{\lambda}$ and $\hat{\lambda}$ as follows:

1. Set $\tilde{\lambda} = 0$ and for each forward price, $f(t,T)$, find $\hat{\lambda}_T$ so that (7.2) holds.
2. Set $\hat{\lambda}_T$ as the weighted average of all contracts with delivery, T, on the same day of the week and the same month.
3. With $\hat{\lambda}_T$ fixed in (7.6), for each forward price, $f(t,T)$, find $\tilde{\lambda}(T-t)$ so that (7.2) holds.
4. Set $\tilde{\lambda}(T-t)$ as the weighted average of all forward prices with the same time to delivery, $T-t$.

Note that the averages above are weighted by the number of contracts that were traded at each forward price, so that we give higher weight to the most liquid forward prices.

As $\hat{\lambda}$ remains inside the near-linear region of Figure 2, there is a near-linear one-to-one relationship between $\hat{\lambda}_T$ and the risk premia in Figure 9. However, the results for $\tilde{\lambda}$ are significantly more regular than the risk premia in Figure 10

and are displayed in Figure 11, together with a polynomial fitted to the values for $\tilde{\lambda}(T-t)$. The regularity of this curve means that it can be used to price futures contracts and options on the spot and futures price, which is done in the next section.

8 Pricing

8.1 Monthly Futures

Using the $\mathbb{Q}$-calibration of the spot model, we price monthly futures contracts. The results for six days in April and May 2009 are shown in Figure 12 along with a histogram of the percentage errors in the predicted prices in Figure 13. We can see that the median error is between 2 and 4% while most errors are less than 10%. Ten out of thirty-five of the prices have errors greater than 10% and from Figure 12, it would appear that these mostly occur in contracts with relatively long times to delivery i.e. 5 or 6 months.

In particular, the predicted prices for months 5 and 6 (M5 and M6) in April are consistently too large, which would suggest that the market price of risk for long times to delivery has been overestimated. However, the errors for M5 and M6 during May are less stark, which adds credibility to the specified price of risk. Furthermore, while the error for M6 during April is large, the error for the same contract during May, M5, decreases over time until it is negligable on the last date shown, 19th May. This begs the question: was M6 mis-priced during April?

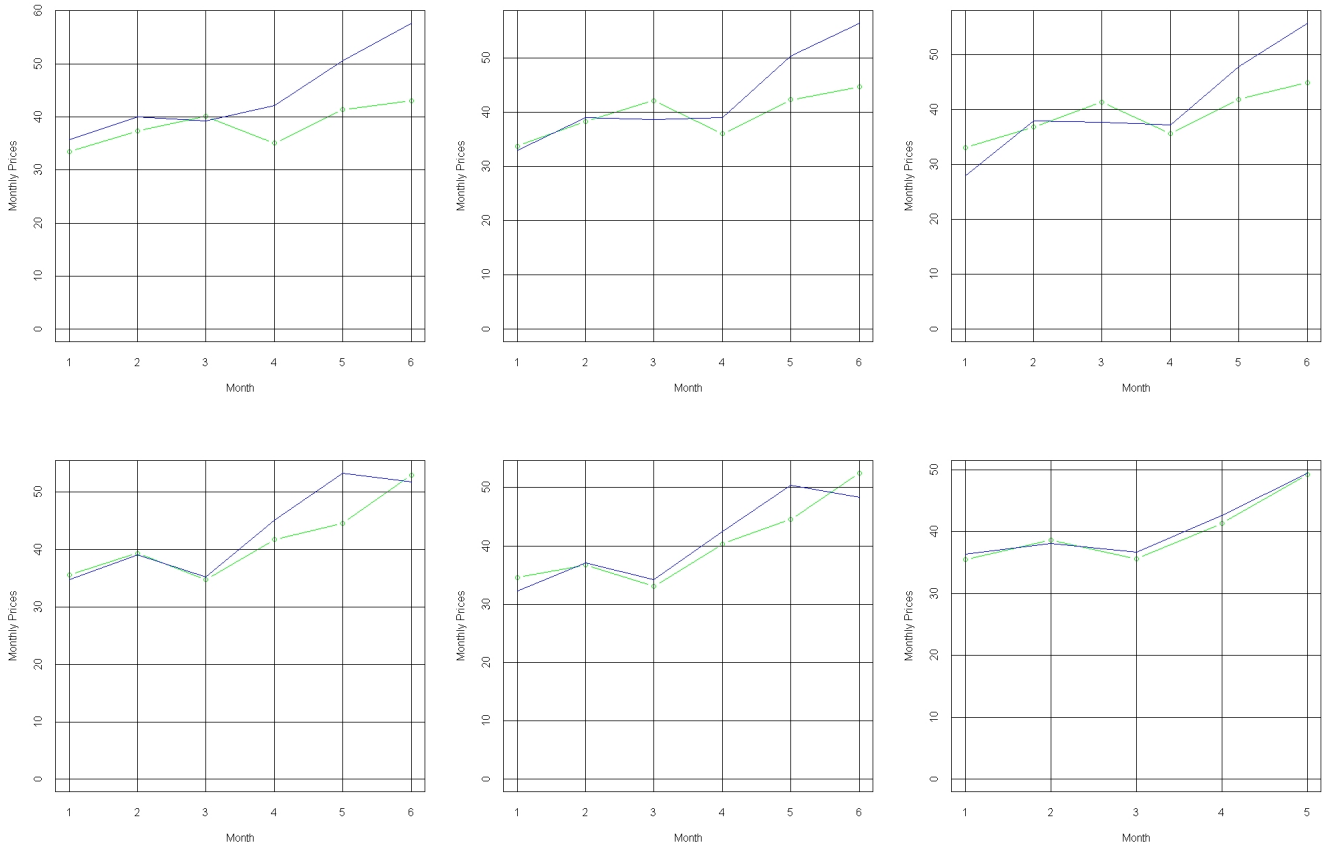

Figure 12: Predicted monthly futures prices (blue) compared with market prices (green) on 2nd, 9th, 16th April 2009 (top) and 5th, 12th, 19th May 2009.

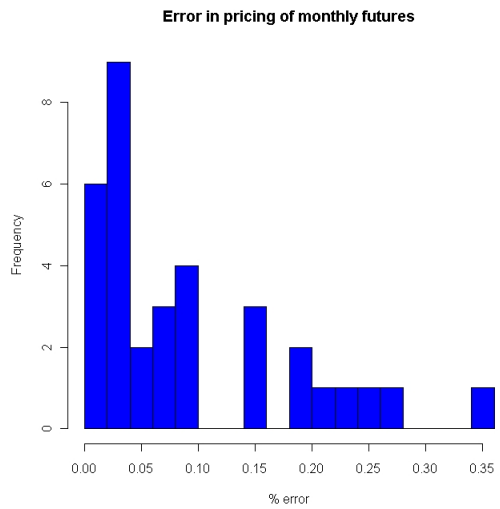

Figure 13: Percentage errors of predicted prices of contracts displayed in Figure 12.

Another example is given in Figure 14 where prices on almost the same days in 2008 and 2009 are compared. The errors for M2 and M3 in 2008 are small (4.6% and 5.4% respectively) but in 2009 the errors are high (33% and 24% respectively). The reason for this is unclear, but given the small errors in 2008, the cause is probably not mis-specification of the market prices of risk or the seasonality function. Rather, deterministic information, not included in the filtration $(\mathcal{F}_t, t \geq 0)$ could have been available to market participants, causing the depressed market prices for M2 and M3 in 2009 e.g. it could be that more power plants than usual were scheduled to be online in September and October 2009. In this light, an improvement in the model could include such deterministic information in another filtration, which could then be used to adjust the expected spot price appropriately (see [5] for examples).

8.2 Options on spot price

We next price vanilla options on the electricity baseload spot price. Such contracts are traded in small quantities on the OTC market and have maturity at 9a.m. daily, before the start of the auction which sets the DA price (i.e. fifteen hours before the start of the delivery period).

The results contained in this section are also useful for valuing the optionality inherent in tolling and power purchase agreements. Holders of such contracts must nominate (within contraints) on a daily basis how much power a plant should produce and sell to the market. Thus, the contract holder has implicit European options on the spot price of electricity (see [22] for more details).

We use the Finite Element Method (FEM) to price the options numerically

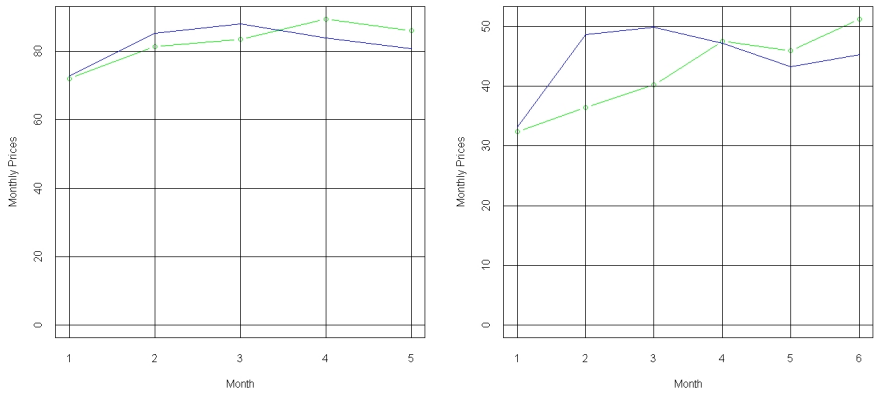

Figure 14: Predicted and market monthly futures prices on 18th July 2008 (left) and 17th July 2009.

and a brief overview of the method is given in Appendix A.

8.2.1 Call option value

We define the value of a generic European type derivative contract on the spot price with payoff function $h : \mathbb{R} \mapsto \mathbb{R}$ as follows:

$$v(t, S_t) = \mathbb{E}^{\mathbb{Q}}[e^{-r(T-t)} h(S(T)) | \mathcal{F}_t]$$

Assuming that the payoff h is that of a European call option, and using the fact that the spot price depends explicitly on the factors X^1, X^2 and X^3 we can rewrite the above as

$$v(t, X_t^1, X_t^2, X_t^3) = \mathbb{E}^{\mathbb{Q}}[e^{-r(T-t)} \left(X_T^1 + X_T^2 + X_T^3 - (K - \Lambda(T))\right)^+ | \mathcal{F}_t]$$

where K is the strike price. In order to simplify matters later on, we rewrite X_T^2 as

$$X_T^2 = X_t^2 e^{-a_2(T-t)} + (m + \frac{\mu_2}{a_2})(1 - e^{-a_2(T-t)}) + \int_t^T e^{-a_2(T-s)} d\tilde{L}_s$$

where $d\tilde{L}(s) = \int_{\mathbb{R}\setminus\{0\}} z\tilde{J}_2(dt, dz)$ is the fully compensated VG process and μ_2 is its $\mathbb{Q}$-expectation. For $s \geq t$ we define

$$\hat{X}_s^2 := X_s^2 - (m + \frac{\mu_2}{a_2})(1 - e^{-a_2(s-t)})$$

so we can write $\hat{X}_T^2$ as

$$\hat{X}_T^2 = \hat{X}_t^2 e^{-a_2(T-t)} + \int_t^T e^{-a_2(T-s)} d\tilde{L}_s. \tag{8.1}$$

Similarily, we define $\hat{X}^3_T := X^3_T - \frac{\mu_3}{a_3}(1 - e^{-a_2(T-t)})$ and write

$$\hat{X}^3_T = \hat{X}^3_t e^{-a_3(T-t)} + \int_t^T e^{-a_3(T-s)} d\tilde{P}_s. \tag{8.2}$$

Finally, as we are pricing under the measure $\mathbb{Q}$, the factor X^1 also contains a drift term, which can be removed by defining

$$\hat{X}^1_t := X^1_t - \sigma_1 \int_0^t \tilde{\lambda}(s) ds. \tag{8.3}$$

Using the above definitions we can rewrite the call option value as

$$v(t, \hat{X}^1_t, \hat{X}^2_t, \hat{X}^3_t) = \mathbb{E}^\theta[e^{-r(T-t)} \left(\hat{X}^1_T + \hat{X}^2_T + \hat{X}^3_T - k(t,T)\right)^+ |\mathcal{F}_t] \tag{8.4}$$

where $k : \mathbb{R}_+ \times \mathbb{R}_+ \mapsto \mathbb{R}$ is defined as

$$\begin{aligned} k(t,T) =& K - \Lambda(T) \\ & - \sigma_1 \int_0^t \tilde{\lambda}(s) ds - (m + \frac{\mu_2}{a_2})(1 - e^{-a_2(T-t)}) - \frac{\mu_3}{a_3}(1 - e^{-a_2(T-t)}). \end{aligned}$$

In the sequel, for convenience we will simply write X^k for $\hat{X}^k$, $k = 1, 2, 3$.

8.2.2 PIDE derivation

Let $X_t = (X^1_t, X^2_t, X^3_t)^\top$ be a vector of the processes defined in (2.2). By applying Itô's Lemma for Lévy processes of finite variation (3.5) to the real-valued function, $f \in C^{1,2}(J \times \mathbb{R}^3)$, where $J = [0, T]$, we get the following dynamics:

$$\begin{aligned} df(t, X_t) =& \partial_t f dt + \sum_{i=1}^3 \partial_{x_i} f dX^i_t + \frac{\sigma_1^2}{2} \partial_{x_1 x_1} f dt \\ & + f(X_{t-} + \Delta X_t) - f(X_{t-}). \end{aligned}$$

Using the Lévy-Itô representation of the processes L and P, we have in differential form

$$\begin{aligned} df(t, X) =& \left(\partial_t f + \frac{\sigma_1^2}{2} \partial_{x_1 x_1} f\right) dt + \sigma_1 \partial_{x_1} f dW_t \\ & + \left(-a_2 x_2 dt + \int_{\mathbb{R}\setminus\{0\}} z \tilde{J}_2(dt, dz)\right) \partial_{x_2} f \\ & + \left(-a_3 x_3 dt + \int_{\mathbb{R}\setminus\{0\}} z \tilde{J}_3(dt, dz)\right) \partial_{x_3} f \\ & + f(X_{t-} + \int_{\mathbb{R}^3\setminus\{\mathbf{0}\}} z J(dt, dz)) - f(X_{t-}). \end{aligned}$$

where $J = (0, J_2, J_3)^\top$ is the 3-d jump measure of the independent Lévy processes W, L and P. We define the corresponding Lévy measure as $\nu := (0, \nu_2, \nu_3)^\top$. By adding and subtracting $\nu(dz)$ and making further simplifications we have

$$
\begin{aligned}
df(t,X) = & \underbrace{\left(\partial_t f + \frac{\sigma_1^2}{2}\partial_{x_1x_1} f - a_2x_2\partial_{x_2} f - a_3x_3\partial_{x_3} f\right)}_{(a)} dt \\
& + \underbrace{\int_{\mathbb{R}^3\setminus\{\mathbf{0}\}} f(X_{t-}+z) - f(X_{t-})\nu(dz)\, dt}_{(a)} \\
& + \underbrace{\int_{\mathbb{R}^3\setminus\{\mathbf{0}\}} f(X_{t-}+z) - f(X_{t-})\tilde{J}(dt,dz)}_{(b)} \\
& + \underbrace{\sigma_1\partial_{x_1} f dW_t + \partial_{x_2} f \int_{\mathbb{R}\setminus\{0\}} z\tilde{J}_2(dt,dz) + \partial_{x_3} f \int_{\mathbb{R}\setminus\{0\}} z\tilde{J}_3(dt,dz)}_{(b)}.
\end{aligned}
$$

The integral form of all terms in (b) are martingales (see [25, Proposition 9.3.1]). Therefore, in order for $f(t,X)$ to be a martingale, (a) must be identically zero. This leads naturally to the following proposition.

Proposition 8.1. *Let $v \in C^{1,2}(J \times \mathbb{R}^3) \cap C^0(\overline{J} \times \mathbb{R}^3)$ with bounded derivatives in x be a solution of*

$$
\begin{aligned}
\partial_t v + \mathcal{A}v - rv = 0 \quad & \text{in } J \times \mathbb{R}^3 \\
v(T,x) = g(x) \quad & \text{in } \mathbb{R}^3
\end{aligned}
$$

where $g : \mathbb{R} \mapsto \mathbb{R}$ is the payoff of a European call option and

$$
(\mathcal{A}v)(x) = \frac{\sigma_1^2}{2}\partial_{x_1x_1} v + \mathbf{b}(x)\cdot\nabla v + \int_{\mathbb{R}^3\setminus\{\mathbf{0}\}} v(x+z) - v(x)\nu(dz)
$$

with $\mathbf{b}(x) = (0, -a_2x_2, -a_3x_3)^\top$. Then v can also be represented as

$$
v(t,x) = \mathbb{E}^{\mathbb{Q}}[e^{-r(T-t)}h(X_T)|\mathcal{F}_t].
$$

We can change to time-to-maturity and remove the interest rate r by setting $u(t,x) = e^{rt}v(T-t,x)$. We then have the *initial* condition $u(0,x) = g(x)$ in $\mathbb{R}^3$ and $u(t,x)$ satisfies

$$
\partial_t u - \mathcal{A}u = 0 \text{ in } [0,T] \times \mathbb{R}^3.
$$

Furthermore, using [25, Lemma 13.2.7], in the special case of the independent Lévy copula, we can re-write the integral term of $\mathcal{A}u$ as

$$\sum_{j=2,3} \int_{\mathbb{R}\setminus\{0\}} u(x+z_j) - u(x)\nu_j(dz_j).$$

Using the densities $k_j(z_j)$ of the jump measures ν_j and by integrating by parts twice with respect to z_j, the resulting integrand is integrable around 0 and we get the final form of $\mathcal{A}$:

$$(\mathcal{A}u)(x) = \frac{\sigma_1^2}{2}\partial_{x_1x_1}u + \sum_{j=2,3}\{-a_jx_j\partial_{x_j}u + \int_{\mathbb{R}} \partial_{z_jz_j}u(x+z_j)k_j^{(-2)}(z_j)dz_j\} \quad (8.5)$$

where $k_j^{(-i)}(z)$ is the i-th antiderivative of k_j:

$$k_j^{(-i)}(z) = \begin{cases} \displaystyle\int_{-\infty}^{z} k^{(-i+1)}(x)dx & \text{if } z < 0 \\ \displaystyle -\int_{z}^{\infty} k^{(-i+1)}(x)dx & \text{if } z > 0. \end{cases}$$

Note that we will ignore the negative jump part of the Lévy measure ν_3 in pricing the option. Negative spikes occur more rarely than positive spikes and, as only a small amount of negative spike data was collected (see Figure 5), the fitted jump size distribution may be unreliable.

8.2.3 Variational Formulation

To obtain the weak formulation of the PIDE, we multiply the equation by a test function and integrate. By then performing partial integration wherever we have a second derivative, we obtain the following variational formulation:

$$\text{Find } u \in L^2(J; H^1(\mathbb{R}^3)) \cap H^1(J; L^2(\mathbb{R}^3)) \text{ such that}$$

$$(\partial_t u, v) + a(u,v) = 0, \forall v \in H^1(\mathbb{R}^3) \text{ a.e. in } J,\ u(0) = g(x)$$

and the bilinear form $a(\cdot\,,\cdot\,) : H^1(\mathbb{R}^3) \times H^1(\mathbb{R}^3) \mapsto \mathbb{R}$ is defined as

$$\begin{aligned} a(u,v) = &\underbrace{\frac{\sigma_1^2}{2}(\partial_{x_1}u, \partial_{x_1}v)}_{(a)} \\ &+ \underbrace{\sum_{j=2,3}\left[(a_jx_j\partial_{x_j}u, v) + \int_{\mathbb{R}^3}\int_{\mathbb{R}} \partial_{y_j}u(y)\partial_{x_j}v(x)k_j^{(-2)}(y_j - x_j)dy_jdx\right]}_{(b)} \end{aligned} \quad (8.6)$$

where we have used the change of variable $y = x + z_j$ such that for fixed j,

$$ye_k = \begin{cases} xe_k + z_j & \text{if } k = j \\ xe_k & \text{otherwise} \end{cases} \quad k = 1,2,3.$$

Remark 8.2. *In order to show that the above variational formuation is well-posed, the bilinear form, $a(\cdot,\cdot)$, must be shown to be continuous and to satisfy a Gårding inequality. This is done for* (8.6)(a) *in [25, Proposition 4.2.1]). However, the space on which* (8.6)(b) *is defined is $L^2(\mathbb{R}^2) \times L^2(\mathbb{R}^2)$ and well-posedness cannot be shown in the same way*[3]*. Nevertheless, as descibed in [25, Section 9.2.3], convergence with optimal rate can still be shown numerically.*

The unbounded domain $\mathbb{R}^3$ places unrealistic integrability constraints on the payoff function g and, numerically, it is not a feasible domain of computation. We therefore truncate the domain to $G = (R_1^l, R_1^u) \times (R_2^l, R_2^u) \times (R_3^l, R_3^u)$ and denote the solution to the localised problem as u_G. As the domain G will need to be much larger than realistic values of X^1, X^2 and X^3, R_1^u, R_2^u and R_2^u will have values between 100 and 400. Therefore, at the upper boundaries of the domain the terms $a_j x_j \partial_{x_j} u, j = 2,3$ will dominate and with the large values for x_j, it is reasonable to assume that $\partial_{x_j} u$ will be small. For this reason, we choose Neumann boundary conditions for our PIDE: $\forall u_G \in H_0^1(G)$, $\partial_{x_i} u_G = 0$, $i = 1,2,3$ on ∂G, where ∂G denotes the boundary of G.

The variational formulation then becomes

$$\text{Find } u_G \in L^2(J; H^1(G)) \cap H^1(J; L^2(G)) \text{ such that}$$

$$(\partial_t u_G, v_G) + a(u_G, v_G) = 0, \forall v_G \in H^1(G) \text{ a.e. in } J,\ u_G(0) = g(x)|_G.$$

8.2.4 Discretisation

We discretise the domain $J \times G$ with grid points $(t_m, x_i), x_i = (x_{i_1}, x_{i_2}, x_{i_3})$ defined by

$$x_{i_j} = R_j^l + ih_j,\ i = 0,1,\ldots,N_j+1,\ j = 1,2,3,\ h_j = (R_j^u - R_j^l)/(N_j+1)$$

$$t_m = mk,\ m = 0,1,\ldots,M,\ k = T/M$$

and solve the variational formulation of the PIDE in the linear finite element space, $V_N = S_{\mathcal{T}}^3 \cap H_0^1(G)$, with $S_{\mathcal{T}}^3$ spanned by the product of hat-functions:

$$S_{\mathcal{T}}^3 = \text{span}\{\prod_{j=1}^{3} b_{i_j}(x_j) : 1 \le i_j \le N_j\}$$

In order to specify the form of the stiffness matrix $\mathbf{A} \in \mathbb{R}^{N\times N}, N = N_1N_2N_3$, we compute the value of an arbitrary element:

$$\begin{aligned}\mathbf{A}_{ij} &= a(b_{j_1}b_{j_2}b_{j_3}, b_{i_1}b_{i_2}b_{i_3}) = \tfrac{\sigma_1^2}{2}\int b'_{j_1}b_{j_2}b_{j_3}b'_{i_1}b_{i_2}b_{i_3}dx \\ &+ \int a_2x_2b_{j_1}b'_{j_2}b_{j_3}b_{i_1}b_{i_2}b_{i_3}dx + \int b_{j_1}b'_{j_2}b_{j_3}b_{i_1}b'_{i_2}b_{i_3}k_2^{(-2)}(y_2-x_2)dxdy_2 \\ &+ \int a_3x_3b_{j_1}b_{j_2}b'_{j_3}b_{i_1}b_{i_2}b_{i_3}dx + \int b_{j_1}b_{j_2}b'_{j_3}b_{i_1}b_{i_2}b'_{i_3}k_3^{(-2)}(y_3-x_3)dxdy_3\end{aligned}$$

[3] If $k_2^{(-2)}$ was the density of a CGMY process with $Y > 1$, instead of a VG process, it would be possible show well-posedness for that factor in the usual way.

$$= \frac{\sigma_1^2}{2}\mathbf{S}^1_{ij}\mathbf{M}^2_{ij}\mathbf{M}^3_{ij} + \mathbf{M}^1_{ij}\mathbf{B}^{a_2x_2}_{ij}\mathbf{M}^3_{ij} + \mathbf{M}^1_{ij}\mathbf{L}^2_{ij}\mathbf{M}^3_{ij}$$

$$+\mathbf{M}^1_{ij}\mathbf{M}^2_{ij}\mathbf{B}^{a_3x_3}_{ij} + \mathbf{M}^1_{ij}\mathbf{M}^2_{ij}\mathbf{L}^3_{ij}$$

where for $l = 1, 2, 3$, the elements of the matrices $\mathbf{S}^l$, $\mathbf{M}^l$, $\mathbf{B}^{a_lx_l}$ and $\mathbf{L}^l$ are defined as follows:

$$\begin{aligned} \mathbf{S}^l_{ij} &= \int b'_{j_l} b'_{i_l} dx_l, & \mathbf{M}^l_{ij} &= \int b_{j_l} b_{i_l} dx_l, \\ \mathbf{B}^{a_lx_l}_{ij} &= \int a_l x_l b'_{j_l} b_{i_l} dx_l, & \mathbf{L}^l_{ij} &= \int b'_{j_l} b'_{i_l} k_l^{(-2)}(y_l - x_l) dx_l dy_l. \end{aligned}$$

Hence the form of the stiffness matrix can be written down as

$$\begin{aligned} \mathbf{A} =& \frac{\sigma_1^2}{2}\mathbf{S}^1 \bigotimes \mathbf{M}^2 \bigotimes \mathbf{M}^3 \\ &+ \mathbf{M}^1 \bigotimes [(\mathbf{B}^{a_2x_2} + \mathbf{L}^2) \bigotimes \mathbf{M}^3 + \mathbf{M}^2 \bigotimes (\mathbf{B}^{a_3x_3} + \mathbf{L}^3)]. \end{aligned} \tag{8.7}$$

Note that in the derivation of the (8.7), the stiffness matrix tensorises so that it can be written as a sum of tensor products of matrices that each only act in one dimension. This is a direct consequence of using independent stochastic factors and a product basis for the space $S^3_{\mathcal{T}}$.

By replacing the time derivative with a finite difference approximation and using the θ-scheme in time, the variational formulation is converted to the following matrix problem:

$$\text{Find } \underline{u}^{m+1} \text{ such that } m = 0, \ldots, M-1$$

$$(\mathbf{M} + \theta k \mathbf{A})\underline{u}^{m+1} = (\mathbf{M} - (1-\theta)k\mathbf{A})\underline{u}^m \ , \quad \underline{u}^0_N = \underline{u}_0, \tag{8.8}$$

where $\mathbf{M} = \mathbf{M}^1 \bigotimes \mathbf{M}^2 \bigotimes \mathbf{M}^3$. In our prototypical implementation we build the whole stiffness matrix at once and store it in memory. However, the tensorised form of the matrices $\mathbf{M}$ and $\mathbf{A}$ implies that the discretised problem could be solved recursively in each dimension, thus improving computational efficiency.

We solve the matrix problem on the domain $G = [-100, 100] \times [-100, 100] \times [-50, 400]$ with parameters defined in Table 4. Figure 15 displays the solution with respect x_2 and x_3. We can clearly see that as both variables grow, the solution becomes increasingly oscillatory. This effect is due to numerical instability known as *convection dominance* and this problem and its solution by streamline diffusion stabilization is discussed in the following subsections.

8.2.5 The problem of convection-dominance

Consider the following multidimensional parabolic PDE for an unknown scalar-valued function u on an open domain $G \subset \mathbb{R}^d$:

$$\begin{aligned} -a\Delta u(x) + \mathbf{b}(x) \cdot \nabla u(x) &= s(x) \quad \text{in } G \\ u(x) &= u_G \quad \text{on } \partial G \end{aligned}$$

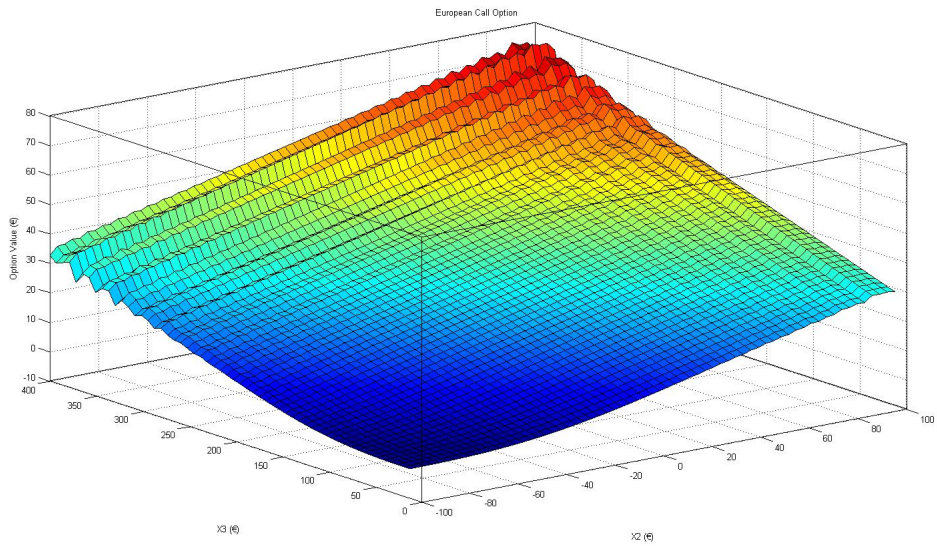

Figure 15: European call option value relative to x_2 and x_3.

where $a \in \mathbb{R}$, s is a scalar-valued function and $\mathbf{b}$ is a vector-valued function of $x \in G$. This type of equation is known as a convection-diffusion PDE and associated with its Finite Element discretisation we can define a *Péclet number* along each coordinate axis as

$$Pe_\kappa = \frac{\|\mathbf{b}\|_{\infty,\kappa} h_\kappa}{2a}$$

where h_κ is the mesh width of the partition along the e_κ axis. When Pe_κ is greater than 1, the diffusion term becomes insignificant and the PDE becomes *convection dominated*. In these circumstances, it is well known that the Finite Element solution is numerically unstable in that oscillations are introduced that are not present in the true solution. The remainder of this section offers some explanation as to why this occurs.

In general, diffusion terms in a PDE have a smoothing effect on the solution. This can be seen by considering the heat equation:

$$u_t - \Delta u = 0 \quad \text{in } \mathbb{R}^d \times (0, \infty). \tag{8.9}$$

If we choose the initial condition $u(x, 0) = \delta(x)$, where δ is the dirac function, the resulting solution is called the *fundamental* solution. It is well known that the fundamental solution of (8.9) is a Gaussian density:

$$u(x, t) = \frac{1}{\sqrt{4\pi t}} e^{-x^2/4t}.$$

Thus the discontinuous initial condition, δ, is instantaneously smoothed out to a $C^\infty(\mathbb{R}^d)$ function for $t > 0$ (see [21] for more details).

When the mesh Péclet number is large, the smoothing properties of the diffusion in the analytic solution cannot be taken into account by FEM and the Finite Element solution in turn over- and underestimates the true solution at each node. The following two examples illustrate these ideas further. The first is taken from [27] and shows how FEM breaks down when the coefficent of the diffusion term is much smaller than the mesh size.

Example 8.3. *Consider the boundary value problem*

$$-\epsilon\partial_{xx}u + \partial_x u = 0, 0 < x < 1;\ u(0) = 1, u(1) = 0$$

where $0 < \epsilon \ll 1$. *The solution can be determined analytically as*

$$u(x) = \frac{1 - e^{-\frac{1-x}{\epsilon}}}{1 - e^{-\frac{1}{\epsilon}}}.$$

Thus, $u(x)$ *is close to 1 except in a neighbourhood of* $x = 1$ *of width* $\mathcal{O}(\epsilon)$ *where it decays from 1 to 0.*

However, if FEM is applied to this problem we get the following system of equations for the finite element solution, U_i, *at the gridpoints,* $x_i = ih$, $i = 0, \ldots, N$, *where* $x_N = 1$.

$$-\frac{\epsilon}{h^2}\underbrace{[U_{i+1} - 2U_i + U_{i-1}]}_{(a)} + \underbrace{\frac{1}{2h}[U_{i+1} - U_{i-1}]}_{(b)} = 0, i = 1, \ldots, N-1 \tag{8.10}$$

$$U_0 = 1, U_N = 0.$$

For small ϵ *(large mesh Péclet number),* (8.10)(b) *dominates and, beginning at* U_0, U_i *is 1 for even* i. *For odd* i, *we have* (8.10)$(a) = 0$ *resulting in* $U_i = 0$ *and so the Finite Element solution oscillates around the true solution.*

A similar example is discussed in [19, Section 2.2.2] and the authors elaborate by determining, using the analytic solution, a discretised form which does not produce an oscillating solution. This particular discretisation differs from (8.10) in the coefficient of the term (a). That is, ϵ becomes $\epsilon + \delta$, which is equivalent to adding an artificial diffusion.

The next example is more closely related to option pricing and illustrates numerical instability when the drift of an asset's price process is disproportionally large.

Example 8.4. *Consider the parabolic PDE on an open domain* $G \in \mathbb{R}$

$$\begin{aligned} \partial_t u(t,x) + bx\partial_{xx}u(t,x) - a\partial_x xu(t,x) &= 0 \quad \text{in } [0,T] \times G \\ \partial_x u(t,x) &= 0 \quad \text{on } \partial G \end{aligned}$$

with initial condition $u(0,x) = (x-K)^+$ *and* $a, b \in \mathbb{R}$. *Such a PDE represents a European call option on an asset with Brownian motion driven mean-reverting*

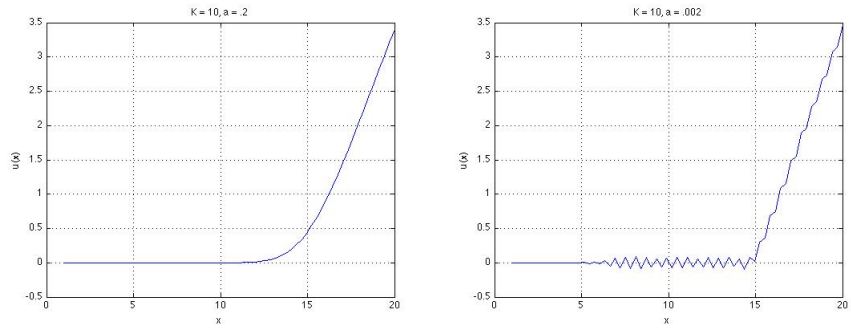

Figure 16: Convection Dominance in European call option values with strike K and coefficient of diffusion a.

price dynamics. Figure 16 shows the result of solving this system using FEM with a basis of hat functions, using $G = [1, 20]$, $T = 2$, $K = 10$, $b = .1$ and a constant mesh width $h = .2969$. The left-hand graph was generated with $a = .2$, corresponding to a Péclet number less than 1 and is quite smooth, as expected. However, the right-hand graph was produced using $a = .001$, corresponding to a large Péclet number, and illustrates numerical instability typical of convection-domination.

8.2.6 Streamline diffusion FEM

As hinted at in the previous section, the problem of convection dominance can be solved by introducing additional diffusion terms. In the multidimensional case, these artificial diffusions must act in the direction of convection, thus giving the method the name 'Streamline Diffusion FEM' (SDFEM). The formulation of the method used here is taken from [11], where a steady state problem similar to the following is considered.

$$\begin{aligned} \mathcal{L} \equiv -\epsilon\Delta u + \mathbf{b}\cdot\nabla u = 0 \quad &\text{in } G \\ u(x) = 0 \quad &\text{on } \partial G \end{aligned} \tag{8.11}$$

The standard weak formulation of the above is

$$\text{find } u \in H_0^1(G) \text{ such that } B_c(u, v) = 0, \forall v \in H_0^1(G)$$

where

$$B_c(u, v) \equiv \int_G (\epsilon\nabla u \cdot \nabla v + (\mathbf{b}\cdot\nabla u)v)dx.$$

To correct for convection-dominance an extra term is added so that the discretised weak formulation becomes

$$\text{Find } u \in H_0^1(G) \text{ such that } B_c(u_h, v_h) + \sum_{\kappa\in\mathcal{T}} \tau_\kappa(\mathcal{L}u_h, \mathbf{b}\cdot\nabla v_h)_\kappa = 0, \forall v_h \in V_h^c$$

where $\mathcal{T}$ is a partition of G with disjoint open elements $\kappa \in \mathcal{T}$. In one dimension, this is equivalent to using test functions $v_h + \sum_\kappa \tau_\kappa (b(x)\partial_x v_h)_\kappa$, which produce artificial diffusion type terms of the form $\tau_\kappa(b(x)\partial_x u_h(x), b(x)\partial_x v_h(x))_\kappa$. In higher dimensions the extra terms are of the form $\tau_\kappa(\mathbf{b}\cdot\nabla u_h(x), \mathbf{b}\cdot\nabla v_h(x))_\kappa$, where the scalar product $\mathbf{b}\cdot\nabla v_h$ ensures that the artificial diffusion acts in the direction of convection.

The parameter τ_κ depends on the mesh width and is typically chosen as $\tau_\kappa = \tau_0 h_\kappa$. In the usual Galerkin discretisation used in this thesis, h_κ will be constant along each co-ordinate axis, thus reducing the number of parameters to the number of dimensions of the vector space. However, as described in [11], exact values for τ_0 cannot be deduced analytically. In general, each τ_0 should be large enough to eradicate any numerical instability in the solution, but not so large that they affect the Finite Element solution.

[27, Theorem 9.2] shows that the optimal convergence rate of FEM is retained with SDFEM.

8.2.7 SDFEM for convection-diffusion-Lévy PIDEs

If we restate (8.11) as

$$\mathcal{L} \equiv -\epsilon\Delta u + \mathbf{b}\cdot\nabla u + \overbrace{\int_{\mathbb{R}\setminus\{0\}} [u(x+z) - u(x) - x\partial_x u(x)]\nu(dz)}^{(a)} = 0 \quad \text{in } G$$
$$u(x) = 0 \quad \text{on } \partial G \tag{8.12}$$

we have a convection-diffusion-Lévy PIDE. In this case, the mesh Péclet number alone cannot give us an indication as to whether or not a FEM solution will be convection-dominated. To judge the influence of the Lévy term on the PDE we need to look at the *order* of the Lévy process. Note that (8.12)(a) can, in general, be written as

$$\int_{\mathbb{R}\setminus\{0\}} \frac{1}{|z|^{1+\alpha}} \left(u(x+z) - u(x) - x\partial_x u(x)\right) f(z)dz \tag{8.13}$$

where $f(z) = k(z)|z|^{1+\alpha}$ and k is the Lévy density of ν. $\alpha \in [0, 2)$ is known as the order of the Lévy process with measure ν. Given the form of the Lévy measure for the Variance Gamma process, it is easy to see that $\alpha = 0$. Compound Poisson processes also have order 0. The CGMY process with parameter $Y \in (0, 2)$, has order Y (see [12]) and other Lévy processes of infinite variation also have order greater than 0 (e.g. NIG process).

The form of (8.13) can be interpreted as a fractional derivate of order α. Similarily, the order of a differential operator $\partial^n, n \in \mathbb{N}$ is n[4]. Thus, the order of a diffusion operator is 2 and the order of a convection operator is 1.

[4]This can be seen by looking at the symbol of the operator ∂^n, which is $(i\xi)^n$, where ξ is the Fourier variable of integration. Similarily, the symbol of a fractional differential operator of order α is $|\xi|^\alpha$.

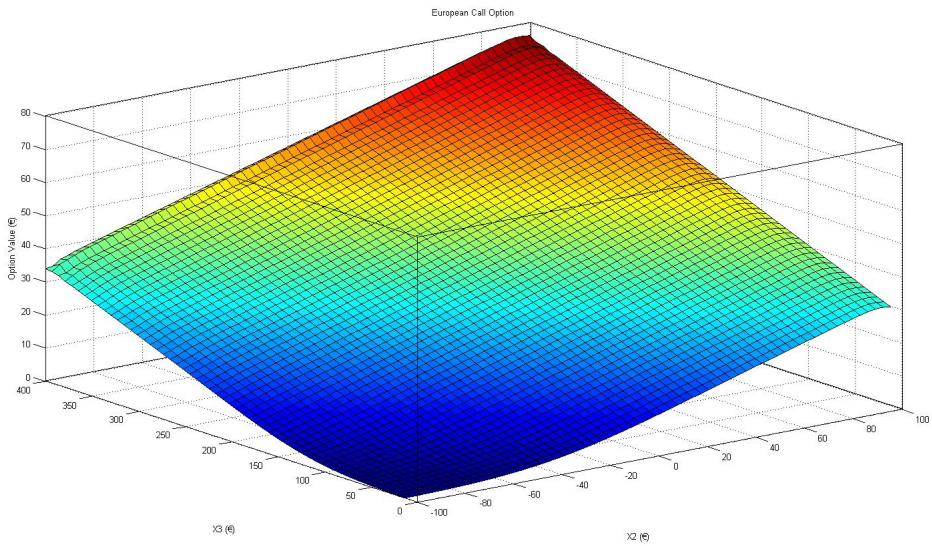

Figure 17: SDFEM solution for European call option value over the whole domain of computation relative to x_2 and x_3.

The closer the order of a Lévy process is to that of a diffusion, the more the Lévy process behaves like the diffusion and smoothes out the solution of the PIDE. Thus, if the order of the Lévy process is less than that of the convection term and the mesh Péclet number is greater than 1, the FEM solution of (8.12) will be convection dominated. In this case, the result of the previous section can be carried over and the convection domination can be compensated by the addition of an artifical diffusion.

8.2.8 SDFEM formulation of pricing problem

We now return to the original problem of a European call option on the electricity spot price. Given the size of the space domain G, realistic values for the mesh widths are at least 1 in each dimension. Moreover, the convection terms in the definition of the bilinear form (8.6) contain the coefficients $a_k x_k$, $k = 2, 3$, and so can have values of up to 400. Furthermore, the Lévy processes (VG and CP) acting along the coordinate axes e_2 and e_3 have order 0. For these reasons, it is clear that convection-dominance is the cause of the numerical instability in Figures 15. To correct this, we employ SDFEM.

As the convection terms only appear in the factors X^2 and X^3, we can ignore the X^1 factor in applying the streamline diffusion method, thus setting $\sigma_1 = 0$ in the PIDE. We can then rewrite the operator $\mathcal{A}$ and its associated bilinear

form as follows:

$$(\mathcal{A}f)(x) = -\mathbf{b}\cdot\nabla f + \int_{\mathbb{R}^2\setminus\{\mathbf{0}\}} f(x+z) - f(x)\nu(dz)$$

where $\mathbf{b} = [a_2x_2, a_3x_3]^T$ and $a(\cdot\,,\cdot) : H^1(\mathbb{R}^2)\times H^1(\mathbb{R}^2) \mapsto \mathbb{R}$ is defined by

$$a(\psi,\phi) = (\mathbf{b}\cdot\nabla\psi,\phi) + a^{\mathcal{J}}(\psi,\phi) \tag{8.14}$$

where

$$a^{\mathcal{J}}(\psi,\phi) = \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} \partial_z\psi(x+z)\partial_z\phi(x)\nu(dz)dx$$

In [26], the streamline diffusion method is applied with a bilinear form that also contains a Lévy component such as $a^{\mathcal{J}}$. We add artificial diffusion terms to (8.14) acting in the discretised space to get a modified bilinear form defined by:

$$a_\tau(u_h,v_h) = a(u_h,v_h) + \sum_{\kappa\in\mathcal{T}} \tau_\kappa[(\mathbf{b}\cdot\nabla u_h, \mathbf{b}\cdot\nabla v_h)_\kappa + a^{\mathcal{J}}(u_h, \mathbf{b}\cdot\nabla v_h)_\kappa] \tag{8.15}$$

Using [26, Remark 5.3], we can omit the term $a^{\mathcal{J}}(u_h, \mathbf{b}.\nabla v_h)_\kappa$ in (8.15) without affecting the stability of the method. Furthermore, using the fact that the mesh width is constant along each coordinate axis we can reduce the number of parameters τ_κ, so that (8.15) becomes:

$$\begin{aligned} a_\tau(u_h,v_h) = a(u_h,v_h) &+\tau_2[(a_2x_2\partial_{x_2}u_h, a_2x_2\partial_{x_2}v_h) + (a_3x_3\partial_{x_3}u_h, a_2x_2\partial_{x_2}v_h)] \\ &+\tau_3[(a_2x_2\partial_{x_2}u_h, a_3x_3\partial_{x_3}v_h) + (a_3x_3\partial_{x_3}u_h, a_3x_3\partial_{x_3}v_h)] \end{aligned}$$

$a(u_h,v_h)$ can be discretised as before and the additional 4 new terms adjust the form of the stiffness matrix so that it becomes

$$\begin{aligned} \mathbf{A}_{ij} &= a(b_{j_2}b_{j_3}, b_{i_2}b_{i_3}) \\ &\quad +\tau_2[\int a_2^2x_2^2 b'_{j_2}b_{j_3}b'_{i_2}b_{i_3}dx + \int a_3x_3b_{j_2}b'_{j_3}a_2x_2b'_{i_2}b_{i_3}dx] \\ &\quad +\tau_3[\int a_2x_2b'_{j_2}b_{j_3}a_3x_3b_{i_2}b'_{i_3}dx + \int a_3^2x_3^2b_{j_2}b'_{j_3}a_2x_2b_{i_2}b'_{i_3}dx] \\ &= \mathbf{B}_{ij}^{a_2x_2}\mathbf{M}_{ij}^3 + \mathbf{L}_{ij}^2\mathbf{M}_{ij}^3 + \mathbf{M}_{ij}^2\mathbf{B}_{ij}^{a_3x_3} + \mathbf{M}_{ij}^2\mathbf{L}_{ij}^3 \\ &\quad +\tau_2[\mathbf{S}_{ij}^{a_2^2x_2^2}\mathbf{M}_{ij}^3 - \mathbf{B}_{ij}^{a_2x_2}\mathbf{B}_{ij}^{a_3x_3}] + \tau_3[\mathbf{M}_{ij}^2\mathbf{S}_{ij}^{a_3^2x_3^2} - \mathbf{B}_{ij}^{a_2x_2}\mathbf{B}_{ij}^{a_3x_3}] \end{aligned}$$

where we have used the fact that $\int b_{j_k}b'_{i_k}dx = -\int b'_{j_k}b_{i_k}dx$, $k = 2,3$. Hence we can write down the stiffness matrix in dimensions 2 and 3 as follows:

$$\begin{aligned} \mathbf{A} &= (\tau_2\mathbf{S}^{a_2^2x_2^2} + \mathbf{B}^{a_2x_2} + \mathbf{L}^2)\bigotimes\mathbf{M}^3 - (\tau_2+\tau_3)\mathbf{B}^{a_2x_2}\bigotimes\mathbf{B}^{a_3x_3} \\ &\quad +\mathbf{M}^2\bigotimes(\tau_3\mathbf{S}^{a_3^2x_3^2} + \mathbf{B}^{a_3x_3} + \mathbf{L}^3). \end{aligned}$$

Here we can easily see that by setting $\tau_2 = \tau_3 = 0$, we recover the standard form of the stiffness matrix in 2 dimensions.

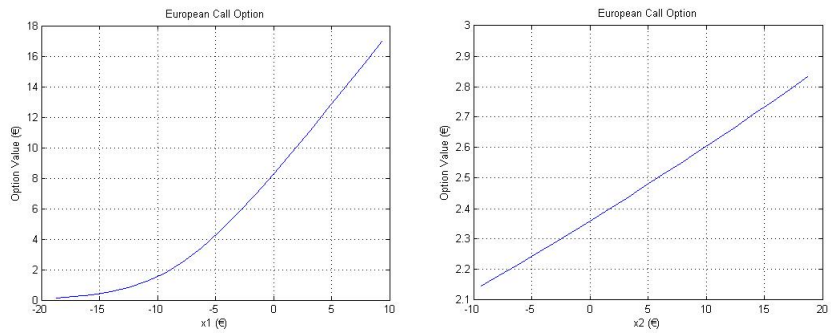

Figure 18: SDFEM solution for European call option value near current values of x_1 (left) and $x2$, with $K = 55$ and $T - t = 2$.

8.2.9 Numerical results

Using this stiffness matrix we solve the matrix problem (8.8) as before, but additionally, appropriate values for τ_2 and τ_3 must be chosen. We use the typical form $\tau_k = \delta_k h_k$ for each $k = 1, 2$ and chose the δ_k's so that the τ_k's are roughly the same. Values of $\delta_2 = .01$ and $\delta_3 = .002$ yield values for each τ_k of about .015, which is enough to eliminate any numerical instability without distorting the solution surface. Figure 17 shows the results over the whole domain of computation and we can clearly see an improvement over the surface in Figure 15.

We price a European call option on 07.10.09 with delivery on 09.10.09. The baseload price on the day of pricing was €55.54 and the seasonality function on the day of delivery had a value of €63.22. We assume a risk-free interest rate of 1% and a contractual strike price of €55, which implies $k(t, T) = -7.94$. Figure 19 shows the results with respect to two dimensions at a time. On 07.10.09, we deduce the values of the factors X_t^1, X_t^2 and X_t^3 given the deseasonalised spot prices, $\bar{S} = \{\bar{S}_s : s = t - 30, t - 29, \ldots, t\}$, as follows:

1. Set X_t^1 to the mean value of $\bar{S}_s, \forall s = t - 30, \ldots, t$.
2. Define $\hat{S}_t := \bar{S}_t - X_t^1$ and $L_{SD} := SD \times L$, where SD is the standard deviation of the whole deseasonalised price time series over the calibration period and L is the number of standard deviations used as the spike 'threshold' described in Section 6.3.
3. Set $X_t^3 = (\hat{S}_t - L_{SD}) \wedge 0$.
4. Set $X_t^2 = \hat{S}_t - X_t^3$.

Hence, we determine the values of the factors to be $X_t^1 = -9.54$, $X_t^2 = 5.68$ and $X_t^3 = 0$. The option value results for X^1 and X^2 around these values are shown in Figure 18 and by combining the option prices for the given values of X^1 and X^2, we calculate the current option price to be €4.21.

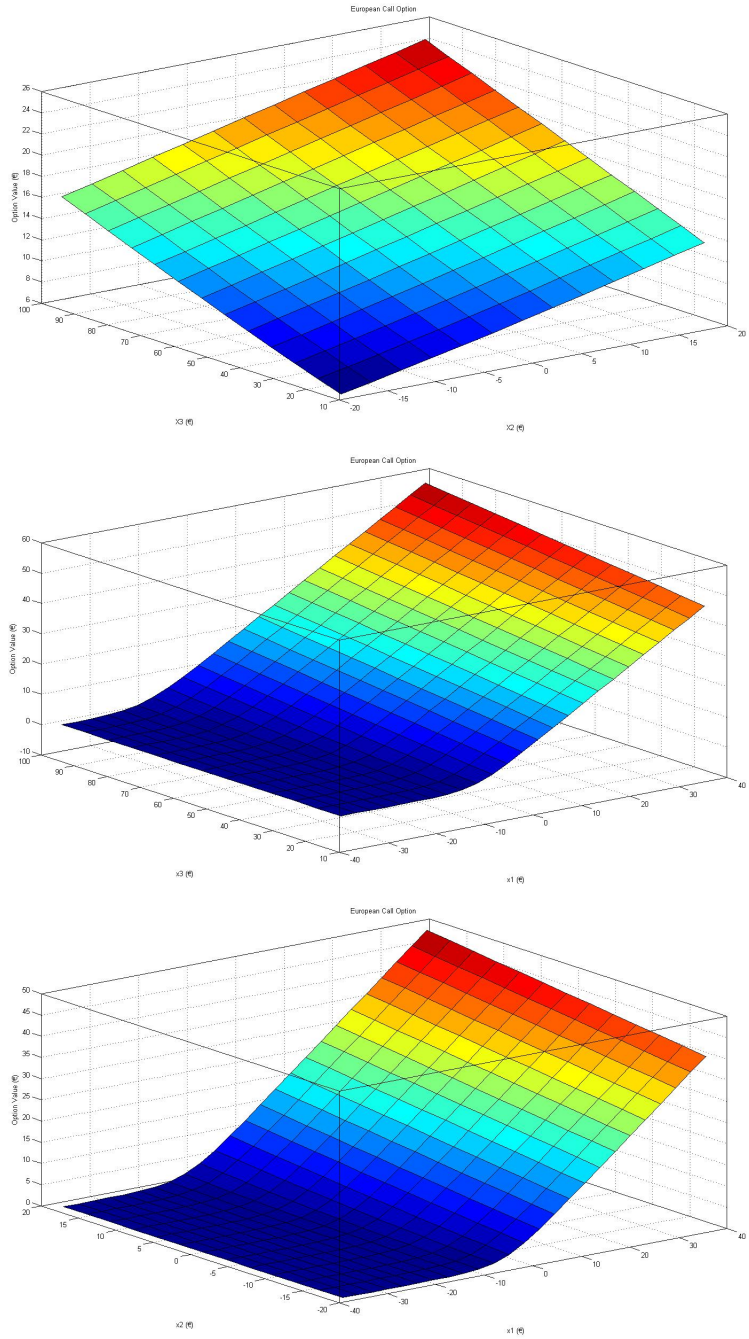

Figure 19: SDFEM solution for European call option value in an appropriate area of interest, with x_1 (top), $x2$ (middle) and x_3 (bottom) set to 0 and $k(t,T) = -7.94$ and $T - t = 2$.

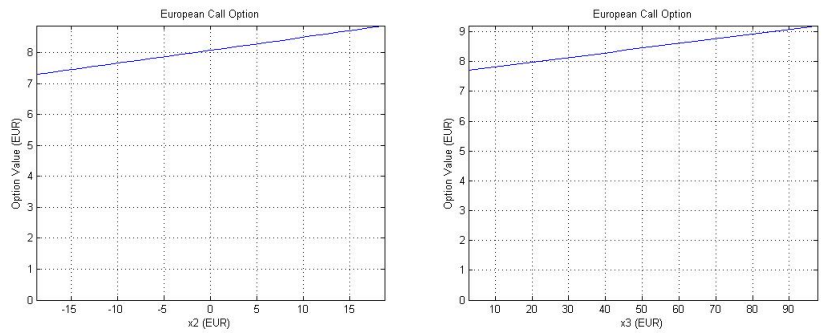

Figure 20: SDFEM solution for European call option value with $T - t = 4$.

Note that the option value with respect to X^1 has a shape similar to that observed for options in Black-Scholes markets, whereas due to strong mean reversion, the option value with respect to x_2 has a more linear form. With increasing time to delivery the option value function will get flatter and more constant for both x_2 and x_3, which is illustrated in Figure 20. Therefore, as time-to-delivery increases, option values will vary mainly with the value of the $X1$ factor.

8.2.10 Convergence

We show convergence of the discretised solution to the exact localised solution by picking a smooth hypothetical solution, $u(x)$, and computing

$$f(x) := (\mathcal{A}u)(x).$$

Knowing $f(x)$, we then find the finite element solution, $u_G(x)$, of the PIDE $\mathcal{A}u = f$, on a bounded domain, G, with an appropriate boundary condition. The error is calculated as

$$\|u_G(x) - u(x)\|_{L^\infty(G')},$$

where G' is an appropriate domain of interest, and is plotted against decreasing mesh widths. We do this for the factors X^1 and X^2 in turn as it is computationally less expensive than computing the error for both factors simultaneously. We choose the exact solutions as

$$u_1(x_1) = \cos(\frac{\pi x_1}{10}),\ x_1 \in (-10, 10),$$

$$u_2(x_2) = (x_2 - 1)(x_2),\ x_2 \in (0, 1),$$

thus implementing Neumann boundary conditions for u_1 and Dirichlet boundary conditions for u_2. The results are displayed in Figure 21, for the VG parameter values $C = 1$ and $M = G = 5$ and the diffusion parameter value $\sigma_1 = .5$. The graphs confirm that the error decreases with the optimal rate 2.

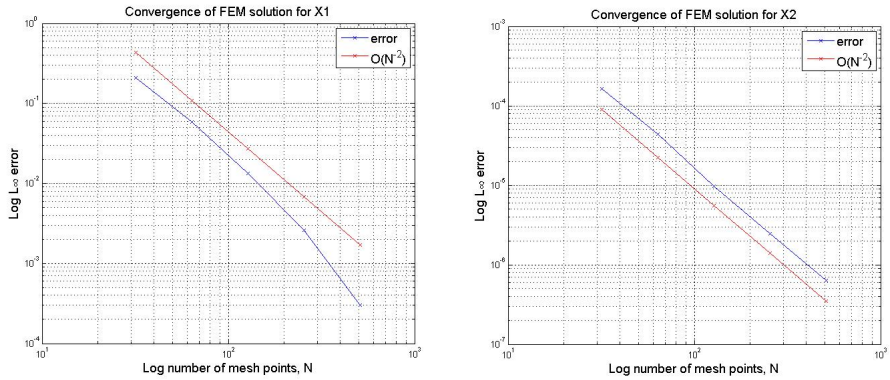

Figure 21: Convergence of FEM solutions for the factors X1 (left) and X2.

8.3 Options on futures

We lastly price European call options with maturity, T, on futures contracts with delivery period $[\tau_1, \tau_n]$. These options have value

$$v(t, S_t) = \mathbb{E}^{\mathbb{Q}}[e^{-r(T-t)}(F(T, \tau_1, \tau_n) - K)^+ | \mathcal{F}_t]$$

where $0 \leq T \leq \tau_1 \leq \tau_2$. Let $n := \tau_n - \tau_1$ and denote the days in the delivery period of the futures contracts as $\tau_i, i = 1, \ldots, n$. Using (7.2) and (7.4), the futures price can be written as

$$F(T, \tau_1, \tau_n) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}^{\mathbb{Q}} \left(S(\tau_i) | \mathcal{F}_T \right).$$

The $\mathbb{Q}$-expected value of the spot price is given in (7.6) and can be substituted into the above equation to give

$$\begin{aligned} F(T, \tau_1, \tau_n) = & \underbrace{\frac{1}{n} \sum_{i=1}^{n} \left(\Lambda(\tau_i) + \int_T^{\tau_i} \tilde{\lambda}(s) ds + (1 - e^{-a_3(\tau_i - T)}) \frac{\mu_P}{a_3} \right)}_{\tilde{\rho}_1} \\ & + \underbrace{\frac{1}{n} \sum_{i=1}^{n} \left((1 - e^{-a_2(\tau_i - T)}) \left[m + \frac{2(A + \hat{\lambda}_{\tau_i})}{a_2 \kappa (B^2 - (A + \hat{\lambda}_{\tau_i})^2)} \right] \right)}_{\hat{\rho}_1} + \sigma_1 W_T^{\mathbb{Q}} \\ & + \underbrace{\left(\frac{1}{n} \sum_{i=1}^{n} e^{-a_2(\tau_i - T)} \right)}_{\rho_2} X_T^2 + \underbrace{\left(\frac{1}{n} \sum_{i=1}^{n} e^{-a_3(\tau_i - T)} \right)}_{\rho_3} X_T^3 \end{aligned} \tag{8.16}$$

The terms $\rho_1 := \tilde{\rho_1} + \hat{\rho_1}$, ρ_2 and ρ_3 above only depend on τ_1 and τ_n and so are contant for any given futures contract. Using the notation of Section 8.2.1, we

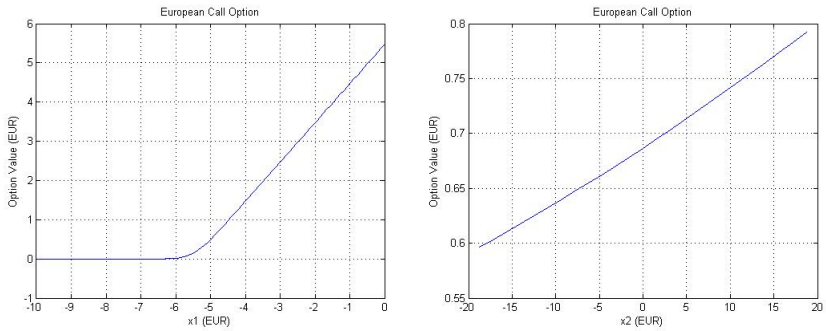

Figure 22: Option on a weekly futures contract with respect to x_1 (left) and x_2, with $k(t,T) = -5.47$, $T - t = 1$ and $\tau_1 - T = 5$.

can rewrite the value of the European call option on a futures contract as:

$$v(t, \hat{X}_t^1, \hat{X}_t^2, \hat{X}_t^3) = \mathbb{E}^{\mathbb{Q}} \left[e^{-r(T-t)} \left(\hat{X}_T^1 + \rho_2 \hat{X}_T^2 + \rho_3 \hat{X}_T^3 - k(t,T) \right)^+ | \mathcal{F}_t \right]$$

where

$$k(t,T) = K - \rho_1 - (m + \frac{\mu_2}{a_2})(1 - e^{-a_2(T-t)}) - \frac{\mu_3}{a_3}(1 - e^{-a_2(T-t)}).$$

As the above options price depends only on t and the factors $\hat{X}_t^1$, $\hat{X}_t^2$ and $\hat{X}_t^3$, the PIDE formulation for options on futures is the same as the PIDE for options on spot in Proposition 8.1, except that the terminal condition for options on futures is

$$v(T,x) = (x_1 + \rho_2 x_2 + \rho_3 x_3 - k(t,T))^+.$$

Therefore, we can price options on futures by reusing the discretised problem derived in Section 8.2.

For monthly futures, $T = \tau_1 - 4$ and for a 31 day delivery period, ρ_2 and ρ_3 are .0027 and .00075, respectively. Thus, the value of an option on a monthly futures contract is practically independent of $\hat{X}^2$ and $\hat{X}^3$. For weekly futures contracts, with $T = \tau_1 - 4$, ρ_2 and ρ_3 are .011 and .0032 respectively, so that there will be some dependence between the value of an option on a weekly futures contract and $\hat{X}^2$.

We consider an option on a futures contract for the week commencing 02.11.09 with contractual strike price €56 and maturity on 28.10.09. We value the option on 27.10.09, when the spot price was €57.01 and $k(t,T) = -5.47$. The results with respect to x_1 and x_2 are displayed in Figure 22. As expected, the option value is almost contant with respect to x_2 and varies only 20 cent over the domain of interest.

9 Conclusions and further work

9.1 Summary

The purpose of the spot model described in this thesis was twofold. Firstly, we described how deseasonalised prices revert to a mean level that is not constant but has a stochastic nature and this stochastic mean level was modelled as a diffusion in a separate factor. Secondly, we showed how a Brownian motion-driven OU process does not accurately fit day-to-day price variations and instead, we used a variance gamma driven process, which was better able to explain the leptokurtosis and heavy tails of the data. By combining these ideas in a three factor model, we showed how a regular market price of risk can be obtained in calibrating the model to the monthly futures market.

The risk-neutrally calibrated model could then be used to price monthly futures and European options on the spot and futures price. In pricing monthly futures, we illustrated how the model could be used to identify mis-pricing in the market. In pricing options on the spot price, we used the Streamline Diffusion FEM to show how option values depend on the various price factors close to maturity. We lastly showed how options on futures contracts can be priced by simply changing the initial condition of the spot option problem.

9.2 Error analysis

The pricing of derivatives described in Section 8 contains a number of sources of error and these fall into two separate categories.

Modelling error is the error caused by the approximations made in the physical and risk neutral calibrations as well as the error inherent in our modelling assumptions and choice of model. We have seen in Figure 13 that the difference between predicted and market prices are typically under 10%, but can be over 30% in some cases. Although it is not clear how much of this disparity is due to modelling error and how much is due to mispricing by the market, we suggest in the following subsection some ways to reduce this difference.

Numerical error is caused by approximating the exact solution to a pricing problem by a solution obtained with a numerical scheme such as FEM. In order for the numerical solution to converge to the exact solution, the scheme must be *stable* and *consistent* (see [25, Sections 3.5 and 3.6]). We saw in Section 8.2 how the streamline diffusion method was used to ensure that the scheme was stable. When a finite element scheme is convergent, its numerical solution contains errors due to the localisation of the problem to a bounded domain as well as the discretisation of the localised solution with finite elements. We showed in Figure 21 how the discretised solution converges to the localised solution with the optimal rate.

Figure 21 also shows that the discretisation error for 2^8 mesh points is of the order of 10^{-3} and, therefore, is considerably smaller than the modelling errors shown in Figure 13. For this reason, it might be tempting to think that numerical error is relatively unimportant. That is, that the care required to

ensure that there is no instability in the numerical scheme, or the computational expense required to apply FEM with a relatively fine mesh of 2^8 points, is unnecessary. However, if we solve a pricing problem numerically, in order to isolate the modelling error, the numerical error must be reduced so that it is almost negligible. It is only then that we can make accurate statements about the overall effectiveness of the model.

9.3 Further work

In addition to what is described in this thesis, the three factor model (2.2) could be further applied and improved in the following ways:

- A relatively simply algorithm was used in Section 7.3.1 to obtain the daily forward curve. It is possible that more regular results for the market prices of risk would be obtained if a more advanced and realistic algorithm were used.

- Figure 14 shows how the model can fail to predict the prices of some monthly contracts and, as suggested in Section 8.1, this could be due to deterministic information available to market participants that is not impounded in the spot price. Further research could investigate how such information could be encapsulated and used in the spot model to reduce the errors in predicted monthly futures prices shown in Figure 13.

- A disadvantage of additive models that contain a diffusion is that they allow prices to become persistently negative, even without considering negative spikes. In this regard, (2.2) could potentially be improved by remodelling $\Lambda(t) + X_t^1$ as $L(t)e^{X_t}$, where L is a multiplicative seasonality function and X is a Lévy process.

- The model was developed with natural gas markets in mind and initial investigations suggest that the same techniques can be used to calibrate the model to European spot and monthly futures gas markets. Furthermore, preliminary analysis shows that almost all of the dependence between gas and electricity is encapsulated in the X_1 factors, with a rank correlation of .5, whereas the other factors are almost independent with rank correlations of less than .1. This is consistent with results in [18] and suggests a convenient way to price the 'spark-spread' between electricity and gas.

A Finite Element Method

The Finite Element Method (FEM) is widely used in many branches of science and engineering to numerically solve partial differential equations (PDEs). In this thesis, we use FEM to solve the parabolic PDEs of derivative pricing problems, which, in general, have one temporal derivative and up to two spatial derivatives. In addition, the PDEs contain integrals, in which case they are referred to as partial integro-differential equations (PIDEs).

One of the main advantages of FEM is that it allows a higher rate of convergence than is achievable with Monte Carlo pricing, as well as permitting straightforward calculation of the 'Greeks'. Furthermore, unlike the Finite Difference Method, FEM permits optimal convergence under low smoothness assumptions on the payoff function and this is particularly relevant in financial applications where the payoff may not be differentiable (as with vanilla options) or even continuous (as with binary options).

We refer to [25] for an indepth treatment of FEM in the financial context and only give the briefest of outlines here.

A.1 PDE form of pricing problems

The link between the value of a contingent claim on an asset whose price is given by a stochastic process and the solution of a parabolic PDE is given by the Feynman-Kac formula which we now recall in the case of Brownian driven dynamics.

Theorem A.1. *Let $X = \{X_t : t \in J\}, J = (0,T)$ be the solution of the stochastic differential equation*

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t$$

where $b, \sigma : J \times \mathbb{R} \mapsto \mathbb{R}$ are Lipschitz continuous and grow at most linearily in x. Given a function $g \in L^2(\mathbb{R})$, the following are equivalent representations of $V \in C^{1,2}(J \times \mathbb{R}) \cap C^0(\overline{J} \times \mathbb{R})$ with bounded derivatives in $x \in \mathbb{R}$:

(i) V is a solution of

$$\begin{aligned} \overbrace{\partial_t V + \mathcal{A}V - rV}^{(a)} = 0 \qquad & \text{in } J \times \mathbb{R}, \\ V(T,x) = g(x) \quad & \text{in } \mathbb{R} \end{aligned} \tag{A.1}$$

where

$$(\mathcal{A}V)(t,x) = \frac{1}{2}\sigma^2(t,x)\partial_{xx}V(t,x) + b(t,x)\partial_x V(t,x).$$

(ii)

$$V(t,x) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r(X_s)ds} g(X_T)|X_t = x\right]$$

where $r : \mathbb{R} \mapsto \mathbb{R}$ is bounded and continuous and $\mathbb{Q}$ is a risk neutral measure.

Proof. See [25, Theorem 4.1.4]. □

The second representation is a $\mathbb{Q}$- or risk neutral expectation and thus defines the fair value of a derivative with payoff g. The idea of the proof is to applying Itô's Lemma to $M_t = e^{-\int_t^T r(X_s)ds} f(t, X_t)$, where $f \in C^{1,2}(\mathbb{R} \times \mathbb{R})$ has bounded derivatives in x. The result is a semi-martingale whose drift is given by (A.1)(a). M_t can be thought of as the value of a derivative, but in order for it to be the fair value, it must be a martingale and therefore, its drift must be identically 0.

Thus, we can see that the form of $\mathcal{A}$, termed the *infinitesimal generator* of X, depends completely on the dynamics of X through an application of Itô's Lemma. Given Itô's Lemma for Lévy processes (3.4), it is not difficult to see how Lévy dynamics for X result in $\mathcal{A}$ containing an integral term. Furthermore, using Itô's Lemma for multidimensional processes, the above theorem can easily be generalised to the case where X is a multidimensional stochastic process.

A.2 Variational Formulation

The next step is to remove the second order derivative of the pricing PDE via integration-by-parts, yielding the *weak* or *variational* form of the PDE. In this light we need the following definitions.

Definition A.2. *Let $G = (a, b) \in \mathbb{R}$ be an open, possibly unbounded domain and $u \in C^1(\overline{G})$. Then the linear functional $D^n u, n \in \mathbb{N}$, is a weak derivative of u if*

$$\int_G D^n u \varphi dx = (-1) \int_G u D^n \varphi dx, \quad \forall \varphi \in C_0^n(\overline{G}).$$

Definition A.3. *For $m \in \mathbb{N}$, define the Hilbert space*

$$H^m(G) = \{u \in L^2(G) : D^n u \in L^2(G), n \leq m\}$$

with the inner product

$$(u, v)_{H^m(G)} = \sum_{n=0}^{m} (D^n u, D^n v)_{L^2(G)}$$

Thus, we consider the derivatives in (A.1) to be weak derivatives and consequently work in the space $H^m(G), m \in \mathbb{N}$, which is an example of a *Sobelov* space. Furthermore, $H^1(G)$ is the closure of $C^\infty(G)$ in the H^1 norm, which will be an important property later when we need to densely embed a discretised space in the solution space. If $\mathcal{A}$ is generated by Lévy-driven dynamics, we need Sobelov spaces of *fractional order*, where $m > 0$. However, in this thesis we only consider Lévy processes in combination with a diffusion, so we only deal with the case $m = 1$.

As alluded to above, we multiply the (A.1) by a generic 'test' function and integrate the second order term by parts to get, $\forall v \in H^1(G)$,

$$\int_G \frac{\sigma^2}{2} \partial_{xx} u(t,x) v(x) dx = \underbrace{\left[\frac{\sigma^2(t,x)}{2} \partial_x u(t,x) v(x) \right]_{x=a}^{x=b}}_{(a)} - \int_G \partial_x u(t,x) \left(\frac{\sigma^2(t,x)}{2} \partial_x v(x) + \sigma(t,x) \partial_x \sigma(t,x) v(x) \right) dx. \tag{A.2}$$

To simplify the weak form of the PDE, we wish to assign a value (usually 0) to (A.2)(a) by choosing a suitable condition on u or v at the boundary of

G. A *Dirichlet* boundary condition assumes $u, v = 0$ at ∂G, which means the Sobelov space that we work in becomes $H_0^1(G)$. Alternatively, we can assume $\partial_x u(x,t) = 0$ at ∂G and this is termed the *Neumann* boundary condition.

Before converting (A.1) to its weak formulation, we change to time-to-maturity ($t \to T - t$) to obtain a forward parabolic problem. We then state the variational form as follows:

$$\text{Find } u \in L^2(J; H_0^1(G)) \cap H^1(J; L^2(G)) \text{ such that}$$

$$(\partial_t u, v) + a(u, v) = 0, \forall v \in H_0^1(G) \text{ a.e. in } J, \quad u(0,x) = g(x)$$

and the bilinear form $a(\cdot\,,\cdot\,) : H_0^1(G) \times H_0^1(G) \mapsto \mathbb{R}$ is defined as

$$a(\varphi, \phi) := \left(\frac{\sigma^2}{2}\varphi', \phi'\right) + (\{\sigma\sigma' - b\}\varphi', \phi) + (r\varphi, \phi)$$

In order to guarantee the existence of a weak solution of the above problem, the bilinear form $a(\cdot,\cdot)$ must be shown to be continuous and must satisfy a Gårding inequality. We refer to [25, Theorem 3.2.2] for more details.

A.3 Localisation

If the PDE is defined on an unbounded domain G (e.g. often the PDE is defined on the whole of $\mathbb{R}^d$), this may place unrealistic integrability constraints on the payoff function g. In this case, the domain is *localised* to $(-R, R), R > 0$ and we solve a modified problem whose solution is u_R. In this case, we must additionally show that $u_R \to u$ as R increases and then R can be chosen suitably large enough so that the difference between u and u_R is negligible.

A.4 Discretisation

We look to approximate the solution to the variational formulation with piecewise polynomial functions. This allows us to discretise the variational formulation in space by approximating $u(t,x)$ with $u_N(t,x) \in V_N$, where V_N is a subset of $H^1(G)$ and is of finite dimension, N. When $G = (a,b)$, the domain can be partitioned with an arbitrary mesh

$$\mathcal{T} = \{a = x_0 < x_1 < \ldots < x_{N+1} = b\}$$

but for the purposes of this thesis we can assume that the mesh points are equidistant. For $1 \le i \le N$ we define the *hat functions* as

$$b_i(x) = \begin{cases} (x - x_{i-1})/h & \text{if } x \in (x_{i-1}, x_i] \\ (x_{i+1} - x)/h & \text{if } x \in (x_i, x_{i+1}] \\ 0 & \text{otherwise} \end{cases}$$

where $h = (b-a)/(N+1)$ is the mesh width. We can then choose $V_N = S_{\mathcal{T}} \cap H_0^1(G)$ where $S_{\mathcal{T}}$ is spanned by the hat functions

$$S_{\mathcal{T}} = \text{span}\{b_i(x) : i = 0, \ldots, N+1\}$$

Consequently, $u_N(t,x)$ will be piecewise linear in x. In the case where $G \subset \mathbb{R}^d$, G will be partitioned along each co-ordinate axis and $S_{\mathcal{T}}$ will be given by a product of hat functions.

The semi-discretised form of the variational formulation now reads

$$\text{Find } u \in C^1(J; V_N) \text{ such that for } t \in J$$

$$\mathbf{M}\dot{\underline{u}}_N(t) + \mathbf{A}\underline{u}_N(t) = 0, \quad \underline{u}_N(0) = \underline{u}_0$$

where $\mathbf{M}_{ij} = (b_j, b_i)$ and $\mathbf{A}_{ij} = a(b_j, b_i)$.

We discretise in time with the finite difference 'θ-scheme' where for $m = 1, \ldots, M$, we have the equidistant time steps $t_0 = 0$, $t_m = mk$, $k = T/M$ and for $\theta \in [0,1]$,

$$\dot{\underline{u}}_N^{m+\theta}(t) = (\underline{u}_N^{m+1} - \underline{u}_N^m)/k$$

$$\underline{u}_N^{m+\theta}(t) = \theta \underline{u}_N^{m+1} - (1-\theta)\underline{u}_N^m.$$

Finally, we state the fully discretised matrix problem as

$$\text{Find } u^m \in V_N \text{ such that for } m = 0, \ldots, M$$

$$(\mathbf{M} + k\theta\mathbf{A})\underline{u}_N^{m+1} = (\mathbf{M} - k(1-\theta)\mathbf{A})\underline{u}_N^m, \quad \underline{u}_N^0 = \underline{u}_0$$

A.5 Stability and Convergence

In [25, Section 3.5], it is shown that for $\theta \in [\frac{1}{2}, 1]$, the θ-scheme is unconditionally stable, whereas for $\theta \in [0, \frac{1}{2})$, k must be sufficiently small.

Furthermore, it is shown in Section 3.6.2 of the same reference that for $\theta = \frac{1}{2}$, optimal second order convergence in both time and space can be achieved in the L^2-norm i.e.

$$\|u^M - u_N^M\|_{L^2(G)} = \mathcal{O}(h^2 + k^2).$$

References

[1] Applebaum, D. (2009). *Levy Processes and Stochastic Calculus.* Second Edition. Cambridge University Press.

[2] Barlow, M., Gusev, Y. and Lai, M. (2004). Calibration of multifactor models in electricity markets. *Intern. J. Theor. Appl. Finance*, **7**(2), pp. 101-120.

[3] Benth, F.E., Ekeland, L., Hauge, R. and Nielsen, B.F. (2003). On arbitrage-free pricing of forward contracts in energy markets. *Appl. Math. Finance*, **10**, 4, pp. 325-336.

[4] Benth, F.E., Kallsen, J. and Meyer-Brandis, T. (2007). A non-Gaussian Ornstein-Uhlenbeck process for electricity spot price modelling and derivatives pricing. *Appl. Math. Finance*, **14**(2), pp. 153-169.

[5] Benth, F.E. and Meyer-Brandis, T. (2009). The information premium for non-storable commodities. *J. Energy Markets*, **2**(3), pp. 111-140.

[6] Benth, F.E., Šaltytė-Benth, J. (2003). The Normal Inverse Gaussian distribution and spot price modelling in energy markets. *Intern. J. Theor. Appl. Finance*, **7**(2), pp. 177-192.

[7] Benth, F.E., Šaltytė-Benth, J. and Koekebakker, S. (2008). *Stochastic Modelling of Electricity and Related Markets.*

[8] Bjørk, T. (1998). *Arbitrage Theory in Continuous Time.* Oxford University Press.

[9] Branger, N., Reichmann, O. and Wobben, M. (2009). Pricing electricity derivatives. *Working paper*, Westfälische Wilhelms-Universität Münster, Germany World Scientific.

[10] Brigo, D. and Mercurio, F. (2001). *Interest Rate Models. Theory and Practice.* Springer Verlag.

[11] Cangiani, A., Georgoulis, E.H. and Jensen, M. (2006). Continuous and Discontinuous Finite Element Methods for Convection-Diffsion Problems: a Comparison. *International Conference on Boundary and Interior Layers*, Göttingen.

[12] Carr, P.P., Geman, H., Madan, D.P. and Yor, M. (2002). The Fine Structure of Asset Returns: An Empirical Investigation. *J. Business*, **75**(2), pp. 305-332.

[13] Cartea, A. and Figueroa, M.G. (2005). Pricing in electricity markets: a mean reverting jump diffusion model with seasonality, *Appl. Math. Finance*, **12**(4), pp. 313-335.

[14] Clewlow, L. and Strickland, C. (1999). Valuing energy options in a one factor model fitted to forward prices. *Working paper*, School of Finance and Economics, Technical University of Sydney.

[15] Clewlow, L. and Strickland, C. (2000). *Energy Derivatives: Pricing and Risk Management.* Lacima Publications.

[16] Cont, R., Tankov, P. (2003). *Financial Modelling with Jump Processes.* Chapman and Hall / CRC Press.

[17] De Jong, C. (2006). The nature of power spikes: a regime-switch approach. *Studies in Nonlinear Dynamics and Econometrics*, **10**(3), Article 3.

[18] De Jong, C. and Schneider, S. (2009). Cointegration between Gas and Power Spot Prices. *J. Energy Markets*, forthcoming.

[19] Donea, J. and Heurta, A. (2003). *Finite Element Methods for Flow Problems.* Wiley.

[20] Doucet, A., de Freitas, N. and Gordon, N. (2001). *Sequential Monte Carlo Methods in Practice.* Springer Verlag.

[21] Evans, L.C. (1998). *Partial Differential Equations.* American Mathematical Society.

[22] Eydeland, A. and Wolyniec, K. (2002). *Energy and Power Risk Management: New Developments in modelling, Pricing and Hedging.* Wiley.

[23] Geman, H. and Roncoroni, A. (2006). Understanding the fine structure of electricity prices. *J. Business*, **79**(3), pp. 1125-1261.

[24] Green, R. and Nossman, M. (2007). A jump diffusion stochastic volatility model for the electricity spot price. *Working paper*, Lund University, Sweden.

[25] Hilber, N., Schwab, C. and Winter, C. (2009). Computational Methods for Quantitative Finance. Finite Element and Finite Difference Methods. *Lecture notes*, ETH Zürich.

[26] Hilber, N. (2009). Stabilized Wavelet Methods for Option Pricing in High Dimensional Stochastic Volatility Models. *PhD. thesis*, ETH Zürich.

[27] Johnson, C. (1988). *Numerical Solution of Partial Differential Equations by the Finite Element Method.* Cambridge University Press.

[28] Kiesel, R., Schindlmyer, G. and Börger, R. (2007). A two-factor model for the electricity forward market. *Preprint*, University of Ulm, Germany.

[29] Lucia, J. and Schwartz, E.S. (2002). Electricity prices and power derivatives: evidence from the Nordic Power Exchange. *Rev. Derivatives Res.*, **5**(1), pp. 5-50.

[30] Madan, D.B, Carr, P.P. and Chang, E.C. (1998). The Variance Gamma Process and Option Pricing *European Finance Rev.*, **2**, pp. 79-105.

[31] Madan, D.B, Seneta, S. (1990). The variance gamma (VG) model for share market returns. *J. Business*, **63**(4), pp. 511-524.

[32] Meyer-Brandis, T., and Tankov, P. (2007). Multi factor jump-diffusion models of electricity prices. *Working paper*, Université Paris VII, France.

[33] McNeil, A. and Saladin, T. (1997). The peaks over thresholds method for estimating high quantiles of loss distributions. *Proceedings of XXVIIth International ASTIN Colloquium*, Cairns, Australia, pp. 23 - 43.

[34] Sato, K.I. (1999). *Levy Processes and Infinite Divisibility.* Cambridge University Press.

[35] Seifert, J. and Uhrig-Homburg, M. (2007) Modelling jumps in electricity prices: theory and empirical evidence *Rev. Derivatives Res.*, **10**(1), pp. 59-85.

[36] Seneta, E. (2004). Fitting the Variance Gamma process to financial data. *J. Appl. Probability*, **41**A, pp. 177-187.

[37] Shreve, S.E. (2004). *Stochastic Calculus for Finance 2.* Springer Science+Business Media.