Stochastic models with random parameters for financial markets

Suren Islyaev

A Thesis submitted for the degree of Doctor of Philosophy

Department of Mathematical Sciences
Brunel University

September 2014

Abstract

The aim of this thesis is a development of a new class of financial models with random parameters, which are computationally efficient and have the same level of performance as existing ones. In particular, this research is threefold.

I have studied the evolution of storable commodity and commodity futures prices in time using a new random parameter model coupled with a Kalman filter. Such a combination allows one to forecast arbitrage-free futures prices and commodity spot prices one step ahead.

Another direction of my research is a new volatility model, where the volatility is a random variable. The main advantage of this model is high calibration speed compared to the existing stochastic volatility models such as the Bates model or the Heston model. However, the performance of the new model is comparable to the latter. Comprehensive numerical studies demonstrate that the new model is a very competitive alternative to the Heston or the Bates model in terms of accuracy of matching option prices or computing hedging parameters.

Finally, a new futures pricing model for electricity futures prices was developed. The new model has a random volatility parameter in its underlying process. The new model has less parameters, as compared to two-factor models for electricity commodity pricing with and without jumps. Numerical experiments with real data illustrate that it is quite competitive with the existing two-factor models in terms of pricing one step ahead futures prices, while being far simpler to calibrate. Further, a new heuristic for calibrating two-factor models was proposed. The new calibration procedure has two stages, offline and online. The offline stage calibrates parameters under a physical measure, while the online stage is used to calibrate the risk-neutrality parameters on each iteration of the particle filter. A particle filter was used to estimate the values of the underlying stochastic processes and to forecast futures prices one step ahead.

The contributory material from two chapters of this thesis have been submitted to peer reviewed journals in terms of two papers:

- Chapter 4: *"A fast calibrating volatility model"* has been submitted to the European Journal of Operational Research.

- Chapter 5: *"Electricity futures price models : calibration and forecasting"* has been submitted to the European Journal of Operational Research.

Acknowledgements

I would like to thank my supervisor Dr. Paresh Date for his support and inspiration. It was great pleasure working with a high level professional researcher.

This thesis would not been possible without support and belief from my family. In particular, I would like to thank my mother Margarita Islyaeva for making the dream come true. My grandfather Karekin Tutelian, who taught me how to manage myself and be on top of the things.

I would like to express my deepest gratitude for my friends and past colleagues from around the world, especially Noemi Peluso, Layal Hakim, Nina Grishina, Emilia Apergi, Konrad Hoppe, Emma Haddi, Anggi Baskara, Vitaliy Rickiy, Alexander Hristolyubov, Farukh Mukhamedov, Mikael Sarkissyan, for making my life beautiful.

Contents

Chapter 1

Introduction

1.1 A brief history of financial modelling

The modern economy has narrowed the boundaries between national markets for financial securities. There is an enormous amount of data available on financial transactions at regulated exchanges, including historical price changes, transaction sizes and trading volumes. Professional investors use this information to model the behaviour of asset prices and devise possible ways to optimise the return on investment or reduce the exposure to the risk of investing in securities with an uncertain or unknown future return.

This thesis looks at asset price behaviour models based on stochastic processes. The contribution of the thesis lies in building models for certain asset pricing applications which are computationally simpler than the existing benchmark ones, but yield a comparable accuracy.

Louis Bachelier [8] was the first to set a price of an asset to be a stochastic process following an arithmetic Brownian motion. The unique assumption allowed the development of a new approach for pricing market contracts. However, the idea was met with scepticism from scientists. With the support of Henri Poincaré, this idea remained alive, but received a little attention in the literature for almost half a century. The development of probability theory and stochastic calculus allowed the construction of more efficient and complex theories about market behaviour.

Investing in only one asset leaves an investor exposed to risk of high losses. To avoid this limitation, Markowitz [57] developed a theory of portfolio optimisation, based on the historical performance of an asset, which was awarded

with a Nobel prize in 1990. This theory has a huge impact on the development in stochastic finance. It uses the fact that, in a portfolio of a bank account, an option and the underlying asset, one can choose weights of the cash and the asset holding at each time, such that the resulting portfolio eliminates risk[58]. Henceforth, the fact that any risk-free portfolio earns a unique risk-free rate of return allows us to price the option in terms of a solution of a partial differential equation (PDE). The solution to the PDE for 'vanilla' put and call options was proposed in [13]. Despite Bachelier's assumption that, the underlying process is adopted by an arithmetical Brownian motion, Black and Scholes specified an asset price to follow a geometrical Brownian motion (GBM). With a positive drift, this model ensures that the company net worth grows over time on average. An extension to the Black-Scholes model was introduced by Merton [60] in 1973. It was concerned about very quick price changes (jumps), which were modelled with a compound Poisson process. However, this approach doesn't have a closed form solution for options. There is an open question of whether there are real jumps on the market and the validity of using the compound Poisson process to model them.

Further study on stochastic models allowed the development of stochastic interest rate models. The first major research was proposed by Vasicec in [77]. A short-rate process was treated as a mean-reversion stochastic process, which is based on an 'Ornstein-Ulenbeck'[1] process. Such an approach is also popular for currency option models because of the nature of currency price behaviour. The weakness of the approach is that this model might produce negative interest rates. Later, Cox, Ingersoll and Ross (CIR) [21] assumed that instantaneous interest rate has a stationary gamma distribution. A CIR process is a square root process, which eliminates the possibility of negative values for interest rates.

All the models discussed so far, viz Black-Scholes and Merton models for asset prices and Vasicec and CIR models for interest rates, assume that volatility has no separate source of uncertainty, other than the asset price or the interest rate itself. These models are called local volatility models[2], with a constant volatility model such as the Vasicec model being a special case. However, local

[1]This process is popular in physics, which usually represents movement of a spring released after a load.

[2]Local volatility means that a volatility is a function of the asset price process and time, i.e. its value at a given time is completely determined by the value of the asset price process at the given time.

volatility models do not explain certain features of the observed market behaviour such as a volatility smile and smirk, i.e. the variation in the volatility implied by option prices with the strike price, at the same expiration date. We will re-visit this issue in chapter 4.

The solution to this problem has been suggested by Wiggins [80] with the assumption that volatility is not constant and can be represented as stochastic process. Hull and White [49] assumed that, in a stock return model stochastic volatility can itself have the form of GBM, which is uncorrelated with the interest rate process. The advantage of this model is the availability of an explicit solution for the option prices in terms of infinite series. Despite having stochastic volatility in this model, one more question for this model was still unanswered: Is the process for volatility chosen right?. Since both processes,*viz.*, asset return and stochastic volatility are GBM, it means that, on average they will grow or decline infinitely. This problem was solved by Stein [74]. He used an OU process for stochastic volatility, which allowed the expected value of the volatility to converge to a pre-defined long run mean value. However, a mean-reverting process is not bounded from below and can yield negative values. Consequently, new assumptions were made about the process driving stochastic volatility. Heston [44] substituted the volatility process in Stein's model with a CIR process. As a result, the possibility of negative values for volatility was eliminated with the square-root term. In addition, a semi-closed form solution was developed through characteristic functions for option prices.

We will look at option pricing in far more detail in chapters 2 and 4. Since we will also consider pricing of certain commodity securities in this thesis, we will discuss commodity markets next.

1.2 Commodity market

A brief introduction to modern commodity price models is provided here. The definition of commodities and commodity markets came from economic studies. According to Marx's idea [59], commodity prices can be evaluated using labour theory of value, with respect to discrete parameters such as labour, rent and profit. This idea was criticised in the works of David Ricardo [69] and Adam Smith [72], where the supply-demand relationship was established to estimate the commodity prices. This approach was later adopted by Mackey [55], where

the variation of the market price of a commodity was presented as a balance of supply and demand. Hence, the most complicated part of this model is that the supply function is stochastic, due to the dependence on the supply schedule.

Understanding commodity markets involves understanding forward rate curves and futures contracts written on commodities. In 1931, Hotelling [46] criticised the overuse of exhaustible resources and the economic situation around them. He introduced a mathematical approach for evaluating exhaustible resources. His hypothesis states that an owner of a commodity would always like to have a maximum profit from selling it and if the interest rate is constant, then the present value of a unit of profit at a future time will be proportional to a negative exponential discount rate. Moreover, he defined the relationship between current commodity price and its price in the future. In 1949, Working [81] introduced an inter-temporal price relation, which is now referred to as a *forward curve*. He analysed the relationship between the shape of the forward curve and provided an explanation of how it depends on the price of commodity storage.

The development of the probability theory and stochastic modelling pushed the development of models for forecasting market variables. Hubbert [47], in 1956, introduced an approach to forecast oil and gas production in the United States and worldwide. He derived a model to predict the dates when the production of these commodities will achieve their peaks. Samuelson in [70] studied the stochastic behaviour of commodity prices. The idea of existence of a free competitive market[3] was criticized in [70]. In addition, a hypothesis was proposed that commodity prices change over time by performing a random walk with no predictable bias. Furthermore, Samuelson presented the fundamental evidence that futures with longer time to maturity have lower volatilities than futures with a short time to maturity. Later this phenomenon was called the 'Samuelson effect'.

Fischer Black in [12] provided a comprehensive study of commodity contracts. He clarified the meaning of futures and forward contracts and introduced his futures option pricing model. An assumption was made that the commodity price is a price at which it can be bought or sold for immediate delivery. In his idea, the commodity price tends to follow a seasonal pattern, e.g., for agricultural commodities, the price before harvest is higher than after

[3]In this context, a free competitive market means that there is a buyer for every seller.

a harvest. Moreover, commodity patterns don't always provide opportunities for profit. For example, a spot price on an agricultural commodity can rise constantly at any rate below the storage cost per unit time for the commodity, without raising a profit opportunity.

Modelling the commodity price behaviour involves understanding a variety of factors such as the evaluation of natural resource investment projects, and this is the difference with other type of assets. In comparison to stock prices, which represent the performance of the company over time, commodities represent a full business cycle associated with them (e.g. investment, production, weather effects, etc). One of the first works on the stochastic control of commodity projects was done by Brennan and Schwartz [14]. They applied techniques of time arbitrage and stochastic control on mining projects. The idea was to find a mathematical representation of a self-financing portfolio with the absence of an arbitrage opportunity. In addition, they introduced interest rate as a deterministic function of time and modelled a portfolio as a problem of stochastic control. This research provided an example of an optimal decision making model for investors and project managers.

Fama and French [34] focused on two theories in futures pricing. The first one was based on the theory of storage, which explains the difference between contemporaneous futures and spot prices according to the changes in interest rates, warehousing costs and convenience yields. The second one is based on two components *viz.*, the expected premium and the expected value of the spot price at maturity time. Their study focused on the behaviour of commodity futures prices. They provided an analysis of the convenience yield and a method to determine whether it was caused by the risk-premium or the storage cost theory.

One of the first stochastic models for commodity pricing was introduced by Gibson and Schwartz [39]. The two-factor model was presented for pricing financial and real assets contingent on the price of oil and short term valuation of futures prices.

An additional study on the behaviour of commodity prices was done by Deaton and Laroque [25]. They applied the standard rational expectations competitive storage model to the real market using yearly data. This was made to compare the prices of commodities with and without storage costs. As a result, differences were reported in the probability density and the skewness of commodity prices. They also provided the evidence that empirical prices are

auto-correlated yearly. This model explains whether price spikes will appear or not during a bad harvest season and how they depend on storage.

Further discussion on commodity and futures price modelling will continue in chapters 3 and 5.

1.3 Foreword

Even though financial models have become more sophisticated with time, many commonly used models are often quite complicated to calibrate and to use in pricing or forecasting. In this work, we aim to to develop new financial models for pricing and forecasting which are simpler to calibrate and to use than the existing benchmark models, while still delivering an acceptable accuracy. We will focus on models for two different purposes: index option pricing and energy commodity futures pricing. In the latter area, we look at pricing storable (oil, gasoline and natural gas) as well as non-storable (electricity) futures contracts separately. Simplification of the models is achieved through replacing complex stochastic processes with random model parameters and modifying the pricing formulae accordingly. It is shown that the newly developed models yield comparable pricing and forecasting performance on real financial data as the relevant benchmark models, while being computationally easier to calibrate and use. Moreover, we use filtering algorithms for parameter estimation and introduce new heuristics to calibrate the models for electricity futures pricing.

The research in this thesis has achieved the following goals:

- We have proposed new financial asset pricing models (specifically, for index option pricing and commodity pricing).
- We have demonstrated through extensive numerical experiments that the new models have explanatory power comparable to existing benchmark models. These new models are computationally far simpler to calibrate and to use in comparison to the existing benchmark models.
- Further, the new heuristic procedures have been developed to simplify and speed up the calibration procedure for these models.

Subsequent chapters are organised as follows.

- Chapter 2 outlines the necessary mathematical and financial preliminary information. The first section is focused on a futures pricing framework and its extension with jump processes. The next section discusses an option pricing framework. The last section provides the preliminary information on the model calibration techniques used in this thesis.

- Chapter 3 discusses the development of a one-factor model with a random parameter for energy commodity price behaviour. My idea is to develop a model with one factor, i.e., with one stochastic process and *one random variable*, which mimics the properties of a model with two factors i.e. with two (non-trivial) stochastic processes. Two-factor models were taken as benchmark models against which the new one-factor models were tested. A simple version of a seasonality function is used for both the one-factor and the two-factor models. This chapter combines theory and practice by carrying out a variety of numerical experiments and benchmarks across the discussed models.

- In chapter 4, we define a new class of random volatility models for option pricing. The new volatility model assumes that volatility, rather than being a stochastic process, is a function of time with a random parameter. The distribution of the volatility is assumed to be log-normal, which can be rigorously justified in a classical stochastic process framework. However, the model class allows us to use almost any distribution with a few restrictions, such as an existence of all the moments. An option price formula is given in terms of an infinite series, which allows for an easy approximation. The results show that the model can be used for pricing European options, path dependant options and for computing hedging parameters. Extensive numerical experiments compare the calibration effort and the pricing accuracy of the new model with those of the Heston model and the Bates model.

- Chapter 5 discusses a new random volatility electricity price model. The new random volatility model is based on assumption that volatility is a random number. We provide an approximate expression for the electricity futures price for this model. In addition, we introduce a new and simplified offline calibration procedure for the two-factor models and the online calibration stage for all the models. Moreover, this chapter is con-

cerned with an empirical investigation into the use of jump-risk premium, which is used to compensate for the risk due to jumps in a compound Poisson process. This approach to account for the jump-risk was recently developed by [16]. A particle filter was used for a one step ahead futures price forecasting in numerical experiments.

- A summary of the contributions of this thesis is provided in chapter 6, along with future research directions.
- The plots regarding evolution of certain variables (noise variances and risk premium) and the expressions for the moments for a two-factor processes with jumps for chapter 5 are provided in Appendix A.
- Technical documentation on attached software used for this research is provided in Appendix B.

Chapter 2

Preliminaries

2.1 Commodity market models

The main requirement for a commodity price model for a sufficient liquid market is the absence of an arbitrage opportunity. Hence, a correct framework for the model has to be defined. For this purpose, the Heath-Jarrow-Morton framework (HJM) [43] and the results from Manoliu and Tompaidis [56] for normally distributed log commodity prices distribution are used. In addition, results by [66], [28], [78] and [16] are used for the jump diffusion models used in modelling the log commodity price. Explicit formulae are provided for commodity and futures prices for up to two factors. The next two sections discuss the futures pricing models with and without jumps.

2.2 Two-factor model for commodities

We use the treatment of commodity price models from [56]. Assume that spot price S_t is driven by the vector process $X_t = [x_t^{(1)}, x_t^{(2)}]^T$ and $t \in [0, T^*]$, where T^* is a time horizon. Let $\log S_t = \sum_i x_t^{(i)}$: $\{X_t \in \Omega\}$ on a probability space $(\Omega, \mathcal{P}, \mathcal{F}_t)$, where Ω is set of all possible realisations of X_t, $\mathcal{P}$ is the objective probability measure defined on Ω and $\mathcal{F}_t$ is the natural filtration. For representing discrete time intervals, a subscript n is used, $(t_n : n = 0, .., N$, $t_0 < t_1 < \cdots < t_N = T^*, \Delta := t_n - t_{n-1})$, where N is the number of time

intervals. The stochastic process for $X_t = \begin{bmatrix} x_t^{(1)} & x_t^{(2)} \end{bmatrix}^\top$ is described as follows:

$$\begin{aligned} dx_t^{(1)} &= (\alpha_1 - \kappa x_t^{(1)})dt + \sigma_1 dW_{1,t}^{\mathcal{P}}, && (2.1) \\ dx_t^{(2)} &= \alpha_2 dt + \sigma_2 dW_{2,t}^{\mathcal{P}}, && (2.2) \\ dW_{1,t}^{\mathcal{P}} dW_{2,t}^{\mathcal{P}} &= \rho dt, && (2.3) \end{aligned}$$

where $\alpha_1, \alpha_2, \kappa, \sigma_1, \sigma_2, \rho$ are real constants and $W_{1,t}^{\mathcal{P}}$, $W_{1,t}^{\mathcal{P}}$ are Wiener processes on $(\Omega, \mathcal{P}, \mathcal{F}_t)$. Let $T = [T_1, \ldots, T_M]$ be a vector of maturities, and let $T_m < T^*, \forall m = 0, \ldots, M$.

A futures price of a commodity with maturity time T_m at given time t is given by the formula:

$$F(t, T_m) = \mathbb{E}^{\mathcal{Q}}(S_{T_m} | \mathcal{F}_t), \tag{2.4}$$

where the futures price is the average of an asset in the future with respect to a risk free rate. One has to apply a change of measure to X_t, to price a futures contract.

Since the process X_t is a two-dimensional stochastic process, solution to its system can be found in terms $\log \tilde{F}(t, T_m) = x_t^{(1)} e^{\kappa t} + x_t^{(2)}$. A multivariate Ito's formula is applied to the latter:

$$\frac{d\tilde{F}(t, T_m)}{\tilde{F}(t, T_m)} = \left[(\alpha_1 e^{\kappa t} + \alpha_2)dt + \sigma_1 e^{\kappa t} dW_{1,t}^{\mathcal{P}} + \sigma_2 dW_{2,t}^{\mathcal{P}} \right]. \tag{2.5}$$

Assume, that (2.5) describes the stochastic evolution of a generic futures contract with maturity date T_m at time t and $\tilde{F}(t, T_m) = 0 \quad \forall t : t \in (T_m, T^*]$. To show the relationship of (2.5) with (2.4), the existence of an equivalent martingale measure has to be discussed.

For instance, let Φ_t be a $M + 1$ dimensional stochastic process adopted to the given natural filtration. $\Phi_t = (\phi_t^1, \ldots, \phi_t^M, \psi_t)$ is a futures trading strategy, where ϕ_t^m is the number of futures contracts with maturity date T_m at the given time t and ψ_t is the bank account. In addition, let $\phi_t^m = 0 \quad \forall t \in (T_m, T^*]$. Therefore, the value of a portfolio V_t is defined with the following equation:

$$V_t = \sum_{m=1}^{M} \phi_t^m \tilde{F}(t, T_m) + \psi_t B_t, \tag{2.6}$$

where B_t is a cash account, which is defined by $dB_t = r_t B_t, B_0 = 1$, where r_t is a deterministic short-term interest rate process.

The futures trading strategy is called self-financing if V_t satisfies the following SDE (see [56] for more details):

$$dV_t = \sum_{m=1}^{M} \phi_t^m d\tilde{F}(t, T_m) + \psi_t dB_t \tag{2.7}$$

Strategy (2.7) is arbitrage-free, if there is no possibility of making profit, without cash injections. More details on self-financing strategies, can be found in [28]. Henceforth, the strategy (2.7) is assumed to be arbitrage-free and, hence, an expression for F can be derived under the risk neutral measure. Hence, there exists a martingale measure $\mathcal{Q}$ equivalent to $\mathcal{P}$. Let the market price of risk be a two-dimensional adapted stochastic process $h_t = [h_t^1, h_t^2]^T$ on $(\Omega, \mathcal{Q}, \mathcal{F}_t)$, such that Novikov's condition (e.g. see [67]) is satisfied and the following holds:

$$\alpha_1 e^{\kappa t} + \alpha_2 = \sigma_1 h_t^1 + \sigma_2 h_t^2. \tag{2.8}$$

Moreover, from Girsanov's theorem [40]:

$$\frac{d\mathcal{Q}}{d\mathcal{P}} = \exp\left\{ -\sum_{i=1}^{2} \left(\int_0^T h_t^i dW_{i,t}^{\mathcal{P}} + \frac{1}{2} \int_0^T |h_t^i|^2 dt \right) \right\}. \tag{2.9}$$

Substitution of (2.8) into (2.5) and the change of measure yield the following relationship between Wiener processes:

$$W_{i,t}^{\mathcal{Q}} = W_{i,t}^{\mathcal{P}} + \int_0^t h_s^i ds.$$

Overall, the futures price process (2.5) under the risk-neutral measure is given as an SDE without a drift component:

$$d\tilde{F}(t, T_m) = \sigma_1 e^{\kappa t} dW_{1,t}^{\mathcal{Q}} + \sigma_2 dW_{2,t}^{\mathcal{Q}}. \tag{2.10}$$

Hence, using the fact that (2.8), $\tilde{F}(t, T_m)$ is related to $F(t, T_m)$ by changing the drift component.

The results in this section are later employed in chapter 3 for numerical experiments.

2.2.1 Two-factor model a practical example

Let S_t be the market commodity price, and let $X_t = [x_t^{(1)}, x_t^{(2)}]^\top$ be a vector-valued stochastic process on a probability space $(\Omega, \mathcal{P}, \mathcal{F}_t)$, where Ω is set of all possible realisations of X_t, $\mathcal{P}$ is the objective probability measure defined on Ω and $\mathcal{F}_t$ is the natural filtration. The elements in X_t are chosen such that $x_t^{(1)}$ is a mean-reverting process which represents a short term price fluctuations and $x_t^{(2)}$ is a Brownian motion representing long term price changes. Now, the market commodity price can be written as an exponent of the sum of elements in X_t: $\log S_t = \sum_{i=1}^{2} x_t^{(i)} + f(t)$, where $f(t)$ is a deterministic function of time, which represents a seasonality factor as before:

$$
\begin{aligned}
dx_t^{(1)} &= (\alpha_1 - \kappa x_t^{(1)})dt + \sigma_1 dW_{1,t}^{\mathcal{P}}, && (2.11)\\
dx_t^{(2)} &= \alpha_2 dt + \sigma_2 dW_{2,t}^{\mathcal{P}}, && (2.12)\\
f(t) &= c_1 + c_2 \sin(c_3 t + c_4), && (2.13)\\
dW_{1,t}^{\mathcal{P}} dW_{2,t}^{\mathcal{P}} &= \rho dt, && (2.14)
\end{aligned}
$$

where $\alpha_1, \kappa, \alpha_2, \sigma_1, \sigma_2$ are constants and $dW_{1,t}^{\mathcal{P}}$, $dW_{2,t}^{\mathcal{P}}$ are Wiener processes with correlation coefficient ρ. A vector $c = [c_1, c_2, c_3, c_4]^\top$ represents the seasonality parameters in $f(t)$.

Using the results from section 2.2, (2.11) and (2.12) can be written under the risk-neutral measure $\mathcal{Q}$ as following:

$$
\begin{aligned}
dx_t^{(1)} &= (\tilde{\alpha}_1 - \kappa x_t^{(1)})dt + \sigma_1 dW_{1,t}^{\mathcal{Q}}, && (2.15)\\
dx_t^{(2)} &= \tilde{\alpha}_2 dt + \sigma_2 dW_{2,t}^{\mathcal{Q}}, && (2.16)\\
dW_{1,t}^{\mathcal{Q}} dW_{2,t}^{\mathcal{Q}} &= \rho dt && (2.17)
\end{aligned}
$$

where:

$$
\begin{aligned}
\tilde{\alpha}_1 &= \alpha_1 - \lambda_1 \sigma_1,\\
\tilde{\alpha}_2 &= \alpha_2 - \lambda_2 \sigma_2,
\end{aligned}
$$

and λ_1, λ_2 are prices of risk. The solution to the equations (2.15) and (2.16 is obtained by using Ito's formula and integration over the interval $[t, t + \Delta]$. It

is easy to show that application of Ito's lemma to (2.15) and (2.16) gives us the following equations:

$$\mathbb{E}^{\mathcal{Q}}[\log S_{t+\Delta} \mid x_t^{(1)}, x_t^{(2)}] = f(t+\Delta) + e^{-\kappa\Delta}x_t^{(1)} + \frac{\tilde{\alpha}_1}{\kappa}\left(1 - e^{-\kappa\Delta}\right) + x_t^{(2)} + \tilde{\alpha}_2\Delta, \tag{2.18}$$

$$\mathrm{Var}^{\mathcal{Q}}[\log S_{t+\Delta} \mid x_t^{(1)}, x_t^{(2)}] = \frac{\sigma_1^2}{2\kappa}\left(1 - e^{-2\kappa\Delta}\right) + 2\frac{\rho\sigma_1\sigma_2}{2\kappa}\left(1 - e^{-\kappa\Delta}\right) + \sigma_2^2\Delta. \tag{2.19}$$

As before, the futures price $F(t, T_i)$ at time t with maturity date T_i is defined under the risk-neutral measure $\mathcal{Q}$:

$$F(t, T_i) = \mathbb{E}^{\mathcal{Q}}(S_{T_i} \mid \mathcal{F}_t). \tag{2.20}$$

Since $\log S_T$ is normally distributed:

$$\log F(t, T_i) = \mathbb{E}^{\mathcal{Q}}(\log S_{T_i} \mid \mathcal{F}_t) + \frac{1}{2}\mathrm{Var}^{\mathcal{Q}}(\log S_{T_i} \mid \mathcal{F}_t) \tag{2.21}$$

This allows us to derive an affine equation for the vector of log futures prices in terms of the log-spot price:

$$vec\{y_t^i\} = \mathbb{E}^{\mathcal{Q}}\left(\log S_{T_i} \mid x_t^{(1)}, x_t^{(2)}\right) + \frac{1}{2}\mathrm{Var}^{\mathcal{Q}}\left(\log S_{T_i} \mid x_t^{(1)}, x_t^{(2)}\right). \tag{2.22}$$

Referring to the filtering terminology mentioned in section 2.7.1, the system of equations (2.11)-(2.14) is the system of transition equations for the log-spot price process, which is treated as an unobserved or latent variable. The $\log F(t, T)$ for the measurement equation is given by (2.18)-(2.19) and (2.22). In addition, two cases when $f(t) = 0$ and $f(t) \neq 0$ are compared as two-factor model (TF) and two-factor model with seasonality (TFS) respectively. Results on numerical experiments for these models are presented in Section 3.4.

2.3 Two-factor model with jumps for commodity price modelling

The model discussed before works well with a range of assets. However, the main assumption is that a (log-) price path generated by the model has a

Gaussian distribution. This assumption of a Gaussian distribution is often inadequate. It was widely reported that the distribution of log-spot electricity prices is observed with fat tails, see e.g. [18], [33], [75] and [45]. The reason is that electricity is a commodity which cannot be stored. As a result, the market prices of electricity commodity exhibit high levels of volatility in short-time intervals. The latter results in very quick and sharp price movements over time and usually these movements have high amplitudes.

One possibility to model this behaviour is to force κ in (2.1) to take high values. This will allow process (2.1) to move around the long-run mean faster over time. However, this solution is temporary, which is limited by the chosen time frame, since an asset behaviour is not always 'explosive'. More modelling approaches to this problem are briefly discussed in chapter 5.

One solution to the problem can also be obtained by adding a compound Poisson (CP) process to the short-term process. The latter is called an affine-jump diffusion (AJD) process. The solution of the AJD process is complicated and require characteristic function based methods.

Here, the inclusion of a compound Poisson (CP) process to (2.1) and the solution to this is also discussed.

Let us use the system of SDEs from the previous chapter, but with some extra components:

$$
\begin{aligned}
dx_t^{(1)} &= (\alpha_1 - \kappa x_t^{(1)})dt + \sigma_1 dW_{1,t}^{\mathcal{P}} + dJ_t, && (2.23)\\
dx_t^{(2)} &= \alpha_2 dt + \sigma_2 dW_{2,t}^{\mathcal{P}}, && (2.24)\\
dW_{1,t}^{\mathcal{P}} dW_{2,t}^{\mathcal{P}} &= \rho dt, && (2.25)
\end{aligned}
$$

where J_t is a compound Poisson process with rate λ_J and $Y = \{Y_1, .., Y_t\} \sim N(\mu_J, \sigma_J^2)$ is the collection of jump amplitudes. The remaining components are defined in section 2.2. Note, J_t is uncorrelated with Wiener processes. The extra jump process in this equation complicates the solution. However, [29] used a jump-transform method to obtain the expression for the futures price. The solution is discussed here for a one dimensional CP process with a constant jump rate.

First, the following vectors and matrices are defined corresponding to equa-

tions (2.23) and (2.24):

$$K_0 = [\alpha_1 \quad \alpha_2]^\top, \tag{2.26}$$

$$K_1 = [\kappa \quad 0]^\top, \tag{2.27}$$

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2 \\ \sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \tag{2.28}$$

Let $v(x)$ be a jump size distribution and $g(x)$ be a corresponding density function respectively. Let $\theta(x)$ be the characteristic function of Gaussian distribution, then $\theta(x) = \int_{\mathbb{R}} e^{xz} dv(z) = \int_{\mathbb{R}} e^{xz} g(z) dz$. Let $\xi \equiv ([K_0, K_1], \Sigma, \theta, r)$ be the structure which captures both the distribution of $x_t^{(1)}$ and the effects of discounting. where r is defined as a constant interest rate. The transform Ψ^ξ is defined as:

$$\Psi^\xi(u, X, t, T) = \mathbb{E}^\xi \left[e^{-r(T-t)} e^{ux} \mid \mathcal{F}_t \right], \tag{2.29}$$

where $\mathbb{E}^\xi$ is the expectation operator with respect to ξ. Duffie et al [29] note that the difference between Ψ^ξ and a conditional characteristic function is the discount factor. They also prove in [29] that Ψ^ξ has the exponential-affine form:

$$\Psi^\xi(u, x, t, T) = \exp\{\alpha(u, t, T) + \beta(u, t, T)x\}, \tag{2.30}$$

where $\alpha(\cdot)$ and $\beta(\cdot)$ are solutions to the following Riccati equations:

$$\begin{aligned} \frac{\partial \alpha(u,t,T)}{\partial t} &= r - K_0^\top \beta(u,t,T) - \frac{1}{2}\beta(u,t,T)^\top \Sigma \beta(u,t,T) \\ &\quad -\lambda(\theta(\beta(u,t,T) - 1) \end{aligned} \tag{2.31}$$

$$\frac{\partial \beta(u,t,T)}{\partial t} = -K_1 \beta(u,t,T) \tag{2.32}$$

with boundary conditions $\alpha(u, T, T) = 0$ and $\beta(u, T, T) = u$, where $u = 1$.

A futures price of a commodity is given as before in section 2.2, now:

$$\begin{aligned} \mathbb{E}^{\mathcal{Q}}(S_{T_m}|\mathcal{F}_t) &= \mathbb{E}^{\mathcal{Q}}(e^{x_{T_m}^{(1)} + x_{T_m}^{(2)}}|\mathcal{F}_t) \\ &= e^{r(T_m - t)} E^{\mathcal{Q}}(e^{-r(T_m - t)} e^{x_{T_m}^{(1)} + x_{T_m}^{(2)}}|\mathcal{F}_t) \\ &= e^{r(T_m - t)} \Psi^\xi(u, x_t^{(1)} + x_t^{(2)}, t, T_m). \end{aligned} \tag{2.33}$$

This summarises the relationship between the jump-transform and futures price formula. An example of an application of a two-factor model with jumps for electricity futures is given in the next section.

2.4 Two-factor jump-diffusion Model a practical example

To define a jump-diffusion model for electricity futures prices, we start by modelling the behaviour of the commodity spot price. As in the previous section, assume a filtered probability space (Ω, $\mathbb{F}$,$\mathcal{P}$) with $\mathcal{P}$ being the historical measure and $\mathbb{F}$ being the natural filtration. The log commodity price is modelled in this case as:

$$\log S_t = f(t) + x_t + \zeta_t, \tag{2.34}$$

$$dx_t = (\bar{\alpha} - \kappa x_t)dt + \sigma_1 dW^{\mathcal{P}}_{1,t} + dJ_t, \tag{2.35}$$

$$d\zeta_t = \mu dt + \sigma_2 dW^{\mathcal{P}}_{2,t}, \tag{2.36}$$

$$f(t) = c_1 + \varsigma \sin(c_2 t + c_3), \tag{2.37}$$

$$\rho dt = dW_{1,t} dW_{2,t}, \tag{2.38}$$

The spot price process S_t consist of thee components: x_t represents a short-term mean-reversion process with price shocks driven by a compound Poisson process J_t which has intensity λ and jump sizes $Y = \{Y_1, .., Y_t\} \sim N(\mu_J, \sigma_J^2)$; ζ_t represents a long-term price process; seasonality $f(t)$ is a deterministic function of time. Instead of using a Fourier series based model, a simple, single sinusoid plus a level term is used. This form of $f(t)$ is used to de-seasonalise data for all the models in the numerical experiments, including the random volatility model described earlier. The coefficients $\lambda, \mu_J, \sigma_J, \alpha, \kappa, \sigma_1, \sigma_2, \varsigma, c_1, c_2, c_3$ are constants, $W^i_t, i = 1, 2$ are Wiener processes with a constant correlation ρ. Note that the models with two Wiener processes are referred as 'two-factor models' in this chapter. If the model has a jump component in addition to two Wiener processes, it is specified separately as a two-factor model with jumps.

Futures price $F(t, T)$ ($t \in [0, T)$) of a commodity with spot price S_t is $F(t, T) = \mathbb{E}^{\mathcal{Q}}[e^{\log S_T} | \mathcal{F}_t]$. Since the spot price process formulae (2.34)-(2.38) are given under a physical measure and the futures price formula is under a risk neutral measure, it is important to specify the change of measure involved. Let x_t be a risk-neutral mean-reversion process:

$$dx_t = (\alpha - \kappa x_t)dt + \sigma_1 dW^{\mathcal{Q}}_{1,t} + dJ_t, \tag{2.39}$$

where $dW^{\mathcal{Q}}_{1,t}$ is a Wiener process under risk-neutral measure and $\bar{\alpha} - \alpha =$

$h_x\sigma_1 + R(\lambda, \sigma_J, \beta)$. h_x is a market price of risk of mean-reversion process and β is a risk-aversion of the jump component. In the former approaches (e.g., see [60],[78]), jump risk premium was treated as an idiosyncratic component, which means that $R(\lambda, \sigma_J, \beta) = 0$. The formulae from [78] for the log futures price is modified to account for the excess rate of return $R(\lambda, \sigma_J, \beta)$:

$$\log F(t,T) = f(T) + e^{-\kappa(T-t)}x_t + \zeta_t + A(T-t) + B(T-t) \quad \text{where} \tag{2.40}$$

$$A(T-t) = (\mu - h_\zeta)(T-t) - \frac{h_x\sigma_1 + R(\lambda, \sigma_J, \beta)}{\kappa}(1 - e^{-\kappa(T-t)}) + \frac{\sigma_1^2}{4\kappa}(1 - e^{-2\kappa(T-t)}) + \frac{\rho\sigma_1\sigma_2}{\kappa}(1 - e^{-\kappa(T-t)}) + \frac{1}{2}\sigma_2^2(T-t), \tag{2.41}$$

$$B(T-t) = \lambda \int_t^T (\exp\{\mu_J + \frac{1}{2}\sigma_J^2 e^{-2\kappa(T-z)}\} - 1)dz, \tag{2.42}$$

where h_ζ is a market price of risk for the process ζ_t and α is set to 0, as before in chapter 3. [16] is used to introduce jump risk effects to the model, as outlined below.

A standard Poisson process J_t with rate λ and $\{Y_t\}$ *i.i.d.* copies of a random variable Y has the following property:

$$\phi(a) := \mathbb{E}[e^{aY}] < \infty$$

for a in some connected interval A containing the origin. Lévy exponent[1] for a compound Poisson process is $\psi(a) = \lambda(\phi(a) - 1)$ and the excess rate of return is then given by:

$$R(\lambda, h_1, h_2) = \lambda(\phi(h_1) + \phi(-h_2) - \phi(h_1 - h_2) - 1), \tag{2.43}$$

where $h_1, h_2 > 0$. Here, Y is assumed to be normally distributed with zero mean and variance σ_J^2:

$$\phi(a) = \exp\left(\frac{1}{2}a^2\right).$$

In this example, a mean is assumed to be zero for the jump size distribution and let the linear components of the futures price formula to take care of jump

[1] If a Lévy process X_t represents the class of General Lévy Models and $\mathbb{E}[e^{\alpha X_t}] < \infty$, then there exists a Lévy exponent $\psi(\alpha)$, such that $\mathbb{E}[e^{\alpha X_t}] = e^{t\psi(\alpha)}$, for more details see [16].

size. Substitution of $\phi(a)$ into the (2.43) yields:

$$R(\lambda, \sigma_J, \beta) = \lambda(e^{\frac{1}{2}\sigma_J^2} + e^{\frac{1}{2}\beta^2} - e^{\frac{1}{2}(\sigma_J - \beta)^2} - 1), \tag{2.44}$$

where β is the risk aversion of the jump component.

As a result, we have a system of equations (2.34)-(2.38) as the transition equations for the log spot price process, which is treated as an unobserved or latent variable. The measurement equation is given by (5.9), with $\log F(t,T)$ given by (2.40), (2.41), (2.42) and (2.44). The new two-factor model includes the jump risk premium derived recently in [16] and also includes a parametric periodic function to explicitly account for seasonality.

The particle filter algorithm, described earlier in section 2.7.2, was used for one step ahead prediction of futures prices for three different models: the random volatility model (equations (5.7)-(5.8)), two-factor jump-diffusion model (equations (2.34)-(2.36) and (2.40)) and the two-factor model without jumps, which is obtained by setting the relevant parameters in the jump-diffusion model to zero.

2.5 Hull-White option pricing framework

The stochastic-volatility model for European call option pricing defined by Hull-White in [49] is discussed briefly in this section. This section will not provide an exact solution to the problem, but will outline the necessary framework for a new stochastic volatility model discussed in chapter 4.

In the Hull-White model, a security price process S with its instantaneous stochastic variance $V_t = \sigma_t^2$ has the following form:

$$\frac{dS_t}{S_t} = \mu dt + \sigma_t dW_t, \tag{2.45}$$

$$\frac{dV_t}{V_t} = \alpha dt + \eta dZ_t, \tag{2.46}$$

where σ_t is an instantaneous volatility, μ, α and η are constants, W_t and Z_t are uncorrelated Wiener processes.

To price a security f depending on (2.45) and (2.46), one can use results derived by Garman [35]. The partial differential equation for f can be written

as follows:

$$\frac{\partial f}{\partial t} + \frac{1}{2}\left[\sigma_t^2 S_t^2 \frac{\partial^2 f}{\partial S_t^2} + \eta^2 V_t^2 \frac{\partial^2 f}{\partial V_t^2}\right] - rf = -rS_t \frac{\partial f}{\partial S_t} - \mu\sigma_t^2 \frac{\partial f}{\partial V_t}, \tag{2.47}$$

where r is a risk-free rate.

The solution to (2.47) for a European call option with the maturity time T is given as follows:

$$f(S_t, V_t, t) = e^{-r(T-t)} \int f(S_T, V_T, T) p(S_T|S_t, V_t) dS_T, \tag{2.48}$$

here $f(S_t, V_t, t) = \max[0, S - K]$, where K is the strike price, and $p(S_T|S_t, V_t)$ is the conditional distribution of S_t. Since $p(\cdot)$ depends on two variables, one can use the following fact. For three related random variables x, y, z the conditional probability density function (pdf) can be defined as follows:

$$p(x|y) = \int g(x|z) h(z|y) dz, \tag{2.49}$$

where $g(\cdot)$ and $h(\cdot)$ are conditional probabilities (see [41] for the definition and the proof). Let $\bar{V}$ be a mean variance:

$$\bar{V} = \frac{1}{T-t} \int_t^T \sigma_t^2 dt, \tag{2.50}$$

then using (2.49) one can write (2.48) as follows:

$$f(S_t, V_t, t) = e^{-r(T-t)} \int \int f(S_T, V_T, T) g(S_T|\bar{V}) h(\bar{V}|\sigma_t^2) dS_T d\bar{V}, \tag{2.51}$$

rearranging the elements in the previous equation gives us the following form:

$$f(S_t, V_t, t) = \int \left[e^{-r(T-t)} \int f(S_T, V_T, T) g(S_T|\bar{V}) dS_T \right] h(\bar{V}|\sigma_t^2) d\bar{V}. \tag{2.52}$$

Elements in brackets in (2.52) is the Black-Scholes formula. Hence, (2.52) can be written as:

$$f(S_t, V_t, t) = \mathbb{E}\left[C_{BS}(\bar{V})\right], \tag{2.53}$$

where C_{BS} is the Black-Scholes price for a European call option.

Chapter 4 discusses a new stochastic volatility model, which uses the result discussed in this section.

The next section discusses popular stochastic volatility models with and without jumps.

2.6 Heston model and Bates (SVJ) model

The formulae are outlined first for pricing European options using the Heston and the Bates (SVJ) models. Later, these two models is used as benchmarks for the new stochastic volatility model defined in chapter 4. All the subsequent discussion is in a (non-unique) equivalent martingale measure and explicit mention of measure is omitted for simplicity. For the Heston model, the asset price dynamics is assumed to be governed by:

$$dS_t = rS_t dt + \sqrt{v_t} S_t dW_t^1, \tag{2.54}$$

$$dv_t = -\theta(\bar{v} - v_t)dt + \sigma_v \sqrt{v_t} dW_t^2, \tag{2.55}$$

where r is the risk-free rate, W_t^1 and W_t^1 are standard Wiener processes with a given correlation $< W_t^1, W_t^2 >= \rho$ and $\rho, \sigma_v, \theta, v_0, \bar{v}$ are known constants. The price of a European call option with strike price K is given by:

$$C_{EUR} = S_t P_1 - K e^{-r(T-t)} P_2, \tag{2.56}$$

where S_t is a spot price at time t, T is a the expiration time and $P_j, j = 1, 2$ are called the pseudo-probabilities:

$$P_j = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty Re \left[\frac{e^{ix \log(\frac{S_t}{K})} e^{\phi_j(v_t, \tau, x)}}{ix} \right] dx. \tag{2.57}$$

Here, $\tau = T - t$ and $\phi_j(v_t, \tau, x) = \exp\{C_j(\tau, x)\bar{v} + D_j(\tau, x)v_t\}$ is the characteristic function, with

$$\begin{aligned}
C_j(\tau, x) &= rxi\tau + \frac{\theta}{\sigma_v^2} \left[(b_j - \rho\sigma_v xi + d_j)\tau - 2 \log \frac{1 - d_j e^{d_j \tau}}{1 - g_j} \right], \\
D_j(\tau, x) &= \frac{b_j - \rho\sigma_v xi + d_j}{\sigma_v^2} \left[\frac{1 - e^{d_j \tau}}{1 - g_j e^{d_j \tau}} \right], \\
g_j &= \frac{b_j - \rho\sigma_v xi + d_j}{b_j - \rho\sigma_v xi - d_j}, \quad d_j = \sqrt{(\rho\sigma_v xi)^2 - \sigma_v^2(2u_j xi - x^2)}, \\
u_1 &= \frac{1}{2}, u_2 = -\frac{1}{2}, \text{ and } b_j = \kappa + \theta - (1_{j=1})\rho\sigma_v.
\end{aligned}$$

Bates in [11] proposed adding a compound Poisson process in the underlying for the above model, which leads to a modification of (2.54):

$$\frac{dS_t}{S_t} = rdt + \sqrt{v_t}dW_t^1 + (e^{\alpha+\beta\epsilon} - 1)dJ_t, \tag{2.58}$$

where J_t is a Poisson process with a known jump intensity λ_J, α, β are known constants and $\epsilon \sim N(0,1)$. The process J_t is uncorrelated with $W_t^i, (i = 1, 2)$. The volatility dynamics is described by equation (2.55). The solution for price of a European call option is given by modifying the characteristic function in the Heston model above:

$$\phi_j(v_t, \tau, x) = \exp\{C_j(\tau, x)\bar{v} + D_j(\tau, x)v_t + E(x)\tau\},$$

where

$$E(x) = -\lambda_J ix(e^{\alpha+\beta^2/2} - 1) + \lambda_p(e^{ix\alpha - x^2\beta^2/2} - 1).$$

While both these models have proved popular and are known to provide good fits to option prices, they have a few shortcomings. Some of these are discussed in [61]. In particular, it was shown that the Heston model usually fails to fit to a short term market skew while the SVJ model usually fails to fit an inverse yield curve. In addition, the option price is given through a fairly involved numerical integral with several parameters, which presents significant difficulties in calibration.

The next section discusses a class of filtering problems and algorithms for applications in finance.

2.7 Filtering Algorithms

In finance, there are many instances when the data can be assumed to be corrupted by noise (e.g. bid-offer spread may be considered as noise in a unique unobserved price) or the variable of interest is unobservable (e.g. the continuously compounded interest rate). Such noisy and/or unobserved variables can be obtained from observed data using filtering algorithms. These algorithms can also be used for estimating the parameters of the underlying stochastic processes.

Harvey [42] suggested representing market data (such as interest rates changes, stock price movements) as finite or infinite time series. These include

a trend and seasonal pattern components. Moreover, time series representation allows to forecast the future asset behaviour.

Schwartz [71] used an Ornstein-Ulenbeck like model for analysing the dynamics of commodity prices. He used a Kalman filer [52], which is described in detail in the next subsection, to estimate the parameters of this model.

Later, Babbs and Nowman [7] introduced the state-space formulation of the Vasicec term structure model. They showed the application of Kalman filter for analysing and predicting of underlying interest rates from observed bond yields. More analysis was made using the Kalman filter on multi-factor extension of the Vasicek model [24].

More results on the Kalman filter application to model commodity prices were presented by Manoliu and Tompadis [56]. They introduced a multi-factor model for pricing of futures where the underlying commodity price is unobserved. The Kalman filter and maximum likelihood methods were used for numerical experiments.

Lautier and Galli [54] compared two versions of the Kalman filter: simple and extended. The extended Kalman filter is a generalisation of the first one, that allows it to work with some types of nonlinear models. Monoyios [63] was exploring the impact of the uncertainty by using the Kalman filter for the optimal hedging.

Thus the Kalman filter is very widely used in finance. The Kalman filter algorithm is outlined in the next section, before moving on to its various extensions in the subsequent sections.

2.7.1 Kalman filter

A general model for the Kalman filter can be represented with a stochastic system with the following state space equations:

$$dx_t = f(x_t, t)dt + R(x_t, t)dW_t \tag{2.59}$$

$$y_t = h(x_t, t) + B(x_t, t)dZ_t, \tag{2.60}$$

where measurement equation(2.60) represent the observed value y_t and transition equation (2.59)represent the unobserved underlying process x_t, functions $f(.)$ and $h(.)$ are deterministic and can be vector valued, $R(.)$ and $B(.)$ are deterministic matrix-valued functions, W_t and Z_t are Wiener processes. Z_t is

also called the measurement noise.

However, in most applications, the above system can be presented in a discrete form. This is possible since in the most applications measurements are discrete, rather than being continuous. Let $t_n : 0 < t_1 < \ldots < t_N$, where N is the total number of measurements, then $x_n \equiv x_{t_n}$ and $y_n \equiv y_{t_n}$. If the f, h are affine in x_t and if R, Q are constant matrices, equations (2.59) and(2.60) can be presented in the discrete linear form as follows:

$$\begin{aligned} x_{n+1} &= Bx_n + g + RW_{n+1} && (2.61)\\ y_n &= Ax_n + d + QZ_n, && (2.62) \end{aligned}$$

where A and B are linear functions of time, g and d are vectors, can also be functions of time, R and Q are variance matrices, can also be functions of time, y_n represent observed value and x_n is the underlying process, presented by transition equation (2.61) and is a subject to be estimated.

The method for solving this problem was proposed by Kalman in [52].

The estimated value of x_n is based on the information up to time t_{n-1} as $\hat{x}_{n|n-1}$, conditional variance of $\hat{x}_{n|n-1}$ as $P_{n|n-1}$ and is is assumed that $\hat{x}_{0|-1}$ and $P_{0|-1}$ are known. Noise components and information about the system are not available. The recursive algorithm for the linear Kalman Filter has the following structure:

$$\begin{aligned} v_n &= y_n - (A\hat{x}_{n|n-1} + d) && (2.63)\\ \Sigma_n &= AP_{n|n-1}A^\top + QQ^\top && (2.64)\\ k_n &= P_{n|n-1}A^\top\Sigma_n^{-1} && (2.65)\\ \hat{x}_{n|n} &= \hat{x}_{n|n-1} + k_n v_n && (2.66)\\ \hat{x}_{n+1|n} &= B\hat{x}_{n|n-1} + g && (2.67)\\ P_{n+1|n} &= BP_{n|n-1}B^\top + RR^\top - BP_{n|n-1}A^\top\Sigma_n^{-1}AP_{n|n-1}B^\top && (2.68) \end{aligned}$$

Here v_n represents the innovations of the Kalman filter, Σ_n represents the covariance matrix of innovations, $\hat{x}_{n+1|n}$ is the best estimate of x_{n+1} based on the information up to time t_n and determined by measurement equation denoted by y_n.

However, the Kalman filter can only be used for a class of problems where x and y have a linear relationship and both have a conditional Gaussian distri-

bution. In the next section filtering algorithms for nonlinear and non-Gaussian problems are discussed.

2.7.2 Particle filter

To set up a nonlinear problem for a particle filter, the following state space system is considered:

$$x_k = f_k(x_{k-1}, v_{k-1}), \tag{2.69}$$

$$y_k = h_k(x_k, \epsilon_k), \tag{2.70}$$

where $f(\cdot,\cdot)$ is a vector function of a state x_{k-1} and $\{v_{k-1}\}(k \in \mathbb{N})$ is an i.i.d. process noise sequence, $h(\cdot,\cdot)$ is a known function, $\{\epsilon_k\}$ is a zero mean i.i.d. sequence which corresponds to measurement noise and k represents the current time-step. As in the previous section x_k is unobservable, and the aim here is to construct the estimate of x_k by combining model prediction with the measurement y_k, at each time step k. (2.69) and (2.70) are referred to as the transition equation and the measurement equation, respectively. To approach this, the pdf $p(x_k|y_{1:k})$ has to be constructed. Assume that the initial pdf (prior) $p(x_0, y_0)$ is known. Then as each new measurement y_k arrives, $p(x_k|y_{1:k})$ can be constructed recursively within two steps: prediction and update. The prediction step is based on the Chapman-Kolmogorov equation:

$$p(x_k|y_{1:k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|y_{1:k-1})dx_{k-1}, \tag{2.71}$$

where $p(x_{k-1}|y_{1:k-1})$ is assumed to be known and the transition probability $p(x_k|x_{k-1})$ is defined by (2.69). After the measurement y_k becomes available, it can be used to update the prior using Bayes' rule:

$$p(x_k|y_{1:k}) = \frac{p(y_k|x_k)p(x_k|y_{1:k-1})}{p(y_k|x_{1:k-1})} \tag{2.72}$$

where

$$p(y_k|x_{1:k-1}) = \int p(y_k|x_k)p(x_k|y_{1:k-1})dx_k \tag{2.73}$$

In case when the posterior density is not available analytically, the approximation can be done with a set of random samples (or particles) with associated probability weights. Assume that the underlying process is a Markov process,

which will be the case in all the processes considered in this thesis. A short description is provided for the recursion to update a probability measure $\{x_k^i, w_k^i\}$ at time t_k to a corresponding probability measure at time t_{k+1}, (See, e.g. [6] for more details).

- Sample x_{k+1}^i from $q(x|x_k, y_{k+1})$. q is called a *proposal density* and serves as an approximation to the posterior density. The choice of q is crucial in terms of the quality of estimates. A common choice is $q(x|x_k, y_{k+1}) = p(x_{k+1}|x_k)$, although other choices are possible (e.g. Gaussian density generated using the extended Kalman filter is also frequently employed).

- The weight update can be done using the following relation:

$$\omega_{k+1}^i \propto \omega_k^i \frac{p(y_{k+1}|x_{k+1}^i)p(x_{k+1}^i|x_k^i)}{q(x_{k+1}^i|x_k^i, y_{k+1})}, \tag{2.74}$$

  and the posterior filtered density is given by:

$$p(x_{k+1}|x_k) \approx \sum_{i=1}^{N_s} \omega_k^i \delta(x_{k+1} - x_{k+1}^i), \tag{2.75}$$

  where δ is the Dirac-Delta function.

An issue which frequently arises in using a particle filter is the degeneracy phenomenon, when after a few iterations particle weights might start pinning around one value. However, this can partially be solved by introducing a measure for the degeneracy problem:

$$N_{\text{eff}} = \left(\sum_{i=1}^{N_s} (\omega_k^i)^2 \right)^{-1}, \tag{2.76}$$

where N_s is the actual number of samples. If N_{eff} falls below some pre-defined threshold, a resampling algorithm is applied. See [6] for more details on resampling algorithms.

The Kalman filter and the particle filter are great algorithms to estimate the value of unobserved processes or variables, once the model parameters are available. However, methods to estimate the model parameters from data are needed. The next section will discuss such algorithms.

2.8 Optimisation algorithms

This research is focused on three popular methods for estimating model parameters. These methods include the weighted least squares method, the maximum likelihood method and the method of moments. In the following subsection the weighted least squares method is discussed.

2.8.1 Weighted least squares

The method of least squares was developed in the 17th century. The study of the behaviour of celestial objects was performed by Gauss. He was the first to introduce the method of least squares, which is widely used now by scientists.

Nowadays, there are two types of least-squares problems: linear and nonlinear. Linear least-squares methods can be solved using systems of linear equations and specify an important class of problems in statistics. However, this research deals with nonlinear problems, and nonlinear least squares is our priority.

The nonlinear least squares method can be defined in the following way. Let $f(x, \Theta)$ be a real valued function, where $x \in \mathbb{R}^n$ is a coordinate vector and $\Theta \in \mathbb{R}^m$ is a vector of a constant real valued parameters, so $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$. Let $y \in \mathbb{R}^n$ be a measurement vector, with corresponding coordinates x. Then the nonlinear least-squares optimisation problem is:

$$\min_{\Theta} ||y - f(x, \Theta)||. \tag{2.77}$$

In real life applications with nonlinear function $f(x, \Theta)$, cost function (2.77) can produce poor results. These is the result of equal weights for all the measurements in y, while some elements in y can be less important than others. To account for possibly differing importance of different measurements, a weighted least squares method is often used.

A good example for this problem is an option price market. Option price data on an asset consists of three main parameters: a price of the underlying asset, strike price and option maturity. Each option has it's bid and ask prices updated at each unit of time. The difference between them is called a bid-ask spread. The spread indicates the liquidity of the option. Smaller values indicates higher liquidity and visa versa. Let $bid \in \mathbb{R}^n$ and $ask \in \mathbb{R}^n$ be the vectors of market closing bid and ask quotes respectively, and let $y =$

$\frac{1}{2}(bid - ask)$ be a closing mid price. These vectors can be used in calibration of parameters for an option pricing model by adjusting (2.77) with bid-ask weights vector $\omega = 1/(bid - ask)^N$, where N is an arbitrary natural number (usually empirically chosen). Then problem (2.77) is:

$$\min_b \sum_{i=1}^{n} (y - f(x, \Theta))^2 \omega_i. \tag{2.78}$$

The latter is widely used in calibration of option pricing models. However, parameter N should be chosen carefully, because it affects the sensitivity of weighting, hence calibration. (2.78) can be solved numerically on any high-level programming language such as Matlab.

Discussion on applications of weighted least-squares are continued in chapter 4.

The next subsection will discuss a different approach for calibration.

2.8.2 Maximum Likelihood

Finding parameters for financial models can be implemented using a Maximum Likelihood method. Let $x \in R^n$ be a vector of identically distributed observations, $f(\cdot|\Theta)$ be a distribution function of x with a vector of parameters Θ defined as before. Since x is a collection of iid samples, $f(\cdot|\Theta)$ can be written as a joint-pdf:

$$f(x|\Theta) = \prod_{i=1}^{n} f(x_n|\Theta). \tag{2.79}$$

Now, let $L(\Theta; x)$ be a likelihood function, where the vector Θ is a free variable, and the vector x is given:

$$L(\Theta; x) = \prod_{i=1}^{n} f(x_n|\Theta), \tag{2.80}$$

such that the value of $L(\Theta; x)$ describes how likely Θ defines the distribution of x.

For example, let $x \sim N(\mu, \sigma^2)$. Then $\Theta = [\mu, \sigma]^\top$. Now, the problem is to find the best estimate $\hat{\Theta}$, which will more likely describe the given sample x.

For a Gaussian distribution:

$$f(x|\Theta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{2.81}$$

then the likelihood function is given by:

$$L(\Theta; x) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \prod_{i=1}^{n} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}. \tag{2.82}$$

Now one can write the optimisation problem for finding $\hat{\Theta}$:

$$\hat{\Theta} = arg\max_{\Theta} \frac{1}{(\sqrt{2\pi}\sigma)^n} \prod_{i=1}^{n} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}. \tag{2.83}$$

However, in real-life applications the log-likelihood is used instead due to computational efficiency. Hence, the logarithm of a product is a sum, which is computationally faster. By taking the logarithm of (2.82) and taking out the constants the following optimisation problem is given:

$$\hat{\Theta} = arg\max_{\Theta} - \left(\frac{n}{2} \log \sigma^2 + \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right). \tag{2.84}$$

In practice, x_t is often modelled as having a multivariate normal distribution and the mean vector and the covariance matrix may be functions of underlying model parameters. This situation is encountered in chapter 3.

The calibration of the objective function can be performed using derivative and non-derivative methods. However, in some situations for non-Gaussian distributions finding the value of the derivative is complicated. In these situations simplex methods are used.

The next subsection will discuss moment matching methods for finding parameters.

2.8.3 Method of moments

Likelihood methods are effective when the distribution of the observation is known. However, some financial models don't have an explicit form of probability distribution. For example, stochastic models with jumps don't have an explicit form for a probability distribution. In this case, the *method of moments* can be implemented. The main idea of this method is matching

the sampled moments of the observed data to the analytical moments of the model.

In chapter 5 the moments of the distribution of the electricity spot prices are used to estimate parameters for the two-factor model with jumps. The fact is used that the knowledge of the characteristic function of the distribution of the model allows one to obtain an explicit form for the moments up to any required order. In combination with least squares methods, this yields the solution to the problem. More detailed discussion on this problem is continued in chapter 5.

Chapter 3

Commodity price forecasting using the Kalman filter

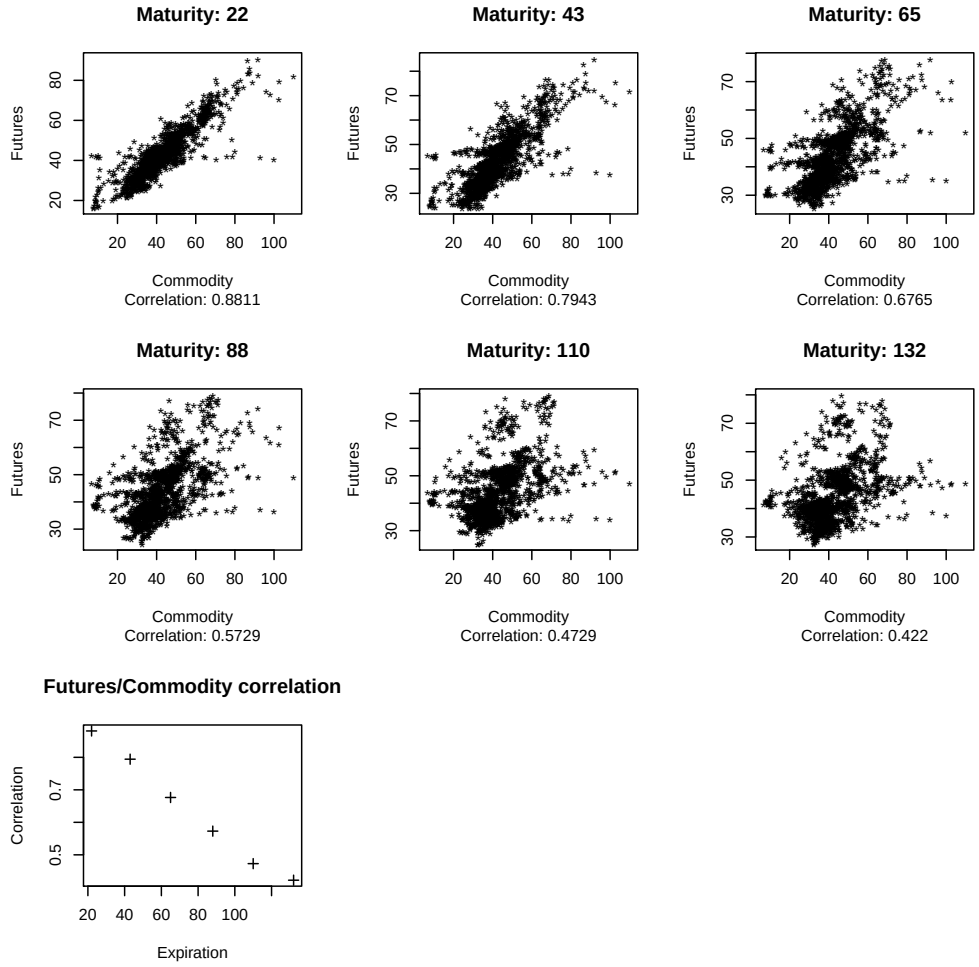
3.1 Introduction

This chapter is concerned with the problem of forecasting commodity spot prices using futures price information. As futures on commodities tend to be more liquid than the commodities themselves in the spot market, they contain more information about the future behaviour of spot prices than the current spot market prices. Scatter[1] and correlation plots on figure 3.1 illustrate that futures prices are not perfectly correlated with spot prices and different futures contracts contain different information, than the spot price history, regarding the future behaviour of spot prices.

This fact can be exploited to infer the future spot price behaviour using futures prices via a Kalman filter, with the spot price as a latent state variable. Filtering has been used in estimating the spot prices from futures prices in [71], [56] and [54]. A multi-commodity implementation is presented in [19], where the futures prices on different commodities are used simultaneously to forecast the commodity prices. The inclusion of jumps to a commodity price process and its subsequent use of particle filter for inference on commodity prices is advocated in [2]. In [20], a three factor model for oil futures prices is suggested, which departs from a Bayesian viewpoint used in filtering and infers prices using a numerical (but simple) optimization instead. [62] provides a model

[1]Natural Gas spot prices were plotted against futures prices with different maturities for an interval of 1500 days, from 29.11.2007 to 12.09.2012.

Figure 3.1: Scatter and correlation plots for Natural Gas commodity and futures (1500 data points).



where the seasonality components are also stochastic and hence allow for a frequency variation.

Besides estimation of the commodity prices, the Kalman filter is discussed in section 2.7.1 and its various modifications have also been used within financial mathematics for modelling and forecasting of interest rates ([7], [24]) and estimating asset price volatility from intra-day stock prices [10]. Linear filtering is used in [63] in the context of hedging in incomplete markets, for updating the estimates of uncertain drift parameters in the price process. [23] provides a review of applications of filtering within finance. Date at al in [22] developed a regime switching model for commodity futures pricing and forecasting under Hidden Markov Model framework.

In this chapter, the approach taken by Manoliu and Tompaidis in [56] has been followed and extended. They [56] introduced a multi-factor model for futures pricing with unobserved underlying commodity prices. The emphasis here is on storable energy commodities with highly liquid futures markets, *viz* crude oil, gasoline and natural gas. The evolution of log-spot price is modelled as a mean-reverting process, resulting in a linear state space system with log futures price vector as observable variable. Unlike [56], which uses a non-parametric seasonality component, the seasonality factor is modelled explicitly in this chapter through a simple parametrized sinusoid. The long run mean of the log spot price is set to be random, which offers a very parsimonious way of adding an extra source of randomness. A detailed empirical study of calibration and out-of-sample forecasting of commodity prices is provided for three different commodities, with two different data-sets considered for each commodity. A comparison of various models in terms of parametric complexity and out-of-sample prediction accuracy is done. To ours knowledge, comparison of out-of-sample forecasting performance has not been reported in the literature on commodity futures price modelling before.

The rest of this chapter is structured as follows. Section 3.2 presents the formulation of the models for evolution of futures prices. In Section 3.3, the state space representation of the same models is described, which is used for Kalman filtering and maximum likelihood based parameter estimation. In section 3.4, the numerical experiments are presented, including information about data, algorithms, validation metrics and the results of experiments on the forecasting ability of the models. Finally, section 3.5 summarises the contributions of this chapter.

3.2 Formulation of the models

In this section, a short formulation of one-factor arbitrage-free model is provided for the futures price dynamics. The formulation is based on [56], to which the reader is referred to for proofs. The added contributions are related to adding a random long run mean and a parametrized seasonality factor, and to a comprehensive empirical evaluation of the forecasting performance of models in subsequent sections. Four different models are described, depending on whether there are one or two-factors and whether or not there is a seasonality component. It is straightforward to extend this work beyond two-factors,

although ours numerical experience indicates that models with three of more factors are rather difficult to calibrate reliably and tend to perform poorly out-of-sample, as far as forecasting is concerned.

3.2.1 One factor model

Assume that a spot price S_t is driven by the process, $x_t = \log S_t : \{x_t \in \Omega\}$ on a probability space $(\Omega, \mathcal{P}, \mathcal{F}_t)$, where Ω is a set of all possible realisations of x_t, $\mathcal{P}$ is the objective probability measure defined on Ω and $\mathcal{F}_t$ is the natural filtration. For representing discrete time, the subscript n is used, $(t_n : n = 0, .., N,\ t_0 < t_1 < \cdots < t_N,\ \Delta := t_n - t_{n-1})$, where N is the total number of time intervals.

The log-spot price is assumed to follow an Ornstein-Ulenbeck type process:

$$dx_t = (\alpha - \kappa x_t)dt + \sigma dW_t^{\mathcal{P}}, \tag{3.1}$$

where κ and α represent mean-reversion speed and long-run mean of x_t, respectively, and $W_t^{\mathcal{P}}$ is a Wiener process. The fundamental theorem of asset pricing states that the absence of arbitrage opportunities on the market implies an existence of the equivalent martingale measure. Hence, process x_t has the following form if generated by a risk-neutral Wiener process $W_t^{\mathcal{Q}}$:

$$dx_t = (\tilde{\alpha} - \kappa x_t)dt + \sigma dW_t^{\mathcal{Q}}. \tag{3.2}$$

The drifts are related by $\tilde{\alpha} = \alpha - \lambda_t \sigma$ for some process λ_t, i.e. $dW_t^{\mathcal{Q}} = dW_t^{\mathcal{P}} + \lambda_t dt$; see, e.g. [28] for the exact conditions on λ_t. λ is assumed to be a constant here, which is a commonly used assumption in the literature.

Under the risk-neutral measure, the process x_t is normally distributed. Using Ito's lemma for the function $f(x_t, t) = e^{\kappa t} x_t$, it can be easily shown that x_t has the following mean and variance:

$$\mathbb{E}^{\mathcal{Q}}(x_{t+\Delta} \mid \mathcal{F}_t) = x_t e^{-\kappa\Delta} + \frac{\tilde{\alpha}}{\kappa}(1 - e^{-\kappa\Delta}), \tag{3.3}$$

$$\mathrm{Var}^{\mathcal{Q}}(x_{t+\Delta} \mid \mathcal{F}_t) = \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa\Delta}\right). \tag{3.4}$$

To add further flexibility to the model at a very modest increased complexity, in (3.1) the assumption that $\alpha \sim N(\mu_0, \Theta^2)$ is used. This allows the

log-spot price to converge to a random mean and potentially improves the predictive ability of the model, at the cost of a single added parameter to the parametric model. The investigation on whether this added parameter improves price prediction or not is continued in section 4. The random mean is assumed to be uncorrelated with the Wiener process. Hence, the expressions[2] for conditional mean and variance can be written as:

$$\mathbb{E}^{\mathcal{Q}}(x_{t+\Delta} \mid \mathcal{F}_t) = x_t e^{-\kappa\Delta} + \frac{\mu_0 + \lambda\sigma}{\kappa}(1 - e^{-\kappa\Delta}), \tag{3.5}$$

$$\mathrm{Var}^{\mathcal{Q}}(x_{t+\Delta} \mid \mathcal{F}_t) = \frac{\Theta^2}{\kappa^2}\left(1 - e^{-\kappa\Delta}\right)^2 + \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa\Delta}\right). \tag{3.6}$$

Next, let $T = \{T_i : i = 1, ..., m, 0 < T_1 < T_2 < \cdots < T_m\}$ be the collection of the futures maturity dates. Then futures price for maturity T_i for a commodity with log-spot price x_t at time $t < T_i$ can be written as a conditional expectation of the commodity price at the maturity time of the futures contract: $F(t, T_i) = \mathbb{E}^{\mathcal{Q}}(e^{x^i} \mid \mathcal{F}_t)$, $i = 0, .., m$, where the expectation is taken under the $\mathcal{Q}$ measure and $x^i := x_{T_i}$, for brevity of notation. In the case if $T_i > t$, the futures price $F(t, T_i) > 0$, otherwise it is zero. The time to expiry of the i^{th} futures contract is represented by $\Delta_t^i = T_i - t$. Since S_t is log-normally distributed, the futures price is given by:

$$F(t, T_i) = \mathbb{E}^{\mathcal{Q}}(e^{x^i} \mid \mathcal{F}_t) = e^{\mathbb{E}^{\mathcal{Q}}(x^i|\mathcal{F}_t) + \frac{1}{2}\mathrm{Var}^{\mathcal{Q}}(x^i|\mathcal{F}_t)}. \tag{3.7}$$

This allows us to derive an affine equation for the vector of log futures prices in terms of the log-spot price:

$$vec\{y_t^i\} = x_t e^{-\kappa\Delta_t^i} + \frac{\mu_0 - \lambda\sigma}{\kappa}(1 - e^{-\kappa\Delta_t^i}) + \frac{\sigma^2}{4\kappa}(1 - e^{-2\kappa\Delta_t^i}) + \frac{\Theta^2}{2\kappa^2}\left(1 - e^{-\kappa\Delta_t^i}\right)^2, \tag{3.8}$$

where $y_t^i = \log F(t, T_i)$ and the *vec* operator is defined by

$$vec(z_i) = \begin{bmatrix} z_1 & z_2 & \cdots & z_n \end{bmatrix}^\top.$$

Note that the convenience yield is not modelled explicitly and assume that it is already reflected in the prices of futures contracts. Again, our approach is consistent with the framework followed in [56]. In contrast, convenience yield is explicitly modelled in [50].

[2]Note that, conditional on x_t and α, $x_{t+\Delta}$ is still Gaussian.

3.2.2 One factor model with seasonality

Since energy futures prices depend on the weather conditions, any seasonality pattern needs to be taken into account. In the literature, a variety of seasonality functions for different financial application is used. For example, Manoliu and Tompaidis [56] used a discrete seasonality function with separate parameters representing each month, while Sorensen [73] used a Fourier series to model seasonality. However, a complicated seasonality function makes parameter estimation more difficult and may lead to poorer estimates, especially when the data set is small relative to the number of the parameters. To reduce the parameter estimation complexity, a simple function for seasonality is considered, which is parametrised as follows:

$$f(t) = \exp(c_1 + c_2 \sin(c_3 t + c_4)), \tag{3.9}$$

where c_1 is a constant level, c_2, c_3 and c_4 are constants representing amplitude, the frequency and the phase of a seasonal pattern respectively. Accordingly, the prices of futures are modified as follows:

$$F(t, T_i) = f(T_i)\mathbb{E}^{\mathcal{Q}}(e^{x^i} \mid \mathcal{F}_t), \tag{3.10}$$

and

$$\begin{aligned} vec\{y_t^i\} = \log f(T_i) + x_t e^{-\kappa\Delta_t^i} + \frac{\mu_0 - \lambda\sigma}{\kappa}(1 - e^{-\kappa\Delta_t^i}) + \frac{\sigma^2}{4\kappa}(1 - e^{-2\kappa\Delta_t^i}) \\ + \frac{\Theta}{2\kappa^2}\left(1 - e^{-\kappa\Delta}\right)^2, \end{aligned} \tag{3.11}$$

which denotes a vector of log futures prices, with i^{th} element of the vector denoting log futures price for time to maturity Δ_t^i, as before. In practice, one may parametrise seasonality using multiple sinusoids. However, in our experience, this complicates parameter estimation without necessary improving the quality of out of sample price forecasting.

3.2.3 Observable commodity prices

Finally, a simple model of commodity price is described, which will be used as a benchmark. In this model the futures prices are *not* used as an extra source of information and the log-spot price in (3.1) is considered as observable, with

$\alpha \sim N(\mu_0, \Theta^2)$. In this case, the process and the model can be discretized, which preserves the exact conditional moments, and the process is given by

$$x_{n+1} = \mathbb{E}^{\mathcal{P}}(x_{n+1} \mid \mathcal{F}_n) + \sqrt{\mathrm{Var}^{\mathcal{P}}(x_{n+1} \mid \mathcal{F}_n)} W_{n+1}, \tag{3.12}$$

where $\mathbb{E}^{\mathcal{P}}(x_{n+1} \mid \mathcal{F}_n)$ and $\mathrm{Var}^{\mathcal{P}}(x_{n+1} \mid \mathcal{F}_n)$ are defined by (3.5) and (3.6) respectively, $\lambda = 0$ (no measure change) , $W \sim N(0,1)$ i.i.d. and x_n represents the value of a variable x at time $t = t_n$.

For a given set of observations $S = \{x_0, x_1, ..., x_N\}$, the joint log-likelihood function can be written for the log-spot price observations as follows:

$$L(\Psi) := -\sum_{i=1}^{N} \left(\frac{(x_{n+1} - \mathbb{E}^{\mathcal{P}}(x_{n+1} \mid \mathcal{F}_n))^2}{\mathrm{Var}^{\mathcal{P}}(x_{n+1} \mid \mathcal{F}_n)} + \log(\mathrm{Var}^{\mathcal{P}}(x_{n+1} \mid \mathcal{F}_n)) \right),$$

where the constant terms in the log likelihood function are ignored. Maximising $L(\Psi)$ will yield parameter estimates $\hat{\Psi} = (\mu_0, \kappa, \sigma, \Theta)$, by using an off-the-shelf solver routine such as *fminsearch* in Matlab. Then (3.12) can be used for predicting the future commodity prices. Note, however, that this model cannot be used for arbitrage-free prediction of futures prices, since the price of risk can not be estimated from the historical commodity prices alone.

3.3 Linear state space representation for latent commodity price models

For the models described in subsections 3.2.2 and 2.2.1, a state space representation is used, with a measurement equation based on the observable time series of futures prices and a discretized transition equation of log-spot commodity price, which is assumed to be unobservable. This allows us to use the Kalman filter to estimate the parameters by constructing and maximising a likelihood function, and to forecast the log-spot price when new futures price measurements become available. The state space equations for one factor with seasonality model in subsection 3.2.2 and two-factors with seasonality model in subsection 2.2.1 are provided below. The models without seasonality are obtained by setting the relevant parameters to zero.

3.3.1 One-factor model with seasonality

The state space equations corresponding to the model in section 3.2.2 can be written as

$$x_{n+1} = Bx_n + g + Rw_{n+1}, \tag{3.13}$$

$$y_n = A_n x_n + d_n + Qz_n, \tag{3.14}$$

where the state space model parameters may be expressed in terms of original model parameters as:

$$f(t_n) = c_1 + c_2 \sin(c_3 t_n + c_4), \tag{3.15}$$

$$B = e^{-\kappa\Delta},\ g = \frac{\mu_0}{\kappa}(1 - e^{-\kappa\Delta}), \tag{3.16}$$

$$R^2 = \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa\Delta}) + \frac{\Theta}{\kappa^2}\left(1 - e^{-\kappa\Delta}\right)^2,\ A_n = \begin{pmatrix} e^{-\kappa\Delta_n^1} \\ \vdots \\ e^{-\kappa\Delta_n^m} \end{pmatrix}, \tag{3.17}$$

$$d_n = \begin{pmatrix} \frac{\mu_0 - \lambda\sigma}{\kappa}(1 - e^{-\kappa\Delta_n^1}) + \frac{\sigma^2}{4\kappa}(1 - e^{-2\kappa\Delta_n^1}) + \frac{\Theta}{2\kappa^2}\left(1 - e^{-\kappa\Delta_n^1}\right)^2 + f(T_1) \\ \vdots \\ \frac{\mu_0 - \lambda\sigma}{\kappa}(1 - e^{-\kappa\Delta_n^m}) + \frac{\sigma^2}{4\kappa}(1 - e^{-2\kappa\Delta_n^m}) + \frac{\Theta}{2\kappa^2}\left(1 - e^{-\kappa\Delta_n^m}\right)^2 + f(T_m) \end{pmatrix}. \tag{3.18}$$

Here, $\Delta_n^i = \Delta_{t_n}^i = T_i - t_n$ for brevity of notation and m is the number of futures prices available at each t_n. $Q = \eta I_m$, where η is a scalar constant indicating the standard deviation of measurements and I_m is an $m \times m$ identity matrix. Recall that $\alpha \sim N(\mu_0, \Theta^2)$.

3.3.2 Two-factor model with seasonality

In this case, the parameters of (3.13)-(3.14) are:

$$f(t_n) = c_1 + c_2 \sin(c_3 t_n + c_4), \tag{3.19}$$

$$B = \begin{pmatrix} e^{-\kappa\Delta} & 0 \\ 0 & 1 \end{pmatrix}, \; g = \begin{pmatrix} 0 \\ \alpha_2\Delta \end{pmatrix}, \tag{3.20}$$

$$R^2 = \begin{pmatrix} \frac{\sigma_1^2}{2\kappa\Delta}(1 - e^{-2\kappa\Delta}) & \frac{\rho\sigma_1\sigma_2}{\kappa}(1 - e^{-\kappa\Delta}) \\ \frac{\rho\sigma_1\sigma_2}{\kappa}(1 - e^{-\kappa\Delta}) & \sigma_2^2\Delta \end{pmatrix}, \tag{3.21}$$

$$A_n = \begin{pmatrix} e^{-\kappa\Delta_n^1} & 1 \\ \vdots & \\ e^{-\kappa\Delta_n^m} & 1 \end{pmatrix}, \tag{3.22}$$

$$d_n = \begin{bmatrix} \frac{-\sigma_1\lambda_1+\rho\sigma_1\sigma_2}{\kappa}(1 - e^{-\kappa\Delta_t^1}) + (\alpha_2 - \sigma_2\lambda_2 + \frac{1}{2}\sigma_2^2)\Delta_n^1 + \frac{\sigma_1^2}{4\kappa}(1 - e^{-2\kappa\Delta_n^1}) + f(T_1) \\ \vdots \\ \frac{-\sigma_1\lambda_1+\rho\sigma_1\sigma_2}{\kappa}(1 - e^{-\kappa\Delta_n^m}) + (\alpha_2 - \sigma_2\lambda_2 + \frac{1}{2}\sigma_2^2)\Delta_n^m + \frac{\sigma_1^2}{4\kappa}(1 - e^{-2\kappa\Delta_n^m}) + f(T_m) \end{bmatrix}. \tag{3.23}$$

The brief outline on how this state space representation is used along with Kalman filter for parameter estimation can be found in section 2.7.1; see, e.g. [31] for more details on financial time series filtering using state space models.

3.3.3 Maximum Likelihood(ML) estimation

For the given log futures prices measurements $F = \{y_1, y_2, \ldots, y_N\}$ up to time t_N, Kalman filter can be applied to calibrate parameters of (3.15)-(3.18) and (3.19)-(3.23). The joint likelihood function for F can be written as follows:

$$\hat{L}(F) = p(y_1) \prod_{i=2}^{N} p(y_i \mid \mathcal{F}_{i-1}), \tag{3.24}$$

which, after substituting for joint probabilities and taking logarithms becomes

$$\log \hat{L}(F) = -\sum_{i=1}^{N} (\log \mid \Sigma_i \mid + v_i^T \Sigma_i^{-1} v_i), \tag{3.25}$$

where v_i, Σ_i are the innovations at time t_i and the covariance of innovations at time t_i respectively, and are as defined in section 2.7.1. The constant terms which do not depend on the model parameters are ignored. For a given vector-valued time series $\{y_1, y_2, \ldots, y_N\}$ and a vector of unknown model parameters Ψ, the optimisation problem can be stated as following:

$$\hat{\Psi} = \arg\max_{\Psi} \log \hat{L}(F), \tag{3.26}$$

$\hat{\Psi}$ is then used for forecasting experiments. Note that the implementation of maximum likelihood based calibration in this case is considerably more complicated than that in observable commodity price model described in section 3.2.3. However, it yields a richer class of models which use far more information in terms of futures prices.

3.4 Numerical experiments

3.4.1 Data

For an empirical study of the efficacy of the models in explaining the behaviour of commodity prices, the daily data is considered for Henry Hub Natural Gas, Gasoline and Light Sweet Oil, which includes:

- Henry Hub natural gas: spot commodity price and 12 different futures prices daily with maturities (5.15y, 5.24y, 5.32y, 5.41y, 5.49y, 5.58y, 5.66y, 5.75y, 5.84y, 5.92y, 6.01y 6.09y).
- Gasoline: spot commodity price and 7 different futures prices daily with maturities (2.65y, 2.74y, 2.82y, 2.92y, 3y, 3.08y, 3.17y).
- Light sweet oil: spot commodity price and 12 different futures prices daily with maturities (5.18y 5.27y, 5.35y, 5.44y, 5.52y, 5.61y, 5.69y, 5.78y, 5.87y, 5.96y, 6.05y, 6.13y).

The data was collected from Bloomberg for each of the commodities and separated into two panels as given in table 3.1. Panels were split for each data set into two parts: one for the model calibration and one for out-of-sample validation. Data was used as it is without any detrending and deseasonalization.

Statistics of the observed log-spot prices in table 3.2 shows that kurtosis values are slightly different from the normal distribution for all the three data

sets. The distributions are negatively skewed, which may possibly be explained by the seasonality factor.

Table 3.1: Data Panels for experiments

	Gasoline	
	Panel A	Panel B
In-sample	19.10.2007- 29.12.2008	28.04.2010- 01.07.2011
Out-of-sample	30.12.2008- 06.03.2010	02.07.2011- 10.09.2012
	Henry Hub Natural Gas	
	Panel A	Panel B
In-sample	29.11.2007- 05.02.2009	28.04.2010- 06.07.2011
Out-of-sample	06.02.2009- 19.04.2010	07.07.2011- 12.09.2012
	Light Sweet Oil	
	Panel A	Panel B
In-sample	19.11.2007- 28.01.2009	20.04.2010- 27.06.2011
Out-of-sample	28.01.2009- 09.04.2010	28.06.2011- 05.09.2012

Table 3.2: Log Spot Price Statistics

	Max	Mean	Variance	Skewness	Kurtosis
Gasoline	5.8478	5.5763	0.0267	-0.4910	2.0788
Natural Gas	1.6465	1.2623	0.0571	-0.7273	2.3704
Light Sweet Oil	4.9787	4.4244	0.0709	-0.9928	4.2581

3.4.2 Methodology

Choice of models

The relative performance is compared in terms of out of sample commodity price forecasting and model complexity for the following models:

- OCP: Observable commodity price model (section 3.2.3);
- OF: One-factor model (section 3.2.2);
- OFS: One factor model with seasonality (section 3.2.2);
- TF: Two-factor model (section 2.2.1, case $f(t) = 0$);
- TFS: Two-factor model with seasonality (section 2.2.1).

The calibration routine was implemented in Matlab using its in-built solver, *viz fminsearch* which uses the Nelder-Mead algorithm. Note that four models OF, OFS, TF, TFS each contain *two* sources of randomness. In one factor models(OF and OFS), the second source is the random long-run mean, while two-factor models have two Wiener processes. Our numerical experiments allow us to test whether adding a richer description of a second source of uncertainty (as in the TF and TFS case) adds value over a more parsimonious, but restricted description as in OF and OFS.

Choice of measures of comparison

For comparison of model performance, the sample mean of the relative absolute error (MRAE) and root mean square error (RMSE) is considered as measures of errors for commodity price models:

$$MRAE = \frac{1}{N}\sum_{i=1}^{N}\frac{|x_i - \hat{x_i}|}{x_i},$$

$$RMSE = \sqrt{\sum_{i=1}^{N}\frac{(x_i - \hat{x_i})^2}{N}}$$

where x_i is the observed commodity price at time i and $\hat{x_i}$ is the best estimate of log spot price at time i and N is the number of observations. These measures of error will be used for out-of-sample data, in each of the two data panels and for each of the three commodities. For in-sample comparison of the increase in explanatory power relative to increase in model complexity, the Akaike Information criterion (AIC) [3] is also used, which is defined as follows:

$$AIC(\hat{\Theta}) = (-2)\log(maximum \quad likelihood) + 2n$$

where n is number of elements in the vector $\hat{\Theta}$, which minimizes log-likelihood value. The AIC value represents the quality of the models by penalising the log-likelihood values with the number of added parameters, in this case the model with the smallest AIC value has a better fitness. For small data sets($K/n < 40$), where K is the number of data points used, one can use a second-order criterion (denoted by AIC_c):

$$AIC(\hat{\Theta})_c = (-2)\log(maximum \quad likelihood) + 2n + \frac{2n(n+1)}{K-n-1}.$$

The results of calibration and model comparison is summarized in the next section.

3.4.3 Results

The results on calibration and forecasting of the models described above to NYMEX data are now discussed to highlight the behaviour of the models with different types of commodities, for the two different data panels mentioned earlier in section 3.4.1.

Light Sweet Oil

The estimated parameters of the OCP model are presented in table 3.6. The estimated parameters for OF and OFS models from calibration are presented in table 3.4. For the Panel A, the values of Θ are relatively small, which indicates that the long-run mean of the log spot price process can be assumed to be constant. The value of σ in the Panel A is twice as high than in the Panel B. This also seems to affect the seasonality amplitude c_2 in the Panel A, which is smaller than for the Panel B. Parameters of two-factor models are presented in table 3.5. One can see that values of volatility σ_1 are dominating seasonality amplitude ζ in both cases for Panels A and B. For all the cases, one can see the low correlation ρ between the two Wiener processes.

For light sweet oil data-sets, table 3.3 shows that increasing of the model complexity (i.e. adding more parameters) does not significantly increase the quality of model fitness and the AIC_c value is the highest for OF. The out-of-sample prediction performance of all the models is similar, except for the RMSE value of the OCP model for panel A, which is significantly higher than that of the other models. The values of Θ for one factor models are small relative to μ_0, indicating that a second source of randomness does not add value in this case.

Figures 3.2 and 3.3 compare the four Kalman filter based models when it comes to out-of-sample performance in forecasting one day ahead futures prices, where MRAE is considered for 11 different futures as computed over the entire out-of-sample data set. These figures show that, in general, the OF model does a fairly good job of predicting futures prices, being better than OFS and TF models in both the cases as well as TFS model in one case.

Table 3.3: Light Sweet Oil Out-of-Sample errors, in-sample Likelihood and AIC

Panel A					
Model	parameters	MRAE	RMSE	Log-Likelihood	AIC_c
OCP	4	1.32	5.21	1024	-2040
OF	7	2.12	1.644	15698	-31367
OFS	11	2.12	1.645	15695	-31346
TF	10	2.12	1.645	14452	-28862
TFS	14	2.11	1.635	15605	-31153
Panel B					
Model	parameters	MRAE	RMSE	Log-Likelihood	AIC_c
OCP	4	1.44	1.74	2063.36	-1725
OF	7	1.44	1.743	15817	-31606
OFS	11	1.44	1.763	15816	-31587
TF	10	1.43	1.745	15538	-31036
TFS	14	1.44	1.743	15750	-31443

Table 3.4: Light Sweet Oil Parameter estimates for One-Factor models

	Panel A		Panel B	
	OF	OFS	OF	OFS
μ_0	-0.0658	-0.0905	0.4441	1.1924
Θ	0.0000	0.0000	0.0243	0.0277
κ	0.0376	0.0378	0.1136	0.1031
σ	0.4309	0.4317	0.2980	0.2886
η	0.0100	0.0100	0.0100	0.0100
P_0	0.0472	20.5919	0.0270	8.1568
λ	0.3874	0.3714	-0.0523	-1.3620
c_4		-0.3426		-0.3484
c_2		0.1063		-1.0350
c_3		-0.1990		0.3840
c_1		1.2237		-0.1504

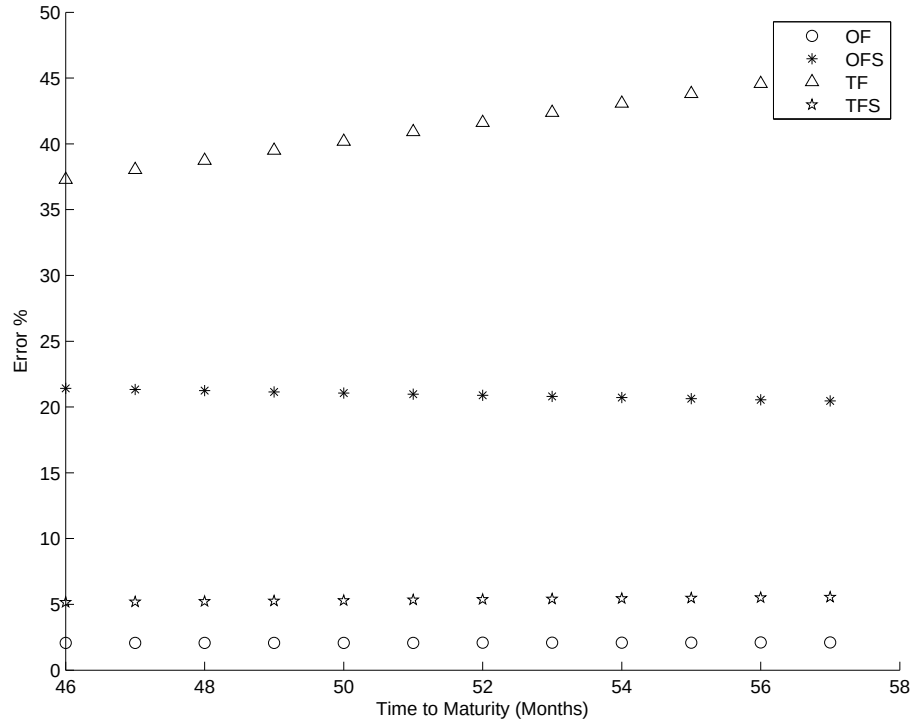
Table 3.5: Light Sweet Oil Parameter estimates for Two-Factor models

	Panel A		Panel B	
	TF	TFS	TF	TFS
α_2	-0.2573	0.3820	0.2496	0.0601
κ	0.0100	0.0100	0.0526	0.1627
σ_1	0.1294	0.3575	0.1494	0.3293
σ_2	0.0242	0.1555	0.0010	0.0008
η	0.0100	0.0100	0.0100	0.0100
λ_1	-0.8669	-1.3187	-2.8441	-0.3547
λ_2	-0.0029	-0.0035	0.0023	0.0027
ρ	-0.0015	-0.0002	0.0011	0.0009
P_{01}	2.5080	8.5736	1.5811	1.1973
P_{02}	2.3636	0.0001	0.5047	0.0001
c_4		0.4503		0.0519
c_2		-0.0082		-0.0077
c_3		0.0077		-0.0013
c_1		0.0032		-0.0014

Table 3.6: Parameter Values for OCP model

	α	κ	σ	Θ
Panel A	-0.6815	-1e10	0.6219	0.0419
Panel B	12.40	2.76	0.31	0.08

Figure 3.2: Light Sweet Oil daily Futures Predictions Errors (MRAE): Panel A

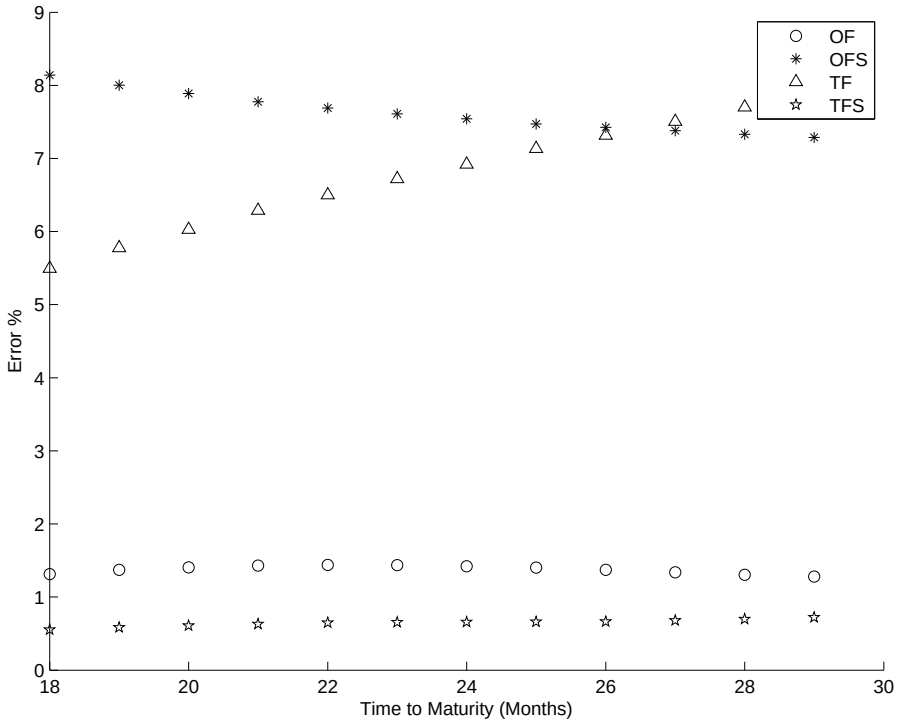


Henry Hub Natural Gas

Two-factor models parameters presented in table 3.9. For all the panels, the values of κ, σ_1 and σ_2 are quite low. From Panel A, one can see that adding a seasonality factor to the model increases price of risk λ_1, and for Panel B it stays on the same level. For both the panels, models without seasonality have weak correlation ρ between factors, while addition of the seasonality factor increases the correlation. The values of the amplitudes c_2 are on the same level, but have a different sign. This effect appears from different data ranges used for the parameter estimation procedure.

Table 3.7 for Panels A and B shows that the OF model has higher AIC values than other models, although the difference between the models is modest (apart from the error values for OCP in data panel B). Note that, for two-factor models, the rate of mean reversion is very small, as is the correlation between the two-factors. Figures 3.4 and 3.5 indicate the one step prediction

Figure 3.3: Light Sweet Oil daily Futures Predictions Errors (MRAE): Panel B



performance of the four filtering-based models. Qualitatively, it is clear that all the models perform very poorly for data panel A, while the OF model may be deemed to be acceptable for data panel B. The poor out-of-sample performance of models in panel A can be explained by the collapse of average natural gas spot prices by 55% during 2009 due to a sharply reduced demand (owing to recession) coupled with an increased US domestic gas output; see, *e.g.* [68].

New York Gasoline

The parameters for one factor models are presented in table 3.12. Panel A, shows that deviation of the long-run mean θ is zero, while the volatility σ is high. Moreover, increasing of the complexity of the model doesn't affect the price of risk values λ. The parameters for the two-factor models are shown in table 3.13. Panel A shows a strong correlation between factors for TFS models as compare to the TF model, this change also drives σ_2 values to the lower

Table 3.7: Henry Hub Natural Gas Out-of-Sample errors, in-sample Likelihood and AIC

Panel A					
Model	parameters	MRAE	RMSE	Log-Likelihood	AIC_c
OCP	4	3.22	0.19	1821.35	-1813
OF	7	3.24	0.186	9845	-19662
OFS	11	3.23	0.186	9845	-19646
TF	10	3.25	0.187	9787	-19533
TFS	14	3.26	0.187	9802	-19546
Panel B					
Model	parameters	MRAE	RMSE	Log-Likelihood	AIC_c
OCP	4	2.84	0.10	940	-1872
OF	7	2.25	0.086	10870	-21712
OFS	11	2.25	0.086	10873	-21702
TF	10	2.26	0.086	10767	-21493
TFS	14	2.26	0.086	10805	-21553

Table 3.8: Henry Hub Natural Gas parameter estimates for One-Factor models

	Panel A		Panel B	
	OF	OFS	OF	OFS
μ_0	-0.1273	-0.3446	0.1108	-0.1418
Θ	0.0440	0.0440	0.0000	0.0438
κ	0.0100	0.0100	0.3233	0.0100
σ	0.1892	0.1894	0.2430	0.1891
η	0.0374	0.0374	0.0283	0.0374
P_0	1.4369	0.0189	0.6855	0.8556
λ	-0.9853	-0.0085	1.5314	-1.0111
c_4		0.0002		2.1105
c_2		0.0153		-2.1321
c_3		0.0058		-0.0020
c_1		3.2133		0.4676

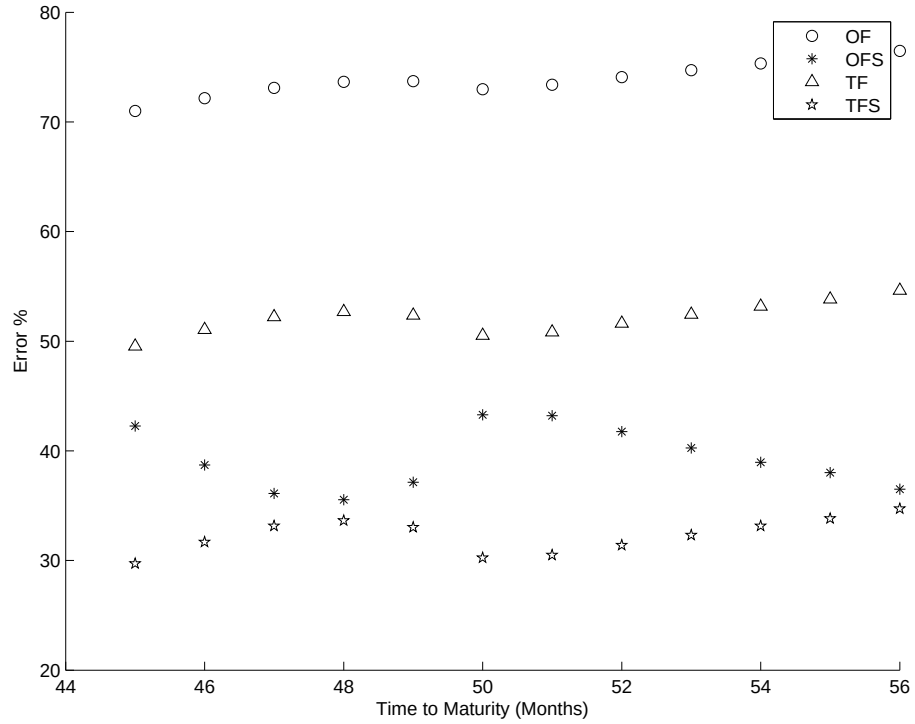
Table 3.9: Henry Hub Natural Gas parameter estimates for two-Factor models

	Panel A		Panel B	
	TF	TFS	TF	TFS
α_2	0.1693	0.3330	-0.1893	-0.1158
κ	0.0100	0.0100	0.0100	0.0100
σ_1	0.1918	0.1579	0.1058	0.0429
σ_2	0.0057	0.1261	0.0010	0.1499
η	0.0376	0.0370	0.0289	0.0283
λ_1	-1.6141	-3.2646	1.3859	1.8230
λ_2	0.0006	-0.0518	-0.0019	-0.2384
ρ	0.0038	0.0991	0.0006	0.0329
P_{01}	0.0001	0.0001	1.7031	0.2694
P_{02}	8.5445	9.0246	0.0001	0.0034
c_4		0.0830		-2.7233
c_2		0.1524		-0.2381
c_3		-0.3997		-0.6534
c_1		0.6816		0.6592

Table 3.10: Natural Gas parameter Values for OCP model

	α	κ	σ	Θ
Panel A	-0.38	4.07E-10	0.46	1.47
Panel B	18.7564	13.0110	0.2273	5.0084

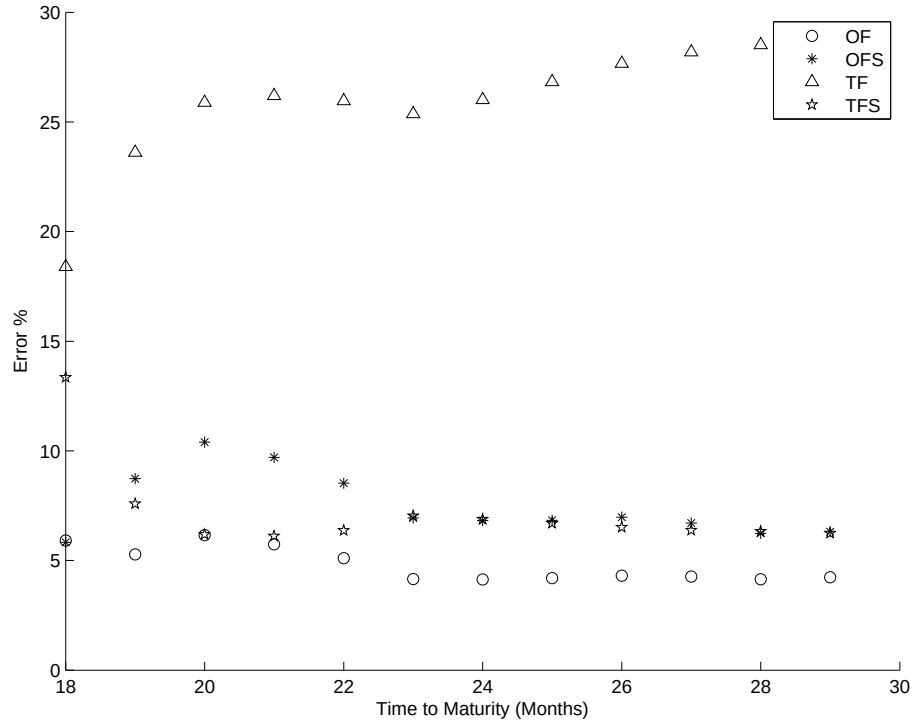
Figure 3.4: Henry Hub Natural Gas daily Futures Predictions Errors (MRAE): Panel A



levels in the TFS model. Increasing the complexity of models also increases the price of risk λ_2. Panel B, yields different parameter values as compared to Panel A, owing to a change in market conditions. TF and TFS model have the same level of σ_1, while σ_2 is increasing, when more degrees of freedom are added. It also can explain the low score for amplitude c_2, since the seasonality factor is overtaken by higher volatility.

In-sample and out-of-sample comparison of models for New York Gasoline is presented in table 3.11. For both the panels A and B, AIC_c values for the OFS model show its dominance over the other models, although the OF model seems to perform best out-of-sample when it comes to predicting the commodity price. The OCP model also seems to perform well out-of-sample in these two data sets. The picture seems to be very different, however, when one looks at the errors in predicting the futures prices. Figures 3.6 and 3.7 show that TFS model outperforms all the other models in both the data sets, when it comes to predicting futures prices (rather than predicting the price of

Figure 3.5: Henry Hub Natural Gas daily Futures Predictions Errors (MRAE): Panel B



the commodity itself).

3.5 Summary

In this chapter, a subclass of the commodity price models was implemented. The implementation is based on extensions of [56]. A comprehensive set of numerical experiments have been carried out to compare different models, on three different commodities and for two data sets for each commodity. The summary of the conclusions from the numerical experiments is provided below:

- Filter-based models do seem to perform well as compared to the simplest, observable commodity price model. They also allow us to make arbitrage-free forecasts of futures prices.

- For light sweet oil, using the one factor model with no seasonality seems to give a good compromise between model complexity and out-of-sample

Table 3.11: Gasoline Out-of-Sample errors, in-sample Likelihood and AIC

Panel A					
Model	parameters	MRAE	RMSE	Log-Likelihood	AIC_c
OCP	4	2.25	4.68	1733	-1725
OF	7	2.23	4.636	7042	-14070
OFS	11	2.21	4.604	7271	-14518
TF	10	2.30	4.764	7032	-14044
TFS	14	2.30	4.779	7039	-14049
Panel B					
Model	parameters	MRAE	RMSE	Log-Likelihood	AIC_c
OCP	4	1.32	5.21	1024	-2040
OF	7	1.33	5.229	7807	-15600
OFS	11	2.01	7.146	8269	-16515
TF	10	1.35	5.297	7671	-15320
TFS	14	1.36	5.314	7678	-15327

Table 3.12: Gasoline parameter estimates for One-Factor models

	Panel A		Panel B	
	OF	OFS	OF	OFS
μ_0	-0.4087	-0.0182	0.2758	4.2419
Θ	0.0000	0.0000	0.0739	0.0394
κ	0.0100	0.0100	0.0100	0.0100
σ	0.3706	0.3690	0.2360	0.2622
η	0.0173	0.0153	0.0122	0.0100
P_0	0.0402	22.8173	0.0115	0.5177
λ	0.8409	0.8065	-1.2990	-17.6784
c_4		1.2627		-0.0793
c_2		-0.8086		0.6721
c_3		-0.6074		-1.0127
c_1		-0.3522		0.0960

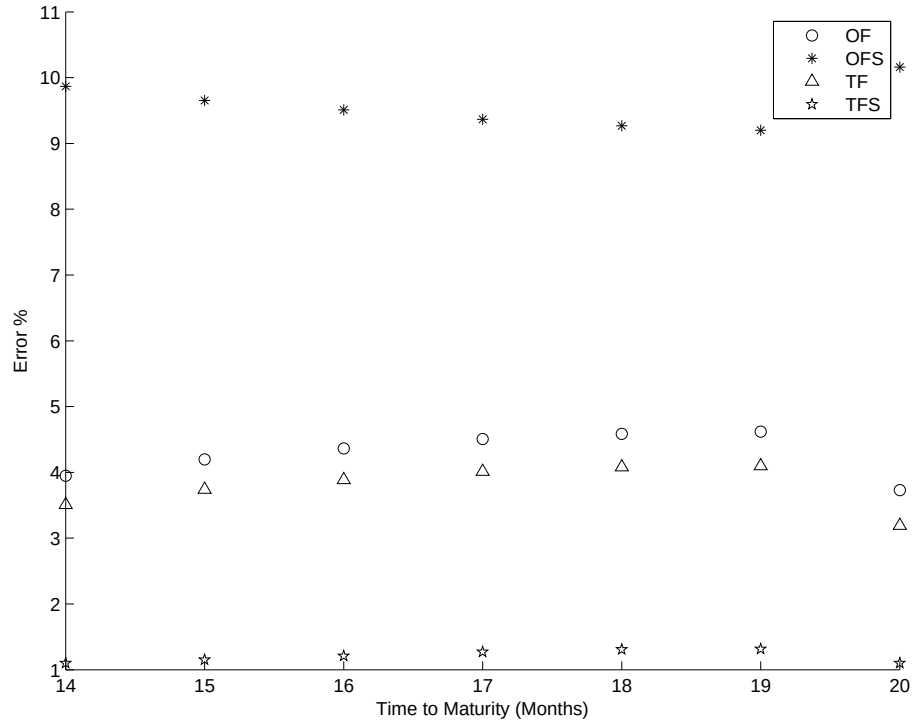
Table 3.13: Gasoline parameter estimates for two-Factor models

	Panel A		Panel B	
	TF	TFS	TF	TFS
α_2	-1.3354	-1.4164	0.7497	0.8348
κ	0.0100	0.0100	0.0100	0.0100
σ_1	0.3272	0.3628	0.2404	0.2042
σ_2	0.2368	0.0121	0.0056	0.1625
η	0.0172	0.0171	0.0129	0.0129
λ_1	3.6756	3.5619	-2.8713	-3.7190
λ_2	0.0498	0.6469	0.0017	-0.1002
ρ	-0.1404	0.9905	0.0031	-0.1398
P_{01}	42.7547	0.0001	0.4134	0.0001
P_{02}	0.0001	0.0001	0.0001	0.0001
c_4		-1.5022		6.3129
c_2		-0.0071		-0.0171
c_3		0.0323		-0.2191
c_1		0.1479		-0.2690

Table 3.14: Gasoline parameter Values for OCP model

	α	κ	σ	Θ
Panel A	-0.80	1.29E-10	0.53	0.33
Panel B	5.5635	0.9774	0.3085	1.8987

Figure 3.6: New York Gasoline daily Futures Predictions Errors (MRAE)



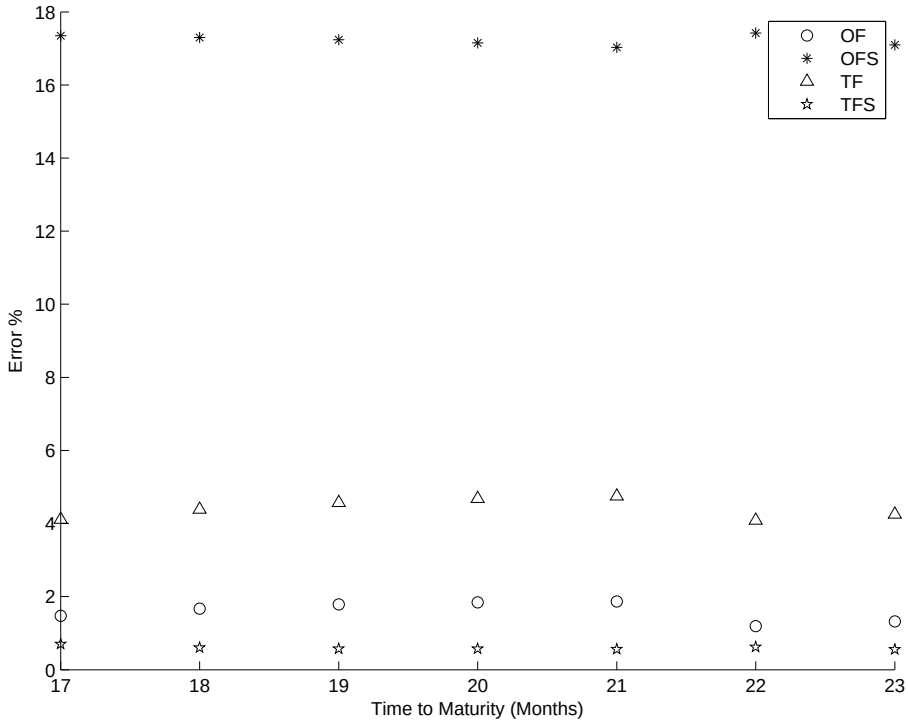
performance, both in terms of commodity price prediction as well as futures price prediction.

- For natural gas data sets, all the models calibrated from 2007-2009 data (for data panel A) gave extremely poor performance when predicting futures prices between 2009-2010. This may be due to the spot gas price collapse in 2009. Clearly, there is a need to re-calibrate the model frequently during periods when market is under stress. For data panel B, the one factor model was again the best performer.

- For the New York gasoline, all the four filtering-based models give very similar performance when considering commodity price prediction, but the two-factor with seasonality model clearly outperforms the other models when it comes to predicting the futures prices.

All in all, different types of energy commodities require different models. It is obvious that it is impossible to construct the model which will explain

Figure 3.7: New York Gasoline daily Futures Predictions Errors (MRAE)



the behaviour of all of the energy market. On the other hand, it is possible to find a model which will suits best to the specific commodity.

Chapter 4

A fast calibrating volatility model for option pricing

4.1 Introduction

The central assumption of the celebrated Black-Scholes formula for European option pricing is that the volatility of the underlying asset is constant [13]. This is known to be untrue in practice. The observed prices of liquid options on the same underlying, for a given set of maturities and strikes, imply different volatilities under the Black-Scholes formulation. Modelling the future evolution of the volatility of the underlying asset, which is consistent with the observed option prices, is obviously essential to price illiquid securities on the same underlying asset. The topic of suitable volatility models which provide a consistent match with the observed prices has resulted in extensive literature over the past few decades.

There are two broad classes of volatility models: local volatility models and stochastic volatility models. Note that this is a rather imprecise taxonomy, but it will be sufficient for the purpose of this chapter. The former class of models does not have an additional source of uncertainty (apart from the sources of uncertainty in the underlying) incorporated in the volatility model and the volatility is assumed to be a deterministic function of the current underlying price and time. Examples of this type of models include the models proposed by Dupire [30], Derman and Kani [27] and Alexander [5]. In contrast, stochastic volatility models include an extra source (or sources) of randomness and provide more flexibility in modelling the dynamics of the volatility surface.

Significant models in this class, with an emphasis on option pricing, include those proposed by Hull and White [49], Merton [60], Heston [44], Bates [11], Kou [53], Duffie et al [29] and Carr et al [17]. Bakshi et al [9] have compared a variety of stochastic volatility models in terms of their pricing and hedging performance. Heston as well as Bates model yields semi-closed form solutions in terms of Fourier transform of European option price and are hence amenable to relatively easy calibration to market data. Gatheral [36] and Javaheri [51] provide comprehensive reviews of the development of volatility models.

In this chapter, a new method is proposed for modelling volatility as implied by option prices. In the new model, the volatility is represented as a deterministic function of time, with its *level* being a random variable on positive support. The proposed volatility model offers the following benefits:

- It provides a very simple approximate pricing function for calibrating the model from option price data. In the experiments performed, the evidence suggests that the proposed model requires only around 1% of the computational time as the Heston model or the Bates model for calibration, on the same hardware.

- In fifteen different data sets tested for three different indices and using two different methods of measuring the pricing error, the proposed model is shown to be extremely competitive in terms of accuracy with the popular existing stochastic volatility models.

- When calibrated from the same data-set, the proposed model also yields prices for path-dependent payoffs which are in the same range as the Heston model and the Bates model. This is important since the prices of illiquid payoffs are non-unique under stochastic volatility and any new model which gives significantly different prices from the established models is unlikely to be accepted by the industrial community.

- When calibrated from the same data-set and using the same numerical method, the proposed model yields option price sensitivity parameters which are very close to those found for one of the two benchmark models, for most data-sets. Option sensitivities (or Greeks) are important for risk monitoring and hedging purposes and the numerical experiments show that hedging using the new model is unlikely to provide significantly different results than hedging using the Heston model.

Note that, apart from the Bates model and the Heston model, several other analytically tractable options exist for modelling volatility (as mentioned earlier). The purpose here is to establish that the new model yields accuracy comparable to some of the popular existing models, while being significantly easier to calibrate, and easier to simulate from, than those models. Hence the benchmark comparison has been restricted to the two aforementioned models.

The rest of this chapter is organized as follows. In the next section, the new volatility model is presented. Section 4.3 on numerical experiments is split into three subsections: section 4.3.1 outlines the data used, section 4.3.2 explains the methodology employed in comparing the performance of different models and lastly section 4.3.3 provides the results and a discussion. Section 4.4 summarizes this chapter. Finally section 4.5 contains a visual representation of the results.

4.2 High-Order Moments Stochastic Volatility model

The basic idea of the model is now introduced. Recall that, by definition, European call option is a right to buy an asset at maturity time T for a strike price K. For a non-dividend paying stock, its price at time t is given by discounted expectation of the terminal pay-off:

$$C_t = e^{-r(T-t)}\mathbb{E}[(S_T - K, 0)^+].$$

Under the Black-Scholes framework with constant volatility, this discounted expected value is given by

$$\begin{aligned} C_{BS} &= S_t N(d_1) - e^{-r\tau} K N(d_2), \\ d_1 &= (\sigma\sqrt{\tau})^{-1}[\log(S_t/K) + (r + \sigma^2/2)\tau], \\ d_2 &= d_1 - (\sigma\sqrt{\tau}), \end{aligned}$$

where r is the constant risk-free rate, σ is the volatility, $N(x)$ is the standard normal cumulative distribution function and $\tau = T - t$ is the time to maturity. The derivation of Black-Scholes price also assumes that short-selling as well as trading in continuous time is possible. As outlined before in section 2.5, one of

the simplest frameworks to introduce a stochastic component in the volatility is to consider a Hull-White type model of the asset price process [49]:

$$dS_t = \mu S_t dt + \sqrt{v_t} S_t dW_t^1, \tag{4.1}$$

$$dv_t = f_1(t, v_t)dt + f_2(t, v_t)dW_t^2, \tag{4.2}$$

where W_t^1 and W_t^2 are uncorrelated Wiener processes and f_1, f_2 are smooth functions bounded by linear growth such that v_t remains non-negative almost surely. [49] shows that the price of European vanilla call option at time 0, for a time to maturity τ can be derived as the expectation of Black-Scholes price with respect to the variance rate (see section 2.5 for the proof):

$$C_{EUR} = \mathbb{E}\left[C_{BS}\left(\frac{1}{\tau}\int_0^\tau v_t dt\right)\right] \tag{4.3}$$

where $C_{BS}(x)$ denotes the Black-Scholes price evaluated at variance x. The above formula is independent of the exact process followed by v_t (under normal assumptions about $t-$ continuity and uniqueness). Denoting the variance rate $\frac{1}{\tau}\int_0^\tau v_t dt$ by $\bar{V}_\tau$ and assuming that the moments in question exist, the right hand side of (4.3) can be expanded around $\mathbb{E}(\bar{V}_\tau)$ in Taylor series as:

$$C_{EUR} \approx C_{BS}(\mathbb{E}(\bar{V}_\tau)) + \sum_{i=2}^{M} \frac{\partial^i C_{BS}}{\partial \bar{V}_\tau^{\,i}} \frac{\mathbb{E}(\bar{V}_\tau - \mathbb{E}(\bar{V}_\tau))^i}{i!}, \tag{4.4}$$

where the partial derivatives are evaluated at $\mathbb{E}(\bar{V}_\tau)$. The aim of this chapter is to construct a process for v_t for which the right hand side of the above equation is easy to evaluate (for a reasonably large M), while remaining sufficiently flexible to fit the observed option prices. Note that truncating after the first term will mean that prices of options with *all* strikes for a *fixed* time to maturity should be the same, which is obviously nonsense. This illustrates the need for non-zero higher moments for $\bar{V}_\tau$ (and hence the need for randomness in volatility) in an intuitively simple fashion.

Without loss of generality, let $t = 0$ be the current time and let $t_0 > 0$ be an arbitrary time which is less than the shortest time to maturity of any derivative product which we want to price using the new model. We will allow the diffusion term in the volatility process of (4.2) to be non-zero only within $[0, t_0)$. This will allow us to use a *single random variable*, rather than an evolving

random process, to model the randomness in volatility when pricing securities at time $t = 0$, whose payoffs are beyond t_0. Note that option pricing models are always used for pricing securities with finite, rather than infinitesimal, time to maturity. Further, t_0 itself does not appear in the pricing formulae (only an integrated variance term does, as we shall see) and can be assumed to be arbitrarily small. Next, assume that v_t in (4.2) is governed by the following, specific stochastic process:

$$dv_t = (\mu_t dt + \gamma_t dW_t^2)v_t, \tag{4.5}$$

where μ_t is a positive deterministic and integrable function, γ_t is a positive deterministic function which is piecewise continuous, with $\gamma_t = 0,\ t > t_0$ and W_t^2 is a standard Wiener process uncorrelated with W_t^1. Using Itô's lemma, it is straightforward to show that

$$v_t = \exp\left(\int_0^t \mu_s ds\right)\zeta_t,$$

where ζ_t is a log-normal process with unit mean and a constant variance for $t > t_0$. In particular,

$$\mathrm{Var}(\zeta_t) = \left(\exp\left\{\int_0^{t_0} \gamma_s^2 ds\right\} - 1\right),\ t > t_0.$$

Henceforth it is assumed that $t > t_0$ holds. Let $k = \sqrt{\mathbb{V}ar(\zeta_t)}$. Then the third and the forth centered moments of ζ_t, m_3 and m_4 respectively, can be expressed as:

$$m_3 = k^4(3 + k^2), \tag{4.6}$$

$$m_4 = k^4\{(1 + k^2)^4 + 2(1 + k^2)^3 + 3(1 + k^2)^2 - 3\}. \tag{4.7}$$

The standard deviation k of the lognormal random variable ζ_t is parametrised directly, with no reference to γ_t or t_0. Finally, $\exp(\int_0^t \mu_s ds)$ is parametrised as

$$\exp\left(\int_0^t \mu_s ds\right) = \hat{\sigma}_0^2 e^{-\lambda t} + \hat{\sigma}_1^2 \lambda t e^{-\lambda t} + \hat{\sigma}_2^2,$$

where $\hat{\sigma}_0, \hat{\sigma}_1, \hat{\sigma}_2, \lambda$ are scalar parameters. This allows the following parametriza-

tion as

$$v_t = \zeta_t(\hat{\sigma}_0^2 e^{-\lambda t} + \hat{\sigma}_1^2 \lambda t e^{-\lambda t} + \hat{\sigma}_2^2),\ \zeta_t \sim LN(1, k^2),\ t > t_0. \tag{4.8}$$

Along with (4.1), (4.8) completely specifies the new pricing model within the chosen pricing measure, which is implicitly specified by the data used for calibration. The model defined here is called high order Moments-based Stochastic Volatility (MSV) model, since it is based on the use of higher order moments of the aforementioned random variable. With this definition of v_t:

$$\bar{V}_\tau := \frac{1}{\tau}\int_0^\tau v_t dt = \underbrace{\zeta_t \left(\frac{\hat{\sigma}_0^2 + \hat{\sigma}_1^2}{\lambda\tau} + \hat{\sigma}_1^2 + \frac{\hat{\sigma}_2^2 - \hat{\sigma}_1^2}{1 - e^{-\lambda\tau}} \right)(1 - e^{-\lambda\tau})}_{Q_\tau}, \tag{4.9}$$

where Q_τ is a deterministic function. Q_τ is actually the equation for Nelson-Siegel [65] spot rate curve used in interest rate modelling. While the application discussed here is unrelated to modelling interest rates, this parametrization is chosen for its known ability to represent a variety of relevant shapes of term structure (both concave and convex), with a suitable choice of parameters. Since, European option price for any $\tau > 0$ is a smooth function with respect to $\bar{V}_\tau$, one can apply Taylor series expansion to the Black-Scholes option price C_{BS} around a point $\mathbb{E}(\bar{V}_\tau) = Q_\tau$:

$$\begin{aligned} C_{EUR}(\bar{V}_\tau) \approx\ & C_{BS} + \frac{\partial^2 C_{BS}}{\partial \bar{V}_\tau^2}\frac{\mathbb{E}(\bar{V}_\tau - Q_\tau)^2}{2} + \frac{\partial^3 C_{BS}}{\partial \bar{V}_\tau^3}\frac{\mathbb{E}(\bar{V}_\tau - Q_\tau)^3}{6} \\ & + \frac{\partial^4 C_{BS}}{\partial \bar{V}_\tau^4}\frac{\mathbb{E}(\bar{V}_\tau - Q_\tau)^4}{24}, \end{aligned} \tag{4.10}$$

where C_{BS} and its partial derivatives are evaluated at $\bar{V}_\tau = Q_\tau$. These partial derivatives for a European call option can be evaluated as:

$$\begin{aligned} \nu :=\ & \frac{\partial C_{BS}}{\partial \bar{V}_\tau} = K e^{-r\tau}\phi(-d_2)\sqrt{\tau}, \\ & \frac{\partial^2 C_{BS}}{\partial \bar{V}_\tau^2} = \nu\frac{d_1 d_2}{Q_\tau}, \\ & \frac{\partial^3 C_{BS}}{\partial \bar{V}_\tau^3} = \frac{-\nu}{Q_\tau^2}\left[d_1 d_2(1 - d_1 d_2) + d_1^2 + d_2^2\right], \\ & \frac{\partial^4 C_{BS}}{\partial \bar{V}_\tau^4} = \nu\frac{12 d_1 d_2 + 3\tau Q_\tau^2(1 - d_1 d_2) - d_1^2 d_2^2(9 - d_1 d_2)}{Q_\tau^3}, \end{aligned} \tag{4.11}$$

with $d_1 = \frac{\log(S_0/K)+(r+Q_\tau^2/2)\tau}{Q_\tau\sqrt{\tau}}$, $d_2 = d_1 - Q_\tau\sqrt{\tau}$ and
$\phi(x) = (\sqrt{2\pi})^{-1}\int_0^x e^{-0.5u^2}du$. Similar expressions can easily be derived for an approximation to the price of a European put option.
The first four moments of $\bar{V}$ can be rewritten as the following:

$$\begin{aligned}
\mathbb{E}(\bar{V}_\tau) &= Q_\tau, \\
\mathbb{E}(\bar{V}_\tau - Q_\tau)^2 &= k^2 Q_\tau^2, \\
\mathbb{E}(\bar{V}_\tau - Q_\tau)^3 &= k^4(3+k^2)Q_\tau^3, \\
\mathbb{E}(\bar{V}_\tau - Q_\tau)^4 &= k^4\{(1+k^2)^4 + 2(1+k^2)^3 + 3(1+k^2)^2 - 3\}Q_\tau^4. \qquad (4.12)
\end{aligned}$$

Equations (4.8)-(4.10) together with equations (4.11)-(4.12) define the approximation to the new option pricing model.

Along with the parameters $\hat{\sigma}_0, \hat{\sigma}_1, \hat{\sigma}_2, \lambda$ which appear in Q_τ, the parameter k which characterises the distribution of ζ_t completes the set of parameters for the new volatility model specification.

A few remarks on this model are in order.

- Empirical experiments showed that a third or a fifth order Taylor series approximation, in place of the fourth order approximation used here, makes very little difference. However, using $k = 0$ leads to very poor fits on calibration, again indicating that randomness is necessary to model the volatility dynamics adequately.

- Zero correlation is assumed between the sources of randomness in the underlying and the volatility, and there is no risk premium attached to the randomness in volatility. However, the choice of a simpler volatility model seems to provide a fit which is quite competitive in terms of accuracy when compared to models with non-zero correlation, at a small fraction of calibration cost, over a large number of data sets. The admittedly limited evidence indicates that choosing a sufficiently flexible parametrised function of time can compensate at least partially for *not* modelling the correlation between the volatility and the price of the underlying.

4.3 Numerical Experiments

4.3.1 Data Specification

For calibration and validation of the new model, the option price data {Strike price, Maturity, Implied Volatility, Bid, Ask and underlying values on the date of reading} was obtained from Bloomberg Option Monitor (OMON). Implied risk free rates were calculated using implied volatilities and option prices by simple nonlinear least squares, for each maturity. European call options were chosen with a minimum of 30 days to maturity and up to 3 years to maturity, with strike prices to be both in-the-money and out-of-the money values. The total data consisted of closing option prices on 3 different stock indices {S&P500, FTSE 100 and DAX} on five different days {01 November 2012, 26 November 2012, 25 July 2013, 26 July 2013, 29 July 2013}, with 100 options for each index and day. This gave a total of 15 data sets (one for each index and each day), from two different years, with 100 prices in each data set. [1]

4.3.2 Methodology

To calibrate and validate the models (the Heston model, the Bates model and the MSV model), the option prices were randomly separated with proportion 80 and 20 percent for in-sample and out-of-sample model evaluation respectively, within each of the fifteen data-sets. The in-sample data was used for calibration as well as validation and the out-of-sample data was used for validation only.

For calibration, the following minimization problem was solved for each of the three models:

$$\min_{\Theta} \sum_{i=1}^{N} \frac{|C_i^{market} - C_i^{model}(\Theta)|^2}{|Bid_i - Ask_i|^4},$$

where Θ is the vector of parameters, $C_i^{model}(\Theta)$ is the price given by the model parametrised by Θ, N is the number of options in the in-sample data and Bid_i, Ask_i are closing bid and ask prices of the i^{th} option, respectively. C_i^{market} is the market price of the i^{th} option which is obtained as an arithmetic average of Bid_i, Ask_i for each option. The choice of weight, which is the inverse of (op-

[1] Note that numerical experiments have been carried out over more data-sets and the results presented here are deemed to be representative.

tion price spread)[4], under-emphasizes any illiquid options during calibration. Three different powers of bid-offer spread were tried $(1, 2, 4)$ for the choice of weight and 4 seems to offer the best fit for all of the models. Calibration was done using Matlab 2012b on a Windows 8 laptop, with Intel i7 processor and 8 Gb memory. As mentioned earlier, Heston stochastic volatility model and Bates, i.e. stochastic volatility with jumps model (SVJ) [11] are used as benchmarks for option pricing models. For Heston and Bates models, 8192 point FFT was used in approximating the option price evaluation integral.

The calibrated models are compared with each other in three different ways:

1. For each in-sample and out-of-sample data set after calibration (30 data-sets in all - with each of 15 data-sets split into in-sample and out-of-sample subsets), the two commonly used error metrics were used, *viz* Mean Relative Absolute Error (MRAE) and Root Mean Square Error (RMSE). Further, since computational speed is one of the main selling points of the new method. The computational time for model calibration was compared between the benchmark models and the new model. The two error metrics are defined below:

$$MRAE = \frac{1}{N}\sum_{i=1}^{N}\frac{|C_i^{market} - C_i^{model}|}{C_i^{market}},$$

$$RMSE = \sqrt{\sum_{i=1}^{N}\frac{(C_i^{market} - C_i^{model})^2}{N}},$$

   where N is the number of data points. These two error metrics and the computation time will be reported for all the data-sets.

2. Since the Heston and the SVJ models are treated as 'benchmark' models, one expects that any new, *sensible* model calibrated from the same data-set as one of these models will yield similar prices for illiquid or non-traded payoffs. We test whether this is the case for the new model by pricing down-and-out-call barrier options for a range of strikes, barriers and expiration, using the three models calibrated from the same data-set. The experiments were repeated with floating strike, arithmetic average Asian calls. Note that in both these cases, there are no 'true' or unique prices and we are simply expecting the models calibrated from the same data to yield similar prices for illiquid securities.

3. Finally, one also expects the models calibrated from the same data to yield similar option price sensitivity parameters, which are crucial in risk monitoring and hedging purposes. This fact is also tested by numerically calculating $\Delta = \frac{\partial C}{\partial S}$ and $\Gamma = \frac{\partial \Delta}{\partial S}$ for options for each of the models, over all the data sets.

One may ask if it is possible to calculate an option Vega, using the MSV model. However, Vega, *i.e.* the rate of change of option price with respect to change in volatility, is undefined for stochastic volatility models, since the volatility is a process rather than a constant parameter. Besides the obvious - and common - way of ignoring Vega hedging altogether, one way the market deals with the problem of sensitivity with respect to volatility is to use the initial value v_0 as a proxy for volatility, and use the first partial derivative of option price relative to this parameter as a proxy for Vega. Unfortunately, this presents very significant model risk and a small perturbation in parameters often leads to very different Vega values. This is a known fact in the market and was confirmed in the numerical experiments as well. As Vega values are extremely sensitive to model error and, in addition, vary wildly from model to model for the same option parameters (strike and expiration). Hence, the comparison would have been uninformative in the sense of a contribution to this research.

The next subsection and the accompanying tables and figures in section 4.5 provide representative results to support the arguments discussed in this chapter.

4.3.3 Results

The application of the new model to the real market data is now discussed. As mentioned above, three different sets of results were considered: the accuracy in matching the traded option prices, a comparison of illiquid option prices via simulation and a comparison of the sensitivity parameters via numerical approximation.

- The in-sample and the out-of-sample errors (as measured by MRAE and RMSE in both the cases) of all the data-sets are presented in the section 4.5. The in-sample errors are denoted by MRAE-I, RMSE-I and the out-of-sample errors are denoted by MRAE-O, RMSE-O. Sample parameters presented for MSV, SV and SVJ models for one data set are presented in

tables 4.9 and 4.8. Tables 4.1-4.5 provide the achieved errors for data on five different days, with each table reporting in-sample as well as out-of-sample error metrics for the three indices for that day. Boldface numbers in each column indicate the worst value for the error metric obtained for that data subset (in-sample or out-of-sample subset, for each data-set). With three indices, five days, two data subsets for each index on each day and two error metrics, we have a total of 60 error columns to compare the three models (Heston, Bates and MSV) with. From tables 4.1-4.5, the MSV model has the worst performance (out of the three models) only 9 out of 60 times, with one of the two benchmark models being the worst performer in all of the remaining 51 cases. This supports the modest claim of this chapter that the new model is very competitive in terms of accuracy with the benchmark models. The other important set of numbers is the calibration times. Tables 4.1-4.5 show that the MSV model can be calibrated within 1.25 seconds in all the fifteen cases, while the lowest calibration time for the other two models is 41.32 seconds. In summary, tables 4.1-4.5 indicate that one can obtain a very good fit to option prices with the MSV model at a fraction of the calibration cost, as compared to some of the existing popular models.

- Next, the three models were compared for prices of illiquid options, when calibrated from the same data set. Table 4.6 outlines the prices obtained for down and out barrier call options, priced using each of three models calibrated from the 1st November 2012 FTSE options data-set. It may be recalled that down-and-out call barrier option with strike K and barrier B has a payoff $\max(S_T - K, 0)$ at expiration time T unless $S_t < B$ at any point between $t = 0$ and $t = T$, in which case the option ceases to exist.

  The option prices were simulated using Euler discretisation for all the models with 10000 steps for each sample path and with 10000 sample paths. The obtained prices and confidence intervals (denoted as CI) for various values of expiration times T, interest rates r, barriers and strike prices are reported in table 4.6. As can be seen, the prices given by the new model are within 10% (in the worst case) of *either* Heston price *or* SVJ price. As there is no unique option price in this case, the aim of this chapter is simply to establish that the MSV model gives *believable*

prices, which are not too far from those given by benchmark models. Moreover, the prices by Heston and SVJ models can themselves differ by 10% or more. It should also be noted that simulation using the new model is computationally somewhat cheaper than that with either of the other two models.

Floating strike, arithmetic average Asian call options were also priced with the three models, calibrated from the 1st November 2012 data-sets (for all the three indices). This generally illiquid option has a payoff $\max(0, S_T - S_{av})$ at expiration, where S_{av} represents the time average of the underlying price between $t = 0$ and $t = T$, T being the expiration. In this case as well, we simulated the option prices using Euler discretisation, 10000 steps for each sample path and 10000 sample paths. The results are reported in table 4.7, along with 95% confidence intervals. As can be seen, the prices obtained by the new model are close to those obtained by SVJ model.

Similar experiments were performed with other data-sets with the same qualitative conclusions; hence the results are omitted for brevity.

- As a final measure of performance, the three models were compared in terms of the sensitivity parameters delta and gamma for the options. These parameters were compared over all the fifteen data-sets. For all the models, approximate values of these parameters are obtained using a central difference approximation scheme as follows:

$$\Delta \approx \frac{C(S+\delta) - C(S-\delta)}{2\delta} \text{ and}$$
$$\Gamma \approx \frac{C(S+\delta) - 2C(S) + 2C(S-\delta)}{\delta^2},$$

where $C(x)$ indicates option price evaluated at the price of underlying equal to x, S is the price of the underlying and δ is a small increment. While more sophisticated methods to calculate these parameters exist (and it is trivial to find these analytically for the MSV model by differentiation), the purpose of this experiment is to compare whether the values given by the new method are in the same range as the values given by the other two methods. A selection of results is presented in figures 4.1-4.3. The remaining results are qualitatively similar, and are omitted

for brevity. Note that the apparent periodicity is simply a result of the same set of strikes being repeated for different expirations. For FTSE and S & P data-sets, the sensitivity parameter estimates from MSV tends to be close to one of the other two models, except at short maturities. The deviation of MSV delta and gamma from those given by the other two models is the highest for 25 July 2013 DAX data set. This is also the only data-set when the RMSE and MRAE errors for MSV model are the worst among the three models; please see table 4.3. Gamma values of all the three models at short maturities vary quite significantly and it is not immediately obvious which values should serve as benchmark values.

It is worth mentioning that no evidence was found whether the MSV model works consistently better/worse at short or long maturities, or for in-the-money or out-of-the-money options.

4.4 Summary

The contribution of this chapter is threefold. First and the main contribution is that the new random volatility model was proposed, called high order moments-based stochastic volatility model (or MSV model), in which the volatility is a function of time with its level being modulated by a random variable. By using a Taylor series expansion of the option price, it was shown that the model yields an easy formula for approximate option prices and hence can be calibrated extremely fast. The proposed model can even be implemented on a spreadsheet[2].

Secondly, it was demonstrated through comprehensive numerical experiments that MSV model is very competitive in terms of accuracy with the Heston model and the SVJ model, while being computationally significantly cheaper to calibrate. Lastly, the claims for the usefulness of the new model have been backed up with simulation experiments for comparison of exotic option prices as well as comparison of numerically evaluated option price sensitivity parameters. The MSV model thus provides a competitive alternative to the existing option pricing models; it is particularly suitable for high frequency financial trading due to its speed of calibration.

[2]An Excel spreadsheet implementation can be found on the CD.

Note that it is conceptually straightforward to use a semi-parametric model, by using a piecewise linear γ_t in (4.5) which is non-zero for $t > t_0$, to match the observed option prices even more accurately. The use of such semi-parametric models with piecewise constant volatility parameters is quite common in financial modelling, *e.g.* it is used in calibrating a LIBOR forward model to observed caplet prices (see [15] and references therein, for example).

4.5 Tables and graphs

Table 4.1: 01 November 2012

	MRAE-I	RMSE-I	Time (sec.)	MRAE-O	RMSE-O
			FTSE		
Heston	**4.79**	10.43	148	**4.97**	**14.85**
SVJ	3.33	**11.29**	407.70	3.53	3.38
MSV	3.30	9.62	1.24	2.08	5.77
			S&P500		
Heston	**8.20**	**5.99**	605.72	**7.13**	**6.9**
SVJ	1.23	1.38	1419	1.28	0.54
MSV	4.75	3.30	0.44	4.73	3.27
			DAX		
Heston	2.45	9.74	77.35	3.72	9.51
SVJ	**4.75**	**35.38**	771.17	4.38	2.80
MSV	4.37	20.51	0.24	**5.60**	**23.42**

Table 4.2: 26 November 2012

	MRAE-I	RMSE-I	Time (sec)	MRAE-O	RMSE-O
			FTSE 100		
Heston	**6.36**	12.04	109	**6.21**	**10.45**
SVJ	3.09	**12.97**	378.7	2.98	1.73
MSV	3.33	7.82	0.76	4.32	8.63
			S&P 500		
Heston	**4.68**	**4.42**	193	**4.83**	**5.46**
SVJ	3.31	3.13	1115.62	3.2	0.65
MSV	3.81	2.32	0.36	3.84	3.15
			DAX		
Heston	6.32	55.19	95.87	5.78	**46.05**
SVJ	**7.25**	**61.94**	910.39	**6.78**	45.85
MSV	4.79	42.80	0.28	4.42	40.38

Table 4.3: 25 July 2013

	MRAE-I	RMSE-I	Time	MRAE-O	RMSE-O
			FTSE 100		
Heston	**7.34**	**22.26**	1332.88	**6.20**	**14.31**
SVJ	4.50	10.87	671.12	5.15	5.23
MSV	4.15	9.80	0.81	5.27	11.11
			S&P 500		
Heston	**4.78**	**3.26**	252.93	**6.28**	**3.45**
SVJ	3.59	2.63	1156.91	3.80	1.79
MSV	3.97	3.11	1.01	4.30	2.34
			DAX		
Heston	10.32	88.39	145.26	12.56	102.16
SVJ	8.20	89.12	600.71	11.89	34.93
MSV	**13.81**	**110.51**	0.71	**17.18**	**114.35**

Table 4.4: 26 July 2013

	MRAE-I	RMSE-I	Time	MRAE-O	RMSE-O
			FTSE 100		
Heston	**8.32**	**23.82**	1146.46	**7.07**	**14.75**
SVJ	4.50	10.87	671.12	5.15	5.23
MSV	4.45	9.37	0.62	5.83	10.80
			S&P 500		
Heston	**5.00**	**3.34**	257.44	**6.30**	**3.43**
SVJ	3.01	2.65	1337.72	3.06	1.79
MSV	3.86	3.06	1.09	1.00	2.43
			DAX		
Heston	**4.33**	**34.34**	286.69	4.19	**22.75**
SVJ	2.70	26.32	418.64	2.99	4.59
MSV	4.24	17.55	0.65	**8.01**	18.17

Table 4.5: 29 July 2013

	MRAE-I	RMSE-I	Time	MRAE-O	RMSE-O
			FTSE 100		
Heston	**8.66**	**23.58**	667.25	**7.70**	**14.53**
SVJ	4.78	12.38	590.72	4.40	4.65
MSV	4.52	9.35	0.85	5.74	10.73
			S&P 500		
Heston	**5.78**	3.50	41.32	**6.01**	3.20
SVJ	2.57	**19.43**	291.20	3.81	**5.87**
MSV	4.71	2.99	1.00	5.29	2.30
			DAX		
Heston	4.26	**28.49**	134.29	4.47	**19.44**
SVJ	2.70	26.32	418.64	2.99	4.59
MSV	**4.34**	18.35	0.64	**8.03**	18.75

Table 4.6: Down-and-out Call Barrier option prices (models calibrated from 1st November 2012 FTSE data)

DAX S	7281.18			Barrier	7100.00			Strike	7250.00	
		SVJ			Heston			MSV		
T	r	Price	CI		Price	CI		Price	CI	
0.10	0.0051	92.93	90.18	95.67	83.13	80.03	86.24	94.66	91.59	97.72
0.21	0.0049	132.92	122.56	143.27	117.24	112.83	121.65	135.96	130.93	140.99
0.32	0.0052	137.97	132.14	143.80	131.89	126.48	137.30	143.25	137.10	149.41
0.43	0.0056	144.88	138.38	151.39	141.61	135.36	147.86	153.09	145.96	160.23
				Barrier	7200.00			Strike	7300.00	
		SVJ			Heston			MSV		
T	r	Price	CI		Price	CI		Price	CI	
0.10	0.0051	55.16	52.92	57.41	53.24	50.75	55.73	60.75	58.00	63.50
0.21	0.0049	63.16	60.17	66.16	60.67	57.36	63.98	68.46	64.58	72.35
0.32	0.0052	69.14	65.52	72.76	67.40	63.43	71.37	73.65	68.99	78.32
0.43	0.0056	77.26	60.28	94.24	73.40	68.67	78.14	78.96	73.48	84.45
FTSE S	5812.06			Barrier	5750.00			Strike	5820.00	
		SVJ			Heston			MSV		
T	r	Price	CI		Price	CI		Price	CI	
0.06	0.0051	61.63	55.40	67.86	40.07	38.56	41.58	54.88	51.50	58.25
0.20	0.0049	61.26	53.02	69.49	53.11	50.57	55.66	58.66	53.77	63.54
0.31	0.0052	65.69	55.56	75.82	55.68	52.72	58.65	64.78	58.73	70.83
0.42	0.0056	59.17	49.33	69.01	57.67	54.35	60.98	66.69	60.03	73.35

Table 4.7: Arithmetic average Asian option with floating strike (1st November 2012, all indices)

				SVJ			Heston			MSV		
Index	S0	T	r	Price	CI		Price	CI		Price	CI	
FTSE	5812.06	0.42	0.0056	89.31	86.02	92.59	84.94	82.49	87.38	88.41	85.82	91.00
SNP	1412.16	0.31	0.003	21.56	21.04	22.08	14.88	13.84	15.91	20.44	19.84	21.04
DAX	7281.18	0.10	0.001	60.88	59.18	62.58	56.04	53.93	58.15	61.73	59.91	63.55

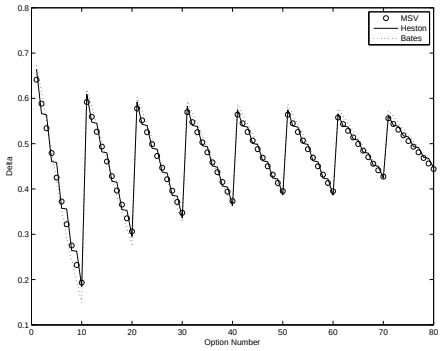
Table 4.8: Heston and SVJ Parameters on 1 November 2012

	κ	θ	ρ	v_0	σ_v	μ_j	σ_j	λ_p
Heston	4.75	0.0085	-0.99	0.0096	0.0118			
SVJ	100	0.0065	-0.99	0.0151	0.99	0.191	0.00001	6.6132

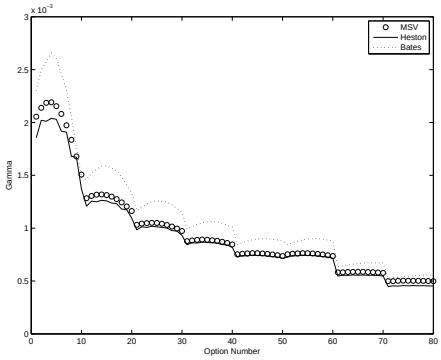
Table 4.9: MSV parameters on 1 November 2012

	k^2	σ_0	σ_1	σ_2	λ
MSV	0.0004	0.1082	0.0027	0.2937	1.3414

Figure 4.1: FTSE100: Delta and Gamma on 25.07.2013

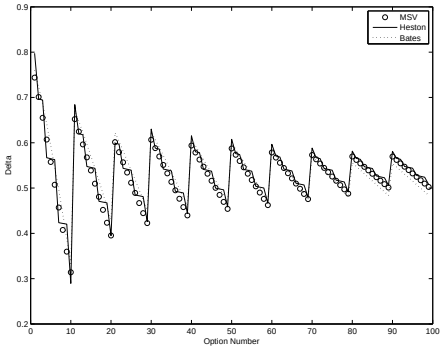


(a) Delta for 25.07.2013

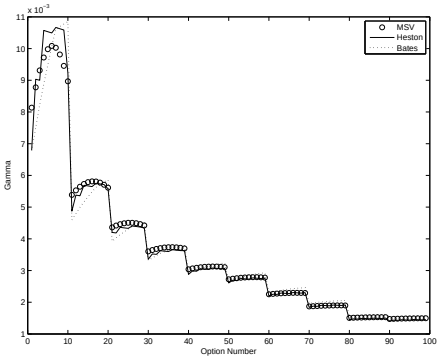


(b) Gamma for 25.07.2013

Figure 4.2: S&P 500:Delta and Gamma on 25.07.2013

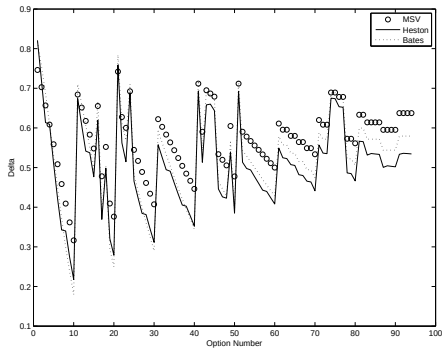


(a) Delta for 25.07.2013

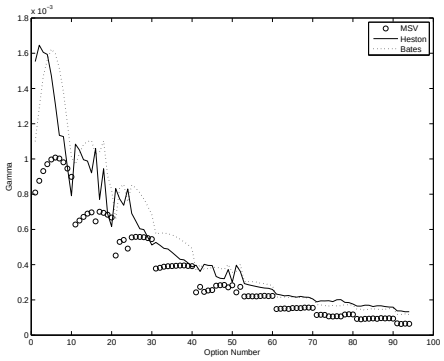


(b) Gamma for 25.07.2013

Figure 4.3: DAX:Delta and Gamma on 25.07.2013



(a) Delta for 25.07.2013



(b) Gamma for 25.07.2013

Chapter 5

Electricity futures price models: calibration and forecasting

5.1 Introduction

Due to its specific structure, the modelling of security prices in the electricity market remains a big challenge for economists and risk managers. As a non-storable commodity, electricity is traded one day ahead and the contract sizes are measured in MWh. Spike structure is usually observed in the electricity spot price time series. Various modern approaches to model this behaviour have been suggested. [26], [32], [78] and [18] use a compound Poisson process coupled with Ornstein-Ulenbeck process for the spot price. Weron et al [79] introduced two different jump processes to account for different jump sizes in positive and negative directions. A jump regime switching model was developed in [48]. Geman and Roncoroni in [37] and [38] use jump direction threshold to force negative jumps if the price exceeds the threshold value. A numerical algorithm based on a continuous time Markov chain was studied by [4]. Different approaches for modelling the electricity prices were surveyed recently in [64]. A drawback of the class of jump-diffusion models used to model electricity prices discussed in most of these papers is a significant complexity of estimating a large number of model parameters from data.

When it comes to a jump risk premium, two ideas exist in the literature. An earlier approach formulated by Merton [60] and followed by [53], [9] and [11] ignores the jump risk premium in general, *i.e.*, treats the jump risk as purely idiosyncratic and assumes that it can be diversified away. The risk premium is

applied only to jump size components of a compound Poisson process. More recently, [66] provides evidence on an implicit jump risk premium. Finally, [16] formulates a general framework for non-Gaussian processes, introduces a risk aversion in non-Gaussian processes and gives an explicit expression for the risk premium in jump processes.

As futures contracts in energy markets are far more liquid than the spot security, it is common to estimate the implied spot price from futures prices. Estimation of the spot commodity price using the Kalman filter was performed by [71], [56] and [54]. However, the classical Kalman filter can be applied only to a linear state space model. [2] used a particle filter to estimate the dynamics of state variables in a two-factor model with jumps and used the maximum likelihood method to obtain the model parameters. [1] used a particle filter to estimate parameters for a two-factor model with jumps.

In this chapter, a new random volatility model is introduced for electricity price modelling. Volatility in this model is taken to be a random variable to explain the non-Gaussian log-spot price behaviour. An approximation to the futures price for this model is derived using the moments of the random volatility, and it is shown that the model performs at least as well as multi-factor models, which are much harder to calibrate and to simulate than the new model.

Further, two methodological contributions are made to the existing literature on two-factor models of commodity prices. Firstly, a simple extension to the two-factor model with jumps is proposed, which was discussed in [78]. This includes adding a sinusoid (rather than a Fourier series based periodic function) to model the seasonality and empirically investigating whether an explicit evaluation of jump risk premium makes a difference to accuracy of pricing. Secondly, we propose a new multi-step calibration procedure to estimate the model parameters, which is computationally simpler than the commonly employed simulation-based estimation in similar problems.

The modified two-factor model (with a seasonality factor and jump risk premium) is calibrated using this simplified calibration procedure and compared with the new model on real electricity futures price data, in terms of accuracy in one step ahead prediction.

The rest of this chapter is structured as follows. Section 5.2 presents the new random volatility model, with a derivation of approximate futures price under this model, while section 2.4 outlines the existing two-factor jump-

diffusion model. Section 5.3 contains the methodology and discussion about the data used for numerical experiments. Section 5.4 includes the results on the comparison of different models. Finally, Section 5.5 provides the summary of this research.

5.2 A new commodity price model

Empirical studies of log spot price of electricity show a significant deviation from normality, principally due to price spikes over time. A simple new model is proposed here which takes into account the resulting fat tailed nature of the log spot distribution. As mentioned before, the traditional approach for modelling price spikes is using a compound Poisson process in addition to a mean-reversion process. However, the resulting two-factor jump-diffusion model has a fairly large number of parameters, which makes it difficult and fairly time-consuming to calibrate. This problem is avoided in the new *random volatility* model, which is discussed below.

A commonly used generic form of the model of energy commodity price behaviour is:

$$S_t = f(t)e^{x_t}, \tag{5.1}$$

where S_t is a commodity price at time t, $f(t)$ is function of time which represents a seasonality pattern and x_t is a stochastic process. Our aim is to construct a stochastic process which can account for all of the properties of electricity price distribution, but preserves a level of complexity which is comparable with a one-factor model.

In the filtered probability space $(\Omega,\ \mathcal{Q},\ \mathcal{F}_t)$ with $\mathcal{Q}$ being the risk-neutral measure and $\mathcal{F}_t$ being the natural filtration, the evolution of de-seasonalised log commodity price is modelled by:

$$dx_t = (\alpha - \kappa x_t)dt + \sigma_t dW^{\mathcal{Q}}_{1,t}, \tag{5.2}$$

$$d\sigma_t = \hat{f}_1(t, \sigma_t)dt + \hat{f}_2(t, \sigma_t)dW^{\mathcal{Q}}_{2,t}, \tag{5.3}$$

where α and $\kappa > 0$ are scalar constants, σ_t is a positive process on the real line, $W_{i,t} \quad (i = 1, 2, t > 0)$ are Wiener processes such that $< W_{1,t}, W_{2,t} >= 0$ and $\hat{f}_1,\ \hat{f}_2$ are real valued smooth functions.

Recall that the arbitrage-free futures price $F(t, T)$ at time t with maturity

time T is given by:

$$F(t,T) = f(T)\mathbb{E}^{\mathcal{Q}}(e^{x_T}|\mathcal{F}_t) \tag{5.4}$$

where $T > t$ is the maturity time. Assume that x_t is a right continuous process generated by $\mathcal{Q}$-Wiener process and all the central moments $m_i\,(i = 1..\infty)$ of x_t exist $\forall t \in [0,T]$.

The expected value of the process e^x can be written in terms of an infinite series as follows, with $x = x_T$ for brevity:

$$\begin{aligned}\mathbb{E}^{\mathcal{Q}}(e^x|\mathcal{F}_t) &= e^{m_1}\mathbb{E}^{\mathcal{Q}}(e^{x-m_1}|\mathcal{F}_t) = e^{m_1}\sum_{i=0}^{\infty}\frac{\mathbb{E}^{\mathcal{Q}}(x-m_1|\mathcal{F}_t)^i}{i!} \\ &= e^{m_1}\left(\left(\frac{m_2}{2}+\frac{3m_2^2}{24}+\ldots\right)+\frac{m_3}{8}+\frac{m_4-3m_2^2}{24}+\frac{m_5}{120}+\ldots\right).\end{aligned} \tag{5.5}$$

Now, one can collect together terms which lead to a Taylor expansion of $e^{\frac{1}{2}m_2}$, noting that odd moments of the Gaussian distribution are zero:

$$\mathbb{E}^{\mathcal{Q}}(e^x|\mathcal{F}_t) \approx e^{m_1}\left(e^{\frac{1}{2}m_2}+\frac{m_4-3m_2^2}{24}+\frac{m_6-15m_2^3}{720}\right). \tag{5.6}$$

In the subsequent discussion, σ_t is assumed to be a log-normally distributed random variable for all $t > 0$, with a constant mean μ and a constant variance η^2. This stationary probability distribution can be obtained at each time t by assuming the following simple stochastic process for $\log\sigma_t$:

$$\log\sigma_t = \mu + \frac{\eta W_{2,t}^{\mathcal{Q}}}{\sqrt{t}}, t > 0, \log(\sigma_0) = \mu,$$

with $< W_{1,t}^{\mathcal{Q}}, W_{2,t}^{\mathcal{Q}} >= 0$. This provides a theoretical justification for the choice of log-normal distribution for σ_t, although other distributions may be chosen in practice. Since we are interested in pricing futures contracts which are path-independent securities, σ_t can be treated as a time-independent random variable and henceforth the time index is omitted from the notation for σ_t.

Under the real world (or physical) measure $\mathcal{P}$, let the log spot price process

be given by

$$\begin{aligned} dx_t &= (\bar{\alpha} - \kappa x_t)dt + \sigma dW^{\mathcal{P}}_{1,t}, \\ \sigma &\sim LN(\mu, \eta^2), \end{aligned} \tag{5.7}$$

where $W^{\mathcal{P}}_{1,t}$ is a Wiener process under a physical measure.

Then, assuming the absence of arbitrage, there exists a price of risk process h_x such that $\alpha - \bar{\alpha} = h_x$, which is assumed to be constant. For de-seasonalised data, the mean reversion level $\bar{\alpha}$ in the real world measure is set to 0, which is in keeping with the convention (see [56], for example). The random variable σ under risk-neutral measure has a log normal distribution as described above.

The fact (5.6) can be used to build an approximation of futures price formula (5.1) which is accurate enough for most purposes. Taking the logarithm it can be easily shown that the log futures price is approximately given by:

$$\begin{aligned} \log F(t,T) \approx \log(f(T)) + x_t e^{-\kappa(T-t)} + \frac{h_x}{\kappa}(1 - e^{-\kappa(T-t)}) + \frac{1}{2}m_2 + \\ \log\left(1 + e^{-\frac{m_2}{2}}\left(\frac{m_4 - 3m_2^2}{24} + \frac{m_6 - 15m_2^3}{720}\right)\right), \end{aligned} \tag{5.8}$$

where m_i is the i^{th} central moment of x_T conditional on x_t and the Taylor series approximation of $F(t,T)$ is truncated after 7 terms (note that all the odd central moments of x_T are 0 for the model defined above). Equation (5.8) defines the log futures price in terms of the new random volatility model. The equations for the moments $m_i = \mathbb{E}^{\mathcal{Q}}(x_T - \mathbb{E}^{\mathcal{Q}}(x_T|x_t))^i$ in the random volatility model are as follows:

$$\begin{aligned} m_1 &= x_t e^{-\kappa\Delta} + \frac{h_x}{\kappa}(1 - e^{-\kappa\Delta}), \\ m_2 &= \frac{e^{2\mu+2\eta^2}}{2\kappa}(1 - e^{-2\kappa\Delta}), \\ m_4 - 3m_2^2 &= \frac{3(e^{4\eta^2} - 1)e^{4(\mu+\eta^2)}}{4\kappa^2}(1 - e^{-2\kappa\Delta})^2, \\ m_6 - 15m_2^3 &= \frac{15(e^{12\eta^2} - 1)e^{6(\mu+\eta^2)}}{8\kappa^3}(1 - e^{-2\kappa\Delta})^3, \end{aligned}$$

where $\Delta = T - t$.

We use the fact that the following identity holds for a Gaussian random

variable x:

$$\log \mathbb{E}(e^x) = \mathbb{E}(x) + \frac{1}{2}\mathrm{Var}(x)$$

holds for a Gaussian random variable x. Note that, if $\eta = 0$ (i.e., if the volatility σ is a deterministic constant), it can be easily shown that equation (5.8) reduces to the corresponding formula for the linear Gaussian model. During calibration, the futures price data is assumed to be observed in noise. This measurement noise can be looked upon as a proxy for the approximation error introduced due to truncation of the Taylor series. Specifically, the measurement equation at each time step t_k is written as

$$\mathrm{vec}\{z(t_k, T_i)\} = \mathrm{vec}\{\log F(t_k, T_i)\} + v_{t_k}, \tag{5.9}$$

where $\log F(t_k, T_i)$ is the log futures price for maturity T_i at time t_k, as given by (5.8), and $v_{t_k} \sim N(0, \Sigma)$ is a measurement noise vector with zero mean and a covariance matrix Σ. vec operator is as defined in chapter 3:

$$\mathrm{vec}(x_i) = \begin{bmatrix} x_1 & x_2 & \cdots & x_N \end{bmatrix}^\top.$$

Having defined the new model, the models are outlined against which the new model will be benchmarked, *viz* the two-factor jump-diffusion model as well as its jump-free special case. The actual numerical experiments performed with the new model and two-factor jump-diffusion models (see Section 2.4) are described next.

5.3 Numerical Experiments

5.3.1 Methodology

Empirical study of the models has the following steps:

- De-seasonalisation: a parametrised seasonality function (2.37) is used to de-seasonalise the data.
- A new multi-step heuristic for parameter estimation in two-factor models: two-factor models (with and without jumps) described in section 2.4 are highly nonlinear and contain a large number of parameters. To alleviate the difficulty of parameter estimation, a new systematic multi-step calibration algorithm is introduced for two-factor models.

– Starting from the characteristic function of the process, the first 9 moments of the log spot price are evaluated analytically at each time t and then the method of moments is used to estimate the parameters of the model in historical measure, from time series data.

– With the acquired parameters, the covariance matrix is estimated for the moments.

– Using the inverse of the covariance matrix as the weight, the parameters are re-estimated.

– Given the parameters of the spot price process in historical measure obtained *offline* as above, the risk premium parameters (assumed to be constant) is estimated along with the observation noise covariance matrix Σ using the least squares method *online*, every ten steps (this number of steps is arbitrary and depends on factors such as the hardware specification). One step ahead prediction of the spot price is achieved through a particle filter. Note that Σ is needed for updating the probability weights in the particle filter at each time-step; please refer to equation (2.74).

More details on this procedure are provided in section 5.3.3.

- Testing for in-sample and out-of-sample prediction ability for all the models (two-factor models with and without jumps as well as the new random volatility model): a particle filter is used to track the latent state process and get one step ahead forecasts for the futures prices. The transition density is set equal to the proposal density in these experiments. Particles for the particle filter were sampled using the transition equation for the relevant process in the following experiments.

5.3.2 Characteristic function and the moments for two-factor models

To derive the characteristic function for the jump-diffusion two-factor model defined in (2.34)-(2.36), a de-seasonalised log price process $S_t^* = \log S_t - f(t)$ is considered. Assume that the jump process is uncorrelated with the Brownian

motion. First, a characteristic function is taken in the following form:

$$f(x,\zeta,t) = \mathbb{E}\{e^{iu(x_T+\zeta_T)}|X_t = x, \zeta_t = \zeta\}.$$

Applying Ito's formula to the $M_t = f(x,\zeta,t)$, and assuming that dM_t is a martingale, a characteristic function of the following form (see, e.g. [76]) is given as follows:

$$\phi_{S_t^*}(u) = \phi_{x_T+\zeta_T}(u)\phi_J(u), \tag{5.10}$$

$$\phi_{x_T^C+\zeta_T}(u) = \exp\{(-u^2(\frac{\rho\sigma_1\sigma_2}{\kappa}(1-e^{-\kappa T}) + \frac{\sigma_1^2}{4\kappa}(1-e^{-2\kappa T}) + \frac{1}{2}\sigma_2^2) + iu(x_0e^{-\kappa T} + \zeta_0 + \mu T)\}, \tag{5.11}$$

$$\phi_J(u) = \exp\{\lambda_J t(e^{iu\mu_J e^{-\kappa t} - \frac{1}{2}\sigma_J^2 e^{-2\kappa t}u^2} - 1)\}. \tag{5.12}$$

From the definition of the characteristic function, one can evaluate moments of the desired process using the following formula:

$$m_n = \frac{1}{i^n}\frac{\partial^n}{\partial u^n}\phi_{S_t^*}(u). \tag{5.13}$$

These moments can be calculated analytically using any symbolic computation software such as Mathematica, Mathcad, Matlab, etc. The exact (and lengthy) expressions for moments can be found in Appendix A. The method of moments is used to estimate the parameters of the model (2.34) from sample moments based on data. However, this will not allow us to find the parameters for the risk premium. As mentioned in the previous subsection, a simple multi-step heuristic is used, where most of the parameters are estimated offline in the historical measure using the method of moments and then the risk premia (which are assumed to be constant) are estimated online by least squares. This is explained in more details in the next section.

5.3.3 Parameter estimation for two-factor models

For parameter estimation, the analytically derived first n moments were used from (5.13), where n is equal to the number of unknown parameters after de-seasonalisation. Let Θ be a vector of unknown parameters. Let $m_n(\Theta)$ be the parameterised n^{th} central moment of the commodity price data set and let m_n^o be the sample n^{th} central moment computed from the observed data

(analytical expressions for the moments can be found in Appendix A). Then the following cost function is minimised:

$$\min_{\Theta} (m_n^o - m_n(\Theta))^T \Lambda^{-1} (m_n^o - m_n(\Theta)), \tag{5.14}$$

where Λ is a weighting matrix. The optimisation is done in three steps:

- Firstly, Λ is assumed to be the identity matrix and the cost function is minimised over the parameter vector Θ. Let $\Theta = \Theta^\star$ be the parameter vector which achieves this optimum.
- Next, the diagonal entries of Λ are set as free variables, with a condition that $|\Lambda| \geq 0$ and minimize the cost function with fixed $\Theta = \Theta^\star$. Let $\Lambda^\star$ be the matrix which achieves this optimum.
- Finally, (5.14) is solved over Θ again, with the weighting matrix set to $\Lambda^\star$.

The vector of optimal parameters is obtained offline using the above procedure, the particle filter can be set up. At each time (or each day) t_i, The one day ahead prediction of the price of each futures contract is based on the arithmetical average of the the predicted prices of the corresponding contract over all the particles generated. The risk aversion parameters *viz.* $\{h_x, h_\zeta, \beta\}$ as well as the noise covariance matrix Σ can be updated online, along with the price predictions made by the particle filter. See section 5.3.5 for more details on this filtering and online calibration stage.

5.3.4 Parameter estimation for the random volatility model

The random volatility model described in section 2 has far fewer parameters in comparison to two-factor models and a simpler (approximate) measurement equation. This allows us to use the maximum likelihood method directly. The parameter vector $\hat{\Theta} = \{\kappa, \mu, \eta, h_x\}$ is estimated along with a diagonal covariance matrix Σ of the measurement noise. From the assumption that log futures price observations available under additive Gaussian noise, a likelihood

function of the observation is constructed, which has the following form:

$$L(\Theta) = \prod_{i=0}^{N} p(\Theta|\zeta)q(\zeta), \tag{5.15}$$

where $p(\cdot)$ and $q(\cdot)$ are normal and lognormal density functions respectively and N is the total number of observations. Taking a logarithm of both sides and substituting p and q with their exact forms gives:

$$\begin{aligned}\log L(\Theta) = &- \frac{N}{2}\left(\log\eta + \frac{\mu^2}{\eta^2}\right) - N\log|\Sigma| - \\ &\frac{1}{2}\sum_{i=0}^{N}\left(y_i - \mathrm{vec}(\log F(t_i, T_k))\right)^{\top}\Sigma^{-1}\left(y_i - \mathrm{vec}(\log F(t_i, T_k))\right),\end{aligned} \tag{5.16}$$

where y_i are observed futures price vectors at time t_i, and vec operator is with respect to the futures maturities T_k, $k = 1, 2, \ldots, n$. Once the parameters are obtained, a particle filter can be used for prediction of futures prices, as it is done for the two-factor models.

The next section describes an online calibration procedure for risk premium parameters and the measurement noise covariance matrix. This procedure is common for the new random volatility model as well as for the two-factor models with and without jumps.

5.3.5 Online calibration stage

Two different quantities are estimated online: the covariance matrix of measurement noise and the parameters reflecting the risk premia.

To estimate the covariance matrix of measurement noise, Σ^0 is initialised as an identity matrix and use the following update for each time step k:

$$\Sigma^k_{i,i} = (\hat{v}^k_i)^2, k = 1, \ldots, N,$$

where

$$\hat{v}^k_i = (\log F^{(market)}(T_k, t_i) - \log \hat{F}^{(theoretical)}(T_k, t_i)), \quad k = 1, \ldots, n \tag{5.17}$$

and $\hat{F}^{(theoretical)}(T_k, t_i))$ represents the average theoretical price using the the

particle filter iteration.

To estimate the risk premia, a vector r_i is defined, whose entries are the risk premia for the relevant model (e.g. $r_i = \begin{bmatrix} h_x & h_\zeta & \beta \end{bmatrix}^\top$ for the two-factor model with jumps).

At the first step, let $r_0 = 0$ (here 0 is a zero vector with $dim(r)$ rows). After obtaining Σ_i at step i, a nonlinear least squares problem can be solved:

$$\min_{r_i} (\log F^{(market)}(T_k, t_i) - \log \hat{F}^{(theoretical)}(T_k, t_i))^2.$$

As mentioned in section 5.1, this problem is solved after each ten time-steps to update the risk premia.

This section and the preceding three sections summarised the calibration procedures for the models used in the numerical experiments. The next section discusses the data used for the numerical experiments.

5.3.6 Data

For the empirical study, the European electricity market was considered as the data source for the experiments, since Nord Pool is the largest market for electrical energy in the world, covering most of the Europe and is traded at NASDAQ OMX Commodities Europe. Over 1500 trading days data were used of spot price and futures contracts (22d, 44d, 66d, 88d, 110d, 132d) : starting from 19/11/2007 to 17/12/2013.

The data was split into 3 data sets of 300 trading days each, with 200 days delay between each data set. Each data set has two equal parts:

- In-sample: 150 observations of futures and commodity prices. This data set is used to estimate the *offline* model parameters.
- Out-of-sample: 150 observations of futures prices to test the behaviour of the models out-of-sample and to estimate (and update) the online model parameters (risk premia and the measurement noise covariance).

The statistics about the data is shown in table 5.1. One can observe that the log-spot price data has different properties when compared to the futures price distribution:

- The standard deviation of the spot price is higher than the average standard deviation of the futures prices by 40%.

Table 5.1: Log futures price statistics

T (days)	Mean	Variance(daily)	Skewness	ExKurtosis
(spot)	3.6962	0.3384	-1.1072	4.5390
22	3.7263	0.2632	-0.0327	0.3396
44	3.7447	0.2393	0.1654	-0.2761
66	3.7562	0.2240	0.3144	-0.4218
88	3.7633	0.2186	0.4328	-0.1573
110	3.7681	0.2087	0.5545	0.0090
132	3.7698	0.2046	0.5813	0.0311

- The log spot price is more skewed and has a higher excess Kurtosis than log futures prices. This high kurtosis is a result of price spikes, which are commonly modelled using a jump-diffusion process as mentioned earlier.

5.3.7 Choice of measures for comparison

For comparison of the performance of models in terms of forecasting, the sample mean of the relative absolute error (MRAE) and root mean square error (RMSE) are considered as the measures of prediction error for the futures price data.

$$MRAE_T = \frac{1}{N}\sum_{i=1}^{N}\frac{|F_{i,T} - \hat{F}_{i,T}|}{F_{i,T}},$$

$$RMSE_T = \sqrt{\sum_{i=1}^{N}\frac{(F_{i,T} - \hat{F}_{i,T})^2}{N}},$$

where $\hat{F}_{i,T}$ are the average values of one step ahead predicted futures prices evaluated for each particle drawn on the i^{th} time step of a particle filter with a corresponding maturity T, and $F_{i,T}$ are the observed futures prices at maturity T. These measures are evaluated for each of the six futures contracts, for each of the three data sets and for both in-sample and out-of-sample data.

5.4 Results and discussion

The two-factor model (TF), the two-factor model with jumps (TFJ) and the new random volatility model (RVM) are compared. The results for in-sample and out-of-sample performance are presented for each model using MRAE and RMSE measurements.

Since two different error metrics were used over six data-sets (three in-sample and three out-of-sample), and since each data set has six futures contracts, a total of 72 columns of errors were obtained to compare the three models with. The bold font indicates the worst (or the highest) error metric in each column (i.e. for each data-set + futures contract + error metric).

Tables 5.2 and 5.3 present the in-sample errors (MRAE and RMSE) for all the three models. As one can see, RVM has the worst performance (i.e., yields the worst value for an error metric) among the three models 2 out of 18 times according to MRAE and 4 out of 18 times according to RMSE, while TFJ has the worst performance 12 out of 18 times for both the error metrics. TF has the worst error 4 out of 18 times according to MRAE and 2 out of 18 according to RMSE. Tables 5.4 and 5.5 present out-of-sample error metrics. When it comes to MRAE, RVM has the worst performance error only 2 out of 18 times, while TFJ and TF have the worst errors 9 times and 7 times respectively. For RMSE, RVM is the worst model out-of-sample only once, while TFJ gives the worst RMSE error 8 times and TF gives the worst error 4 times. The out-of-sample RMSE errors given by at least two of three models are almost indistinguishable in five cases.

In summary, out of a total of 72 error comparisons, RVM is the worst model only in 9 cases, with 'ties' declared in five cases and one of the other two models being the worst model in the remaining 58 cases.

These results support the modest claim of this chapter that the newly proposed RVM (with a single, scalar stochastic process and one random variable) performs at least equally well as more involved models discussed in the literature (with two or more scalar stochastic processes), when the comparison involves predictive ability in terms of one step ahead prediction of the prices of futures contracts. The advantage of RVM over the other two models is its simplicity of calibration and parsimony in terms of parameters. To be more specific, TFJ model has 7 parameters to be calibrated offline and 3 risk premium parameters which are calibrated online. RVM has only 3 offline parameters and a single risk premium parameter which is calibrated online. These numbers exclude 4 seasonality parameters and the measurement noise variances, which have to be calibrated for both the types of models.

When it comes to pricing European style securities, RVM contains a single random variable σ and one stochastic process. Hence it is far easier to simulate from, than the two other models considered here (TF has two correlated

random processes, while TFJ has two correlated random processes and one compound Poisson process). This indicates that pricing of any exotic, European style options via Monte Carlo simulation is computationally far cheaper with RVM, when compared to TF or TFJ.

In addition, an empirical analysis was carried out for two cases for two-factor model with jumps: when the jump risk premium is set to zero and when the jump risk premium is given with formula (2.44). Tables 5.6-5.7 show this comparison for out-of-sample data sets for MRAE and RMSE errors. A comparison was provided only for out-of-sample data since the risk premium was updated online using particle filter in the numerical experiments, as outlined earlier. It can be seen that using explicitly parametrised jump risk premium does not improve the predictive ability, at least for the data sets used, with an improvement in one of the error metrics observed only 17 out of 36 times. This modest set of numerical experiments does not provide any evidence of practical utility of assuming the jump risk to be non-idiosyncratic in electricity markets. However, it is quite conceivable that contrary empirical evidence may be found with pricing other securities with jump diffusion models.

Figures 5.7-5.15 in Appendix A show the evolution of the measurement covariance terms. One can see that, large values appear in two scenarios, either at the end of in-sample data (150 days) or closer to the end of out-of-sample data. This shows, that the electricity price spot price models perform well in-sample. However, the model should be re-calibrated more frequently for a good out-of-sample performance.

Figures 5.16-5.21 in Appendix A show the evolution of the risk premium parameters h_x, h_ζ and β over time. One can see that the price of risk of the mean-reversion process (i.e h_x) is dominating on the plots. At the same time, the GBM risks are low in comparison to the former. The jump risk parameters β are not high in any of the plots for the two-factor model with jumps, again indicating that the use of jump risk premium may not be practically important.

Figures 5.22-5.24 in Appendix A show risk premium evolution for the RVM model. One can see that level of risk premium for this model is much lower in comparison to the previously discussed models.

All the plots supporting the experiments can be found in the Appendix A.

Table 5.2: In-Sample MRAE results

T(days)	22	44	66	88	110	132
Experiment 1						
TFJ	**3.90**	**2.05**	2.29	2.60	1.65	3.14
TF	3.74	1.99	**2.39**	**2.78**	1.75	2.92
RVM	2.22	1.99	2.34	2.65	**3.29**	**3.97**
Experiment 2						
TFJ	**6.12**	3.56	**5.48**	**7.05**	**5.89**	**5.44**
TF	5.97	**4.53**	5.01	5.21	4.00	4.28
RVM	1.39	1.33	1.28	1.24	1.57	2.13
Experiment 3						
TFJ	5.54	**5.95**	**9.77**	**11.74**	**12.68**	**14.40**
TF	**5.70**	4.57	3.42	2.86	3.11	3.37
RVM	1.93	1.76	1.65	1.53	1.60	2.38

Table 5.3: In-Sample RMSE results

T(days)	22	44	66	88	110	132
Experiment 1						
TFJ	**0.17**	**0.10**	0.11	0.12	0.08	0.14
TF	0.16	0.09	0.11	0.12	0.09	0.13
RVM	0.11	0.09	**0.12**	**0.14**	**0.17**	**0.20**
Experiment 2						
TFJ	**0.25**	0.15	**0.26**	**0.35**	**0.34**	**0.32**
TF	0.23	**0.18**	0.20	0.21	0.17	0.18
RVM	0.06	0.06	0.06	0.06	0.08	0.10
Experiment 3						
TFJ	0.26	**0.42**	**0.75**	**1.01**	**1.21**	**1.37**
TF	**0.28**	0.21	0.15	0.14	0.14	0.15
RVM	0.09	0.10	0.11	0.10	0.12	0.15

Table 5.4: Out-of-Sample MRAE results

T(days)	22	44	66	88	110	132
Experiment 1						
TFJ	**2.93**	2.51	1.89	1.92	2.87	**4.28**
TF	2.89	**2.63**	**1.98**	2.03	2.63	3.88
RVM	2.32	1.25	1.22	**2.26**	**2.92**	3.39
Experiment 2						
TFJ	6.08	4.13	3.41	**4.64**	3.91	**4.35**
TF	**7.65**	**5.74**	**4.51**	4.46	**4.08**	3.94
RVM	2.09	1.72	1.17	1.15	1.32	1.98
Experiment 3						
TFJ	4.60	**4.38**	**3.67**	**4.77**	**4.09**	**4.88**
TF	**5.61**	3.13	2.76	3.34	2.99	3.06
RVM	1.87	1.56	1.03	1.19	1.35	2.21

Table 5.5: Out-of-Sample RMSE results

T(days)	22	44	66	88	110	132
Experiment 1						
TFJ	**0.16**	**0.12**	**0.09**	0.10	**0.14**	**0.20**
TF	0.15	**0.12**	**0.09**	0.10	0.12	0.18
RVM	0.10	0.06	0.06	**0.11**	**0.14**	0.16
Experiment 2						
TFJ	0.31	0.21	0.15	**0.19**	**0.17**	**0.19**
TF	**0.39**	**0.29**	**0.20**	**0.19**	**0.17**	0.17
RVM	0.11	0.09	0.06	0.06	0.06	0.09
Experiment 3						
TFJ	0.23	**0.21**	**0.17**	**0.19**	**0.18**	**0.21**
TF	**0.29**	0.15	0.13	0.17	0.14	0.14
RVM	0.09	0.07	0.06	0.08	0.08	0.11

Table 5.6: Out-Of-Sample MRAE results - with and without jump risk premium

T(days)	22	44	66	88	110	132
Experiment 1						
TFJ ($R(\lambda, \sigma_J, \beta) = 0$)	**3.06**	2.27	1.71	**2.07**	2.67	3.84
TFJ	2.93	**2.51**	**1.89**	1.92	**2.87**	**4.28**
Experiment 2						
TFJ ($R(\lambda, \sigma_J, \beta) = 0$)	**6.14**	4.09	**3.66**	**4.84**	**4.00**	**4.37**
TFJ	6.08	**4.13**	3.41	4.64	3.91	4.35
Experiment 3						
TFJ ($R(\lambda, \sigma_J, \beta) = 0$)	4.56	**4.40**	**4.11**	**5.03**	**4.36**	4.59
TFJ	**4.60**	4.38	3.67	4.77	4.09	**4.88**

Table 5.7: Out-Of-Sample RMSE results - with and without jump risk premium

T(days)	22	44	66	88	110	132
Experiment 1						
TFJ($R(\lambda, \sigma_J, \beta) = 0$)	0.15	0.11	0.08	0.10	0.13	0.18
TFJ	**0.16**	**0.12**	**0.09**	0.10	**0.14**	**0.20**
Experiment 2						
TFJ($R(\lambda, \sigma_J, \beta) = 0$)	**0.32**	**0.22**	**0.16**	**0.20**	0.17	0.19
TFJ	0.31	0.21	0.15	0.19	0.17	0.19
Experiment 3						
TFJ($R(\lambda, \sigma_J, \beta) = 0$)	0.23	**0.22**	**0.19**	**0.21**	**0.19**	0.19
TFJ	0.23	0.21	0.17	0.19	0.18	**0.21**

Table 5.8: Parameters for a random volatility model

N	κ	μ	η
1	2.6448	0.8800	0.3897
2	0.8556	2.4659	-0.0122
3	2.5268	0.5902	0.5982

Table 5.9: Parameters for a two-factor model without jumps

N	κ	μ	σ_1	σ_2	ρ
1	0.4200	0.0010	0.1613	0.2015	-0.1019
2	0.4151	0.0498	0.1006	0.1480	-0.0557
3	0.4940	0.0010	0.1839	0.1576	-0.5371

Table 5.10: Parameters for a two-factor model with jumps

N	κ	μ	σ_1	σ_2	ρ	λ_J	μ_J
1	0.7117	1.2316	0.1315	0.1092	-0.2839	0.2852	0.4843
2	0.7275	0.3259	0.1829	0.1215	-0.3210	2.6123	0.0305
3	0.6928	1.2927	0.2145	0.1217	-0.6338	10.7524	0.0303

Table 5.11: Seasonality parameters

N	c_1	ς	c_2	c_3
1	3.8167	0.2778	6.1375	8.7799
2	3.7073	-0.2773	4.5499	12.4057
3	4.0060	0.2798	5.9689	9.9754

5.5 Summary

To summarise, the research presented in this chapter makes three main contributions:

- A new random volatility model was proposed for log spot price in the electricity market, which might be useful in modelling other commodities as well. The model is significantly easier to calibrate and to simulate from, as compared to two-factor models with and without jumps and performs at least as well as these models in the comprehensive numerical experiments from the real electricity market data. This model has the potential to be practically very useful in pricing applications for the electricity market.

- A new systematic multi-step procedure was proposed for calibrating two-factor models with or without jumps, which alleviates some of the difficulty in calibrating models with a large number of parameters. The use of this new procedure has also been tested through numerical experiments.

- The empirical evidence on the use of jump models in electricity markets was added. The empirical results, using three data sets and two error metrics, provide no conclusive evidence that the use of jumps in modelling adds value in terms of prediction, especially out-of-sample. Further, the evidence does not suggest that modelling explicitly parametrised jump risk premium adds value in terms of out-of-sample prediction.

Chapter 6

Conclusion

6.1 Contributions

This section summarises the contributions of this thesis. The major achievement of this thesis was to show that models with random parameters have a useful role to play in financial applications. One of the advantages of these models is that they can be implemented in the variety of applications, but with less computational effort as compared to the existing models with the same number of sources of uncertainty. New models have less parameters, which reduces the risk of over-fitting the data. Specific contributions of different chapters are summarised below.

- In Chapter 3, a one factor model with a random long-run mean was proposed for commodity price modelling. The implementation of this model is based on the extension of the existing two factor model. The new approach is focused on substitution of one factor from two-factor model by a random variable. In addition, we treat a seasonality factor as a simple sinusoid. The proposed simplification allows to reduce a total number of parameters in the model. Hence, reduce the probability of over-fitting and increase the calibration quality. Comprehensive numerical experiments for three different energy commodities and five different models were performed. It was shown that the new model yields a similar level of accuracy as a two-factor model in most cases.

- In Chapter 4, a new random volatility model has been proposed, which is composed of a deterministic volatility term structure modulated by a

log-normal random variable. It is shown that this model can be derived rigorously as a special case of Hull White stochastic volatility framework. Extensive numerical experiments illustrate that the new model is far simpler to calibrate. Moreover, the comparison with existing stochastic volatility models shows at least 99% improvement in terms of required time for calibration. The proposed model is also implemented in Excel spreadsheet. The performance of the new model is at least as good as existing stochastic volatility models in terms of pricing accuracy. In addition, the new model yields similar results in pricing OTC options. Finally, the new model can be used for calculation of option's Delta and Gamma. The numerical results showed that the results are similar to the existing stochastic volatility models.

- In Chapter 5 the new futures pricing formulae were developed using a random volatility parameter for the spot price process. Empirical results show that the new one factor model with random volatility performs as good as existing two-factor models for electricity commodity.

  In addition, a new two stage calibration procedure was proposed. The offline stage of the calibration procedure allows to calibrate model parameters using moments of the data distribution. The online stage of the calibration procedure allowed to update risk parameters and covariance noise matrix during online stage of particle filter. implemented and tested for for the new models in this chapter.

  Finally, the empirical evidence on the use of compound Poisson process for electricity futures models was added. The extensive numerical experiments showed that there is no complete evidence in increasing of prediction accuracy by adding a jump term to the model. Authors also showed that there is no evidence in that the use of the parametrised jump-risk premium adds any value to the prediction accuracy.

6.2 Future research

It is important to outline the possible applications of the developed models and algorithms.

- The new stochastic volatility model introduced in chapter 4 can be used in pricing of currency options. To achieve this one can use a mean-reversion definition for a drift term in the definition of the SDE. In addition, different distributions for a random volatility term can be used. However, it is not necessary to use a parametric distribution for definition of the random term in the model. One can directly define the moments of the volatility term as parameters. A model parametrized directly in terms of moments will require semidefinite programming, since the moment parameters have to form a Hankel matrix which is positive semidefinite.

- The new option pricing model can be used in the high-frequency framework due to its fast calibration capabilities, which allows to recalibrate the model in seconds even while using less sophisticated computer software and hardware. However, it might work faster while used on advanced computing systems such as GPU clusters. Realisation of this model for the parallel computing can be twofold. Firstly, the calibration procedure for the new model should be chosen carefully. For example, the realisation of the weighted least-squares method can be challenging in the parallel framework. Secondly, the realisation of OTC pricing can be performed as well using multi-core systems, in case to speed-up the simulation routine.

- The new futures pricing model in chapter 5 can also be used to price futures on assets other than electricity, which also demonstrate price jumps over time (e.g. currency futures).

- The new random volatility model discussed in chapter 5 can be possibly used for pricing derivatives other than vanilla futures contracts. Moreover, options on futures is a possible topic for a new research. One can also try our approach for pricing interest rate derivatives, such as swaps, caplets and etc.

- The new two-step algorithm for calibration can be used in systems where one can't construct a likelihood function of a desired stochastic process in closed-form.

Appendix A

Expressions for the moments for two-factor process with jumps in chapter 5 are given here. For simplicity of writing, assume following quantities:

$$
\begin{aligned}
M_1 &= \zeta_0 + x_0 e^{-\kappa\Delta} + \mu\Delta, \\
M_2 &= \frac{\sigma_1^2}{2\kappa}(1 - e^{-2\kappa\Delta}) + \frac{\rho\sigma_1\sigma_2}{\kappa}\left(1 - e^{-\kappa\Delta}\right) + \sigma_2^2\Delta, \\
L_1 &= e^{-2\kappa\Delta}\Delta\lambda\sigma_J^2, \\
L_2 &= e^{-4\kappa\Delta}\Delta\lambda\sigma_J^4, \\
L_3 &= 15e^{-6\kappa\Delta}\Delta\lambda\sigma_J^6, \\
L_4 &= 105e^{-8\kappa\Delta}\Delta\lambda\sigma_J^8 + 315e^{-8\kappa\Delta}(\Delta\lambda)^2\sigma_J^8.
\end{aligned}
$$

Now, using above equations, first nine moments of the two-factor process with jumps have the following view:

$$
\begin{aligned}
m_1 &= M_1, \\
m_2 &= M_1^2 + M_2 + L_1, \\
m_3 &= M_1(3L_1 + M_1^2 + 3M_2), \\
m_4 &= 3L_1^2 + 3L_2 + M_1^4 + 6M_1^2M_2 + 3M_2^2 + 6L_1(M_1^2 + M_2), \\
m_5 &= M_1(15L_1^2 + 15L_2 + M_1^4 + 10M_1^2M_2 + 15M_2^2 + 10L_1(M_1^2 + 3M_2)), \\
m_6 &= L_2 + 45L_2M_1^2 + M_1^6 + 45L_2(L_1 + M_2) + 15M_1^4(L_1 + M_2) + 45M_1^2(L_1 + M_2)^2 + \\
&\quad 15(L_1 + M_2)^3, \\
m_7 &= M_1(21L_3 + 105L_2M_1^2 + M_1^6 + 315L_2(L_1 + M_2) + 21M_1^4(L_1 + M_2) \\
&\quad + 105M_1^2(L_1 + M_2)^2 + 105(L_1 + M_2)^3),
\end{aligned}
$$

$$m_8 = L_4 + 28L_3M_1^2 + 210L_2M_1^4 + M_1^8 + 28L_3(L_1 + M_2) + 1260L_2M_1^2(L_1 + M_2) + 28M_1^6(L_1 + M_2) + 630L_2(L_1 + M_2)^2 + 210M_1^4(L_1 + M_2)^2 + 420M_1^2(L_1 + M_2)^3 + 105(L_1 + M_2)^4,$$

$$m_9 = M_1(9L_4 + 84L_3M_1^2 + 378L_2M_1^4 + M_1^8 + 252L_3(L_1 + M_2) + 3780L_2M_1^2(L_1 + M_2) + 36M_1^6(L_1 + M_2) + 5670L_2(L_1 + M_2)^2 + 378M_1^4(L_1 + M_2)^2 + 1260M_1^2(L_1 + M_2)^3 + 945(L_1 + M_2)^4).$$

Note, that the expressions for the moments for the two-factor model without jumps can be obtained by setting jump parameters (λ and σ_J) to zero.

Figures for the electricity commodity experiments in chapter 5 are presented here. Here EDBM is error distribution between models.

Figure 5.1: In-Sample EDBM across contracts (experiment 1)
Bar chart represents the difference in the error metrics across the models.

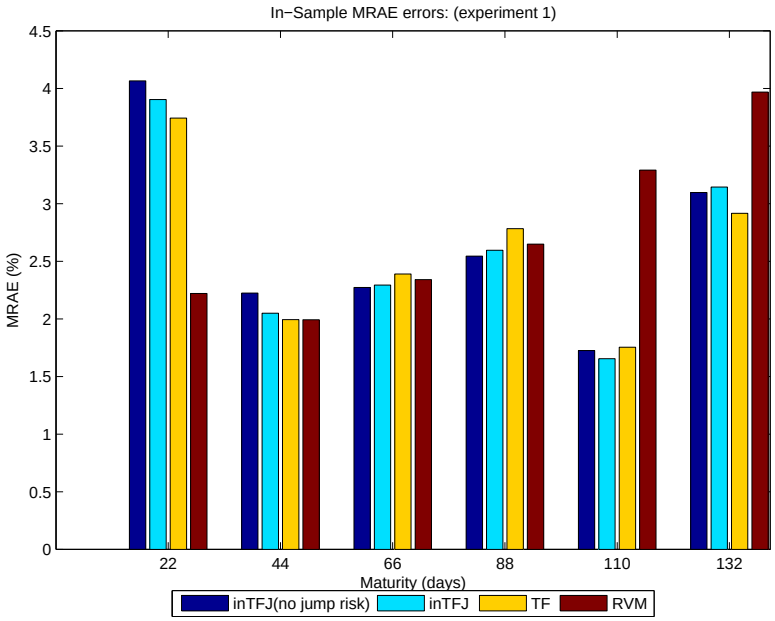


Figure 5.2: In-Sample EDBM across contracts (experiment 2)
Bar chart represents the difference in the error metrics across the models.

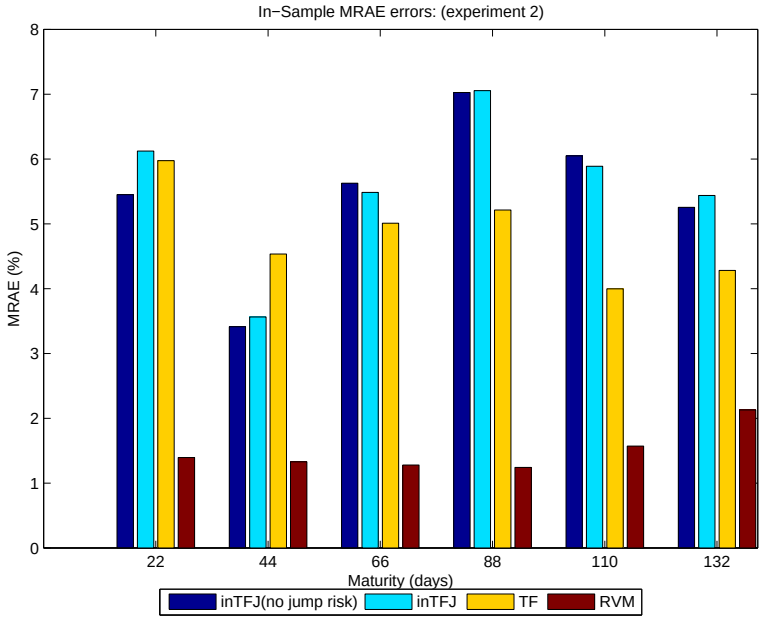


Figure 5.3: In-Sample EDBM across contracts (experiment 3)
Bar chart represents the difference in the error metrics across the models.

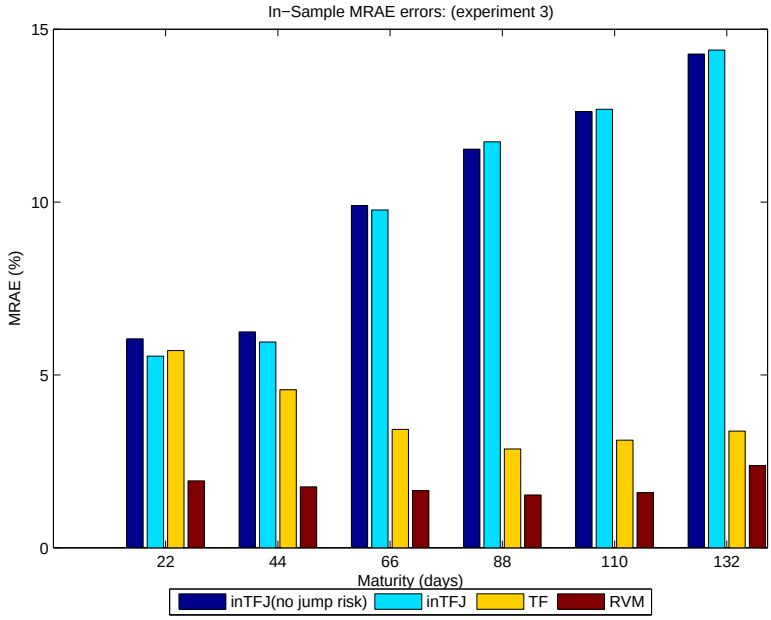


Figure 5.4: Out-of-Sample EDBM across contracts (experiment 1)
Bar chart represents the difference in the error metrics across the models.

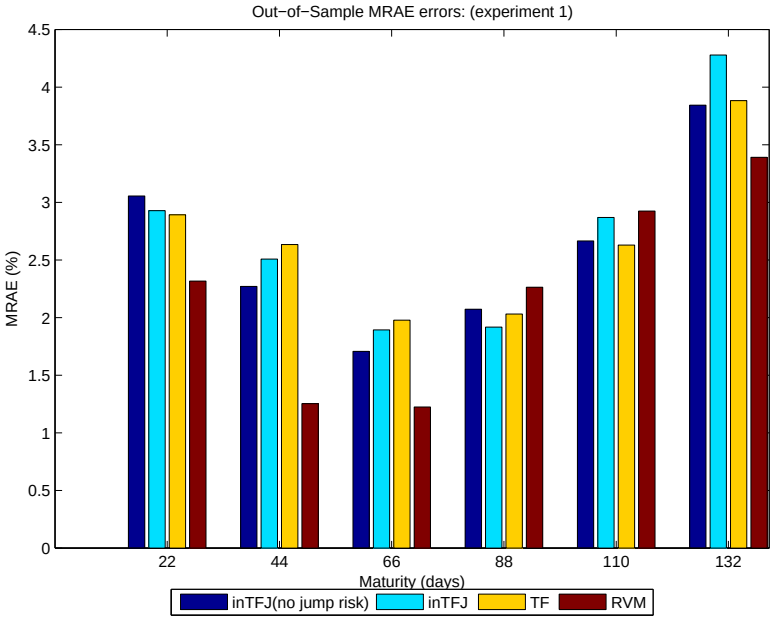


Figure 5.5: Out-of-Sample EDBM across contracts (experiment 2)
Bar chart represents the difference in the error metrics across the models.

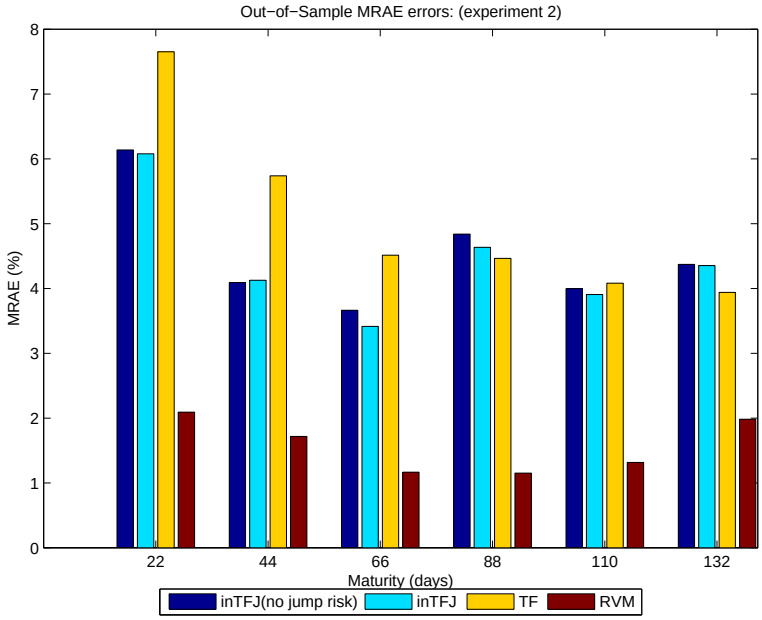


Figure 5.6: Out-of-Sample EDBM across contracts (experiment 3)
Bar chart represents the difference in the error metrics across the models.

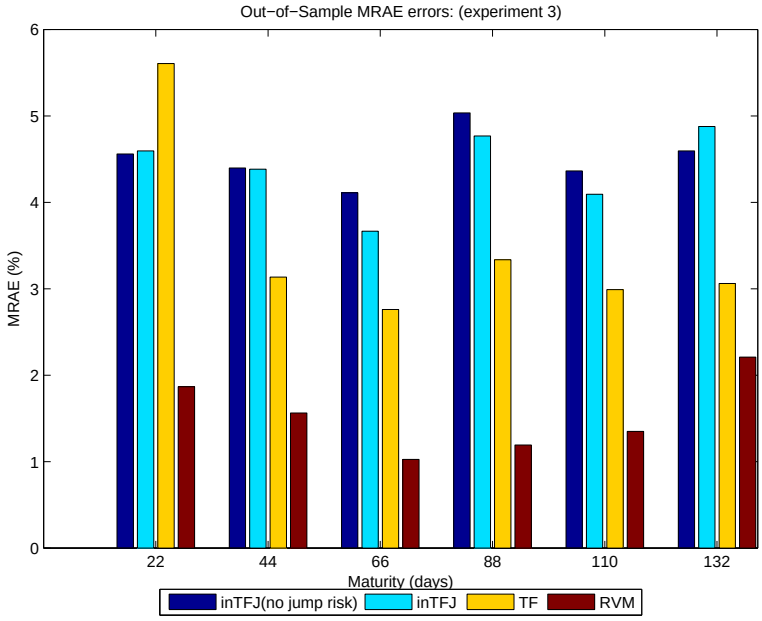


Figure 5.7: Two-factor model with jumps: Noise variance evolution (experiment 1)
The graph shows the evolution of measurement noise variance over time for different futures contracts.

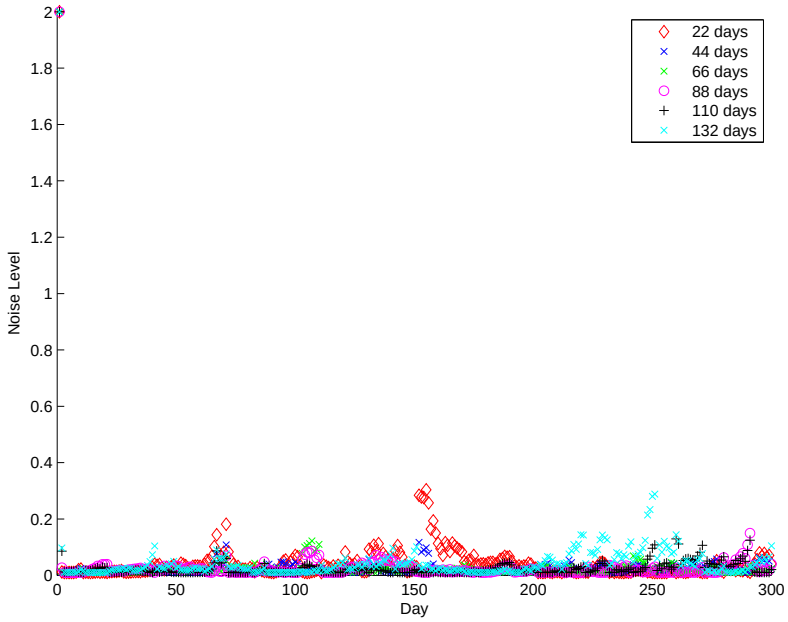


Figure 5.8: Two-factor model with jumps: Noise variance evolution (experiment 2)
The graph shows the evolution of measurement noise variance over time for different futures contracts.

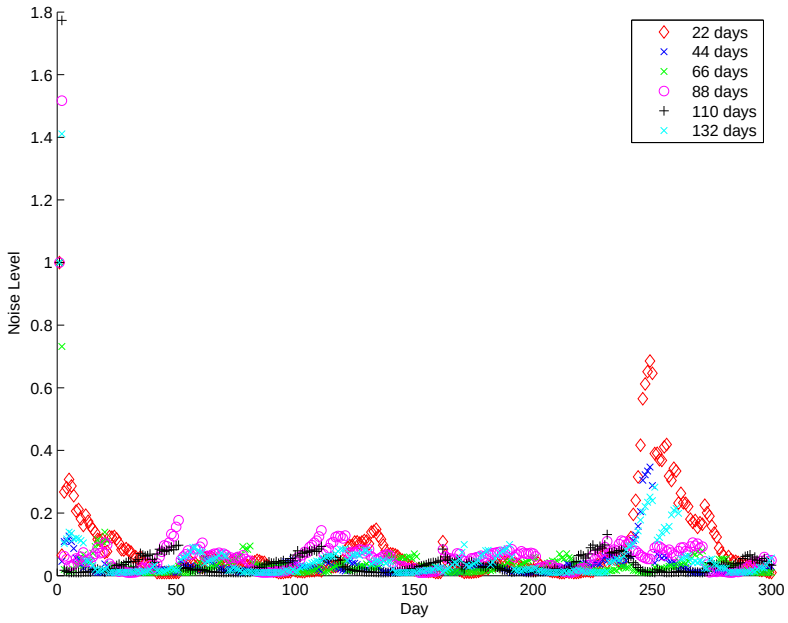


Figure 5.9: Two-factor model with jumps: Noise variance evolution (experiment 3)
The graph shows the evolution of measurement noise variance over time for different futures contracts.

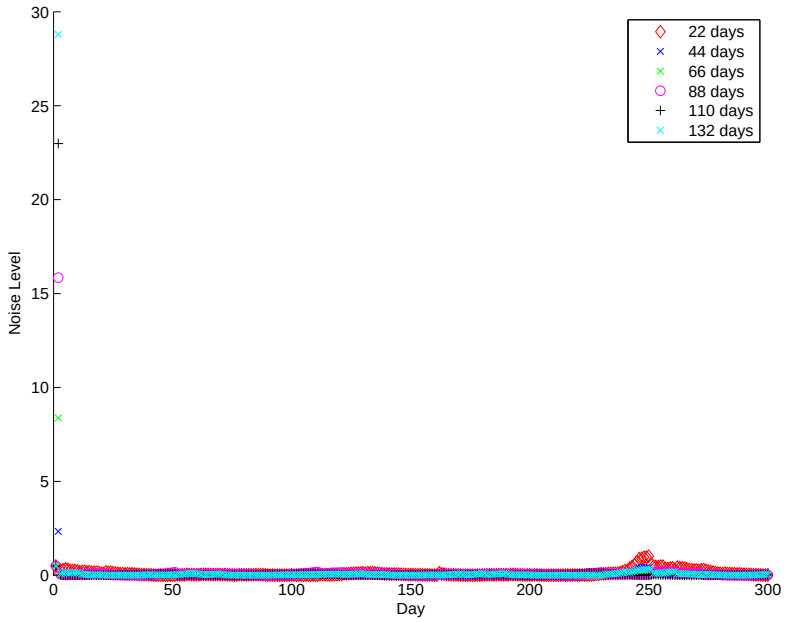


Figure 5.10: Two-factor model without jumps: Noise variance evolution (experiment 1)
The graph shows the evolution of measurement noise variance over time for different futures contracts.

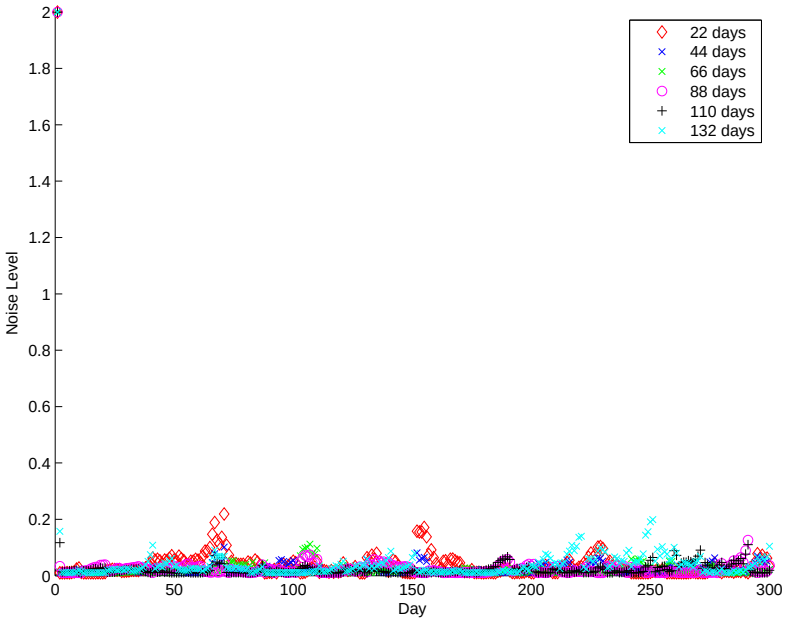


Figure 5.11: Two-factor model without jumps: Noise variance evolution (experiment 2)
The graph shows the evolution of measurement noise variance over time for different futures contracts.

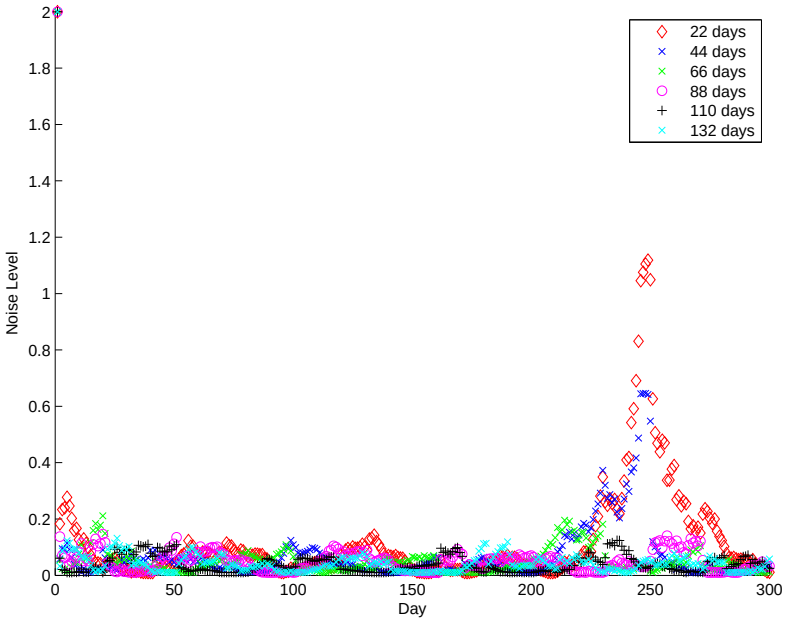


Figure 5.12: Two-factor model without jumps: Noise variance evolution (experiment 3)
The graph shows the evolution of measurement noise variance over time for different futures contracts.

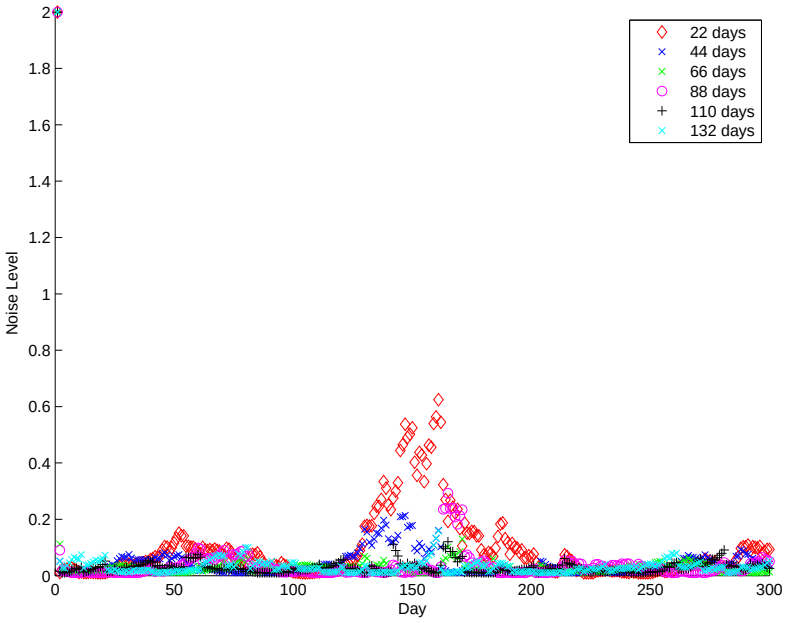


Figure 5.13: RVM model: Noise variance evolution (experiment 1)
The graph shows the evolution of measurement noise variance over time for different futures contracts.

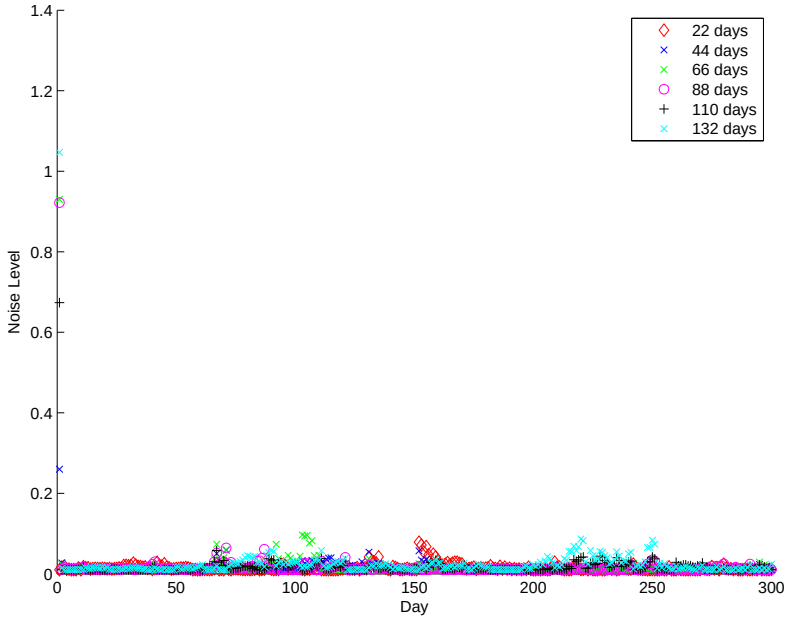


Figure 5.14: RVM model: Noise variance evolution (experiment 2)
The graph shows the evolution of measurement noise variance over time for different futures contracts.

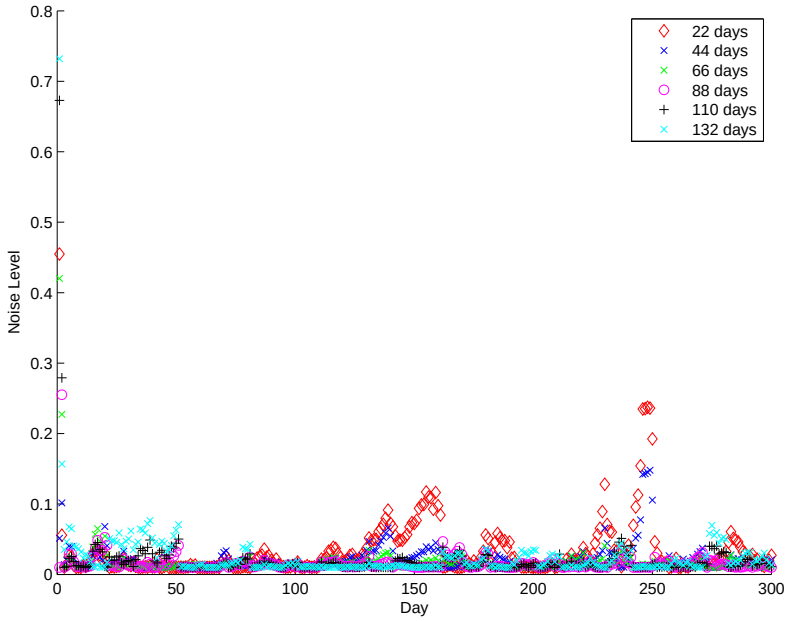


Figure 5.15: RVM model: Noise variance evolution (experiment 3)
The graph shows the evolution of measurement noise variance over time for different futures contracts.

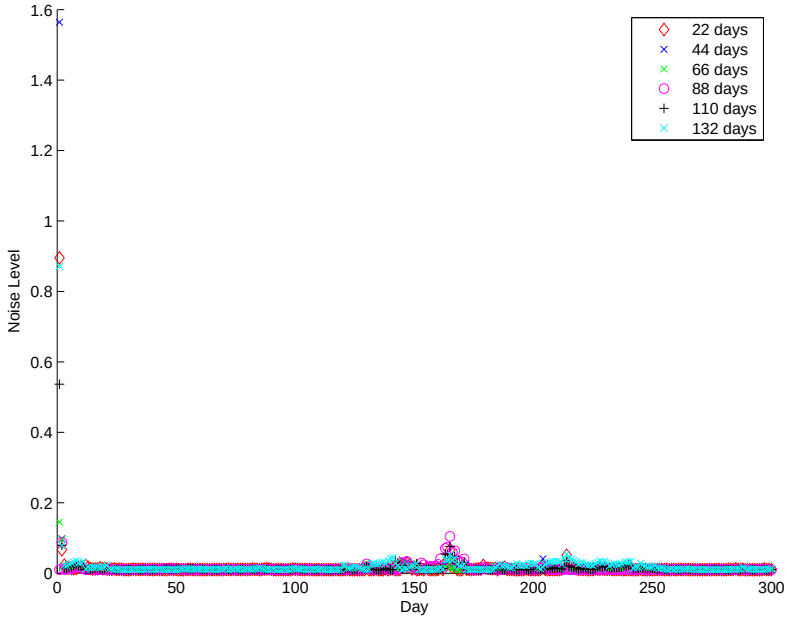


Figure 5.16: Two-factor model with jumps: Risk premium evolution (experiment 1)
The graph shows the evolution of risk premium parameters over time, recalibration time is set to 10 days.

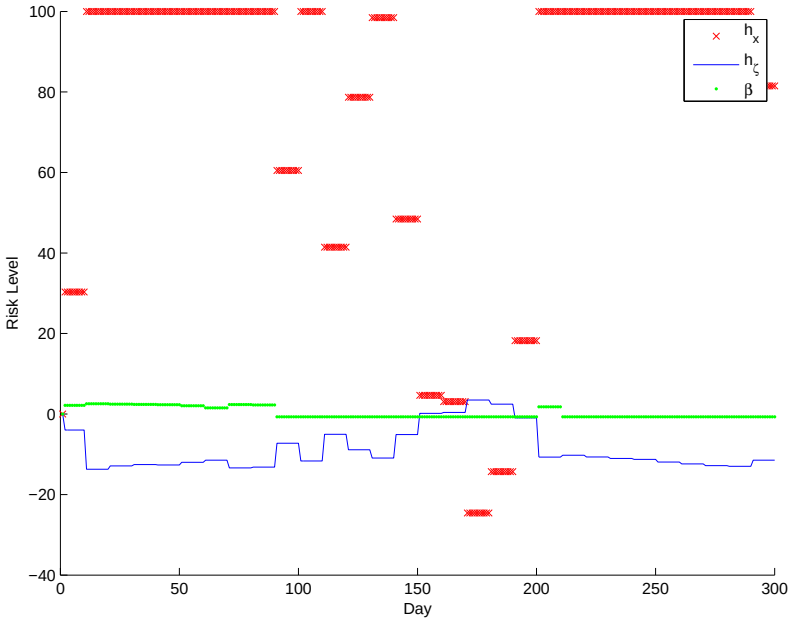


Figure 5.17: Two-factor model with jumps: Risk premium evolution (experiment 2)
The graph shows the evolution of risk premium parameters over time, recalibration time is set to 10 days.

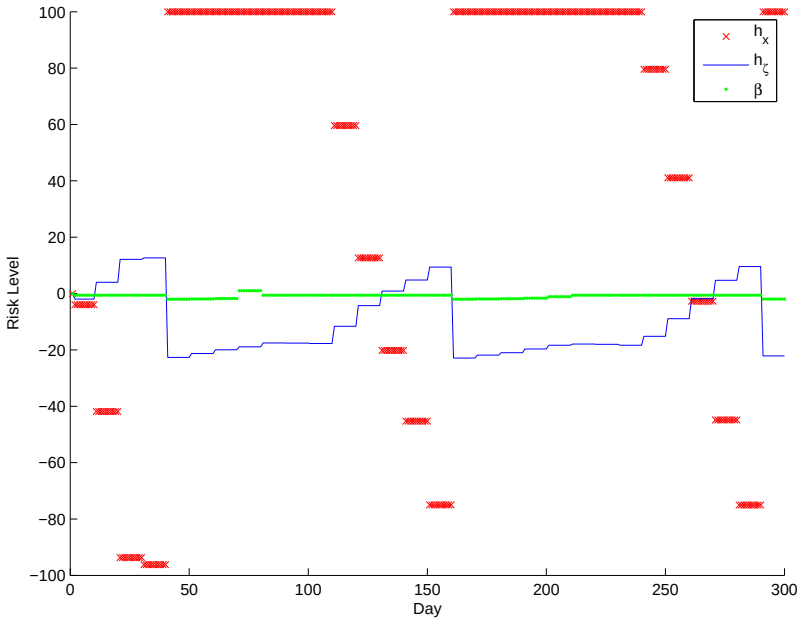


Figure 5.18: Two-factor model with jumps: Risk premium evolution (experiment 3)
The graph shows the evolution of risk premium parameters over time, recalibration time is set to 10 days.

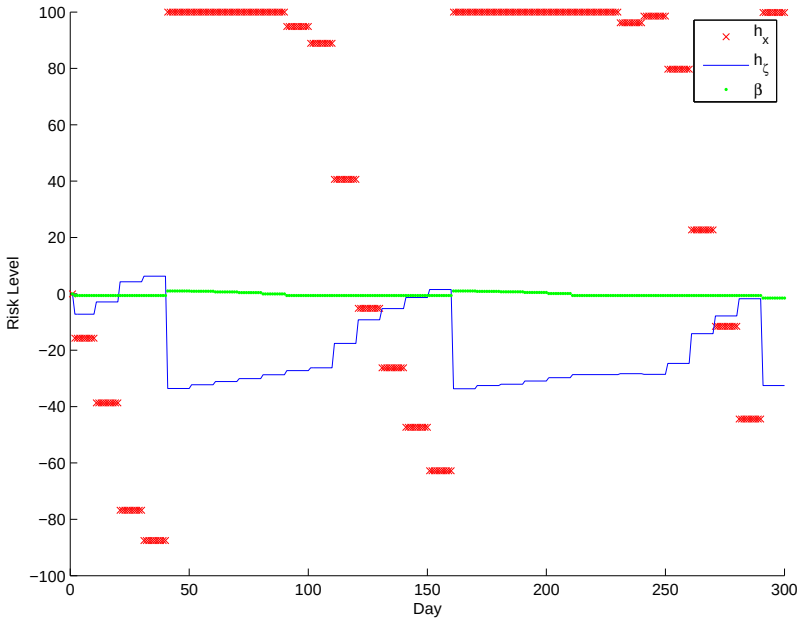


Figure 5.19: Two-factor model without jumps: Risk premium evolution (experiment 1)
The graph shows the evolution of risk premium parameters over time, recalibration time is set to 10 days.

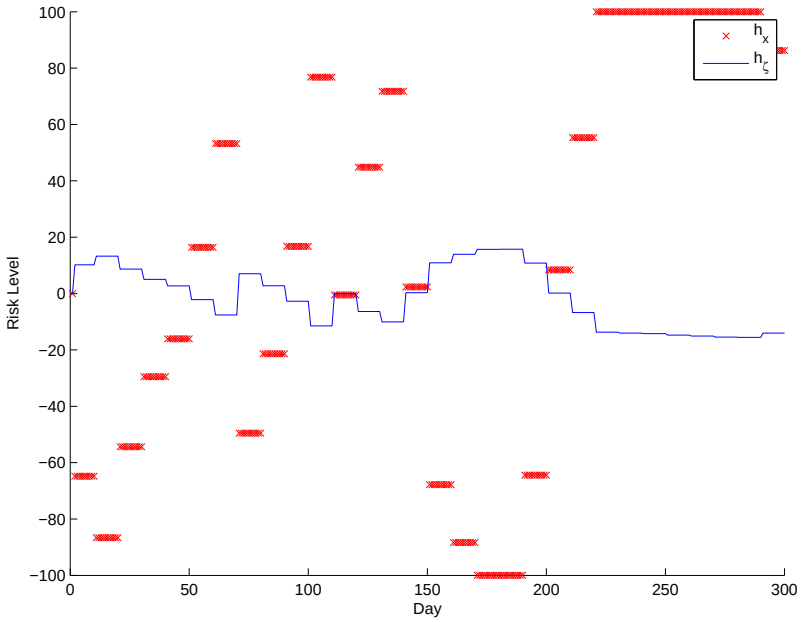


Figure 5.20: Two-factor model without jumps: Risk premium evolution (experiment 2)
The graph shows the evolution of risk premium parameters over time, recalibration time is set to 10 days.

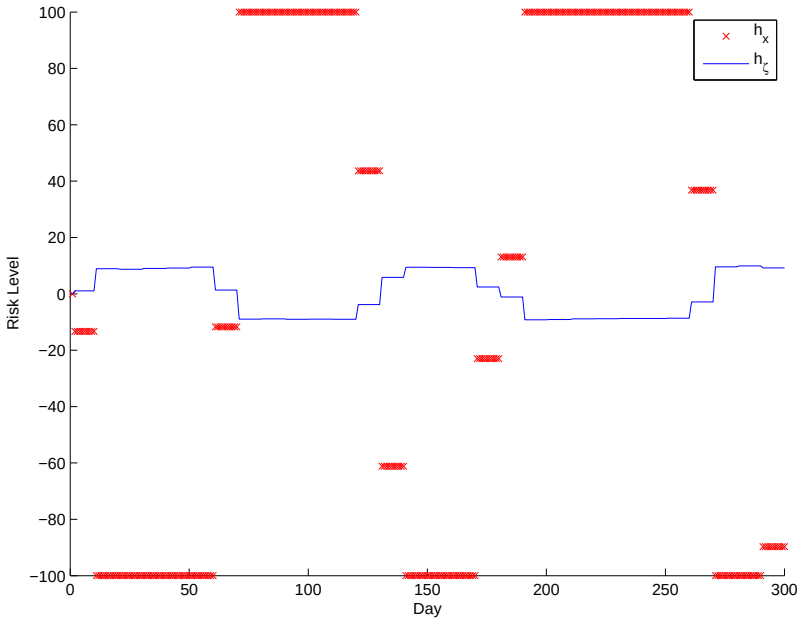


Figure 5.21: Two-factor model without jumps: Risk premium evolution (experiment 3)
The graph shows the evolution of risk premium parameters over time, recalibration time is set to 10 days.

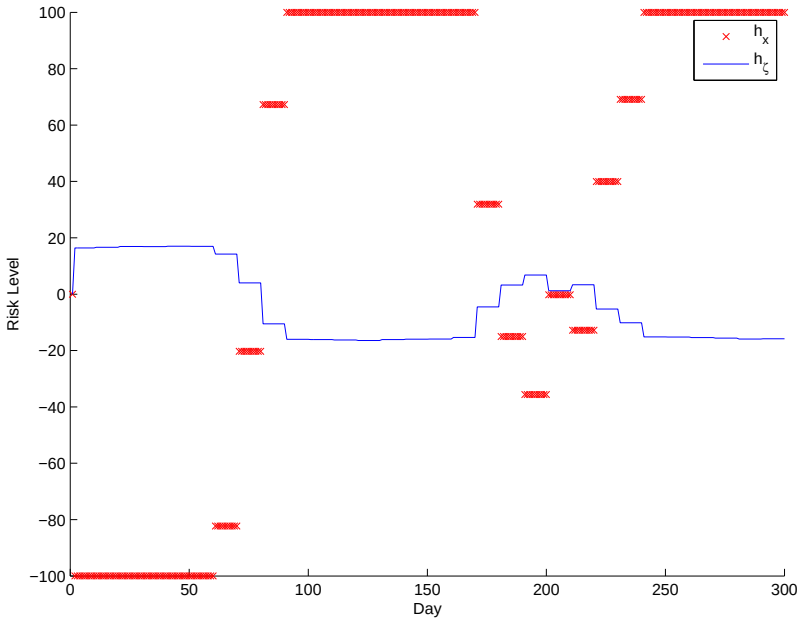


Figure 5.22: RVM model: Risk premium evolution (experiment 1)
The graph shows the evolution of risk premium parameters over time, recalibration time is set to 10 days.

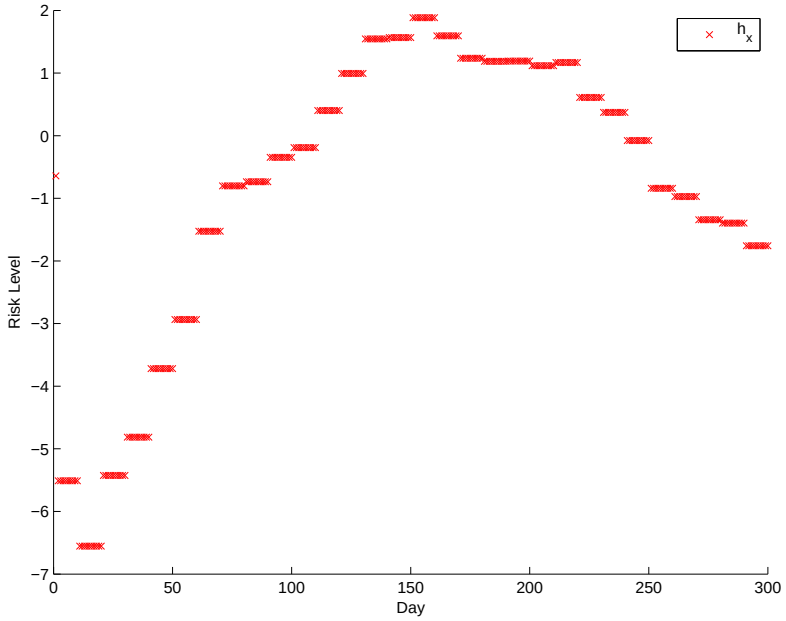


Figure 5.23: RVM model: Risk premium evolution (experiment 2)
The graph shows the evolution of risk premium parameters over time, recalibration time is set to 10 days.

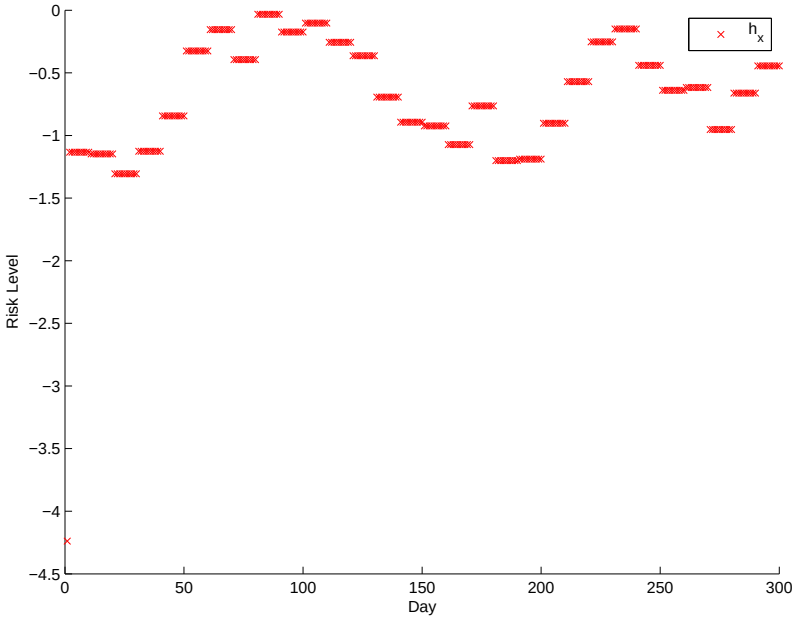
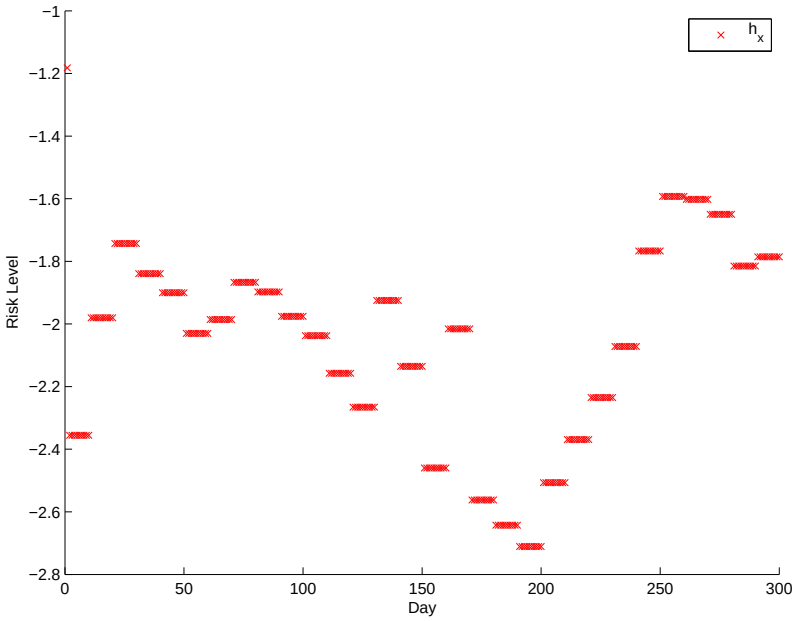


Figure 5.24: RVM model: Risk premium evolution (experiment 3)
The graph shows the evolution of risk premium parameters over time, recalibration time is set to 10 days.



Appendix B

Technical documentation to the developed software is provided here. The CD has following file structure:

- Folder **ElectricityCommodity** contains Matlab files for the experiments discussed in Chapter 5
- Folder **EnergyCommodity** contains Matlab files for the experiments discussed in Chapter 3
- Folder **OptionPricing** contains Matlab files for the experiments discussed in Chapter 4

Technical documentation for storable commodities experiment

Kalman Filter Matlab Experiment has the following structure: Scripts Run1.m and Run2.m access the data from attached .mat files and call function main.m. The two script files refer to two data panels mentioned in the paper. Function main.m performs two step procedure: for transferred function pointer perform calibration of parameters and one step ahead error estimation.

Functions: One-state models:

- onestatefilter.m - one state model without price of risk - used as a first step of optimisation for the nested algorithm
- onestatefilter_rf.m - one state models with price of risk (OF)
- seasonalityfilter.m - one state model with seasonality (OFS)

Two-state models:

- twostatefilter.m - two-state model without seasonality (TF)
- twostatefilterseasonality.m - two-state model with seasonality (TFS)

All the Functions have 2 types of output depending on tag value passed to the function.

All the computation was performed in Matlab R2011b, using optimization toolbox. The exact dates for Bloomberg data are given in table 3.1, section 3.4.1 (day closing prices considered in all cases).

Technical documentation for option pricing experiments

OptionPricing folder has following sub folders:

- Bates - contains files regarding Bates model:
  - **BatesFFT.m**: function to calculate European Call using Bates model, through Fast Fourier Transform. Input parameters are:
    - KAPPA - κ,
    - THETA - θ,
    - SIGMAv - σ_v,
    - RHO - ρ,
    - V0 - v_0,
    - LAMBDA - λ_p,
    - MUJ - μ_j,
    - SIGMAS - $sigma_J$,
    - r - risk-free rate,
    - q - dividend yield,
    - T - time to expiry,
    - S0 - Spot price,
    - K - strike price;
  - **BCCDeltafft.m**: function to calculate the objective function for the weighted least squares method. Input parameters:
    - c - vector of parameters for the model,

* C - vector of European Call prices,
* S - Spot price,
* K - vector of Strike prices,
* T - vector of option maturities,
* r - vector of risk-free rates,
* w - vector of weights;

– **sample_run_script_Bates.m** - script launching the experiment for the Bates model

– **rffix.m**: function is to calibrate the vector of from the data risk-free rates. Input:

* Data - Matrix, containing data abour options, with the following entries: Column 1: Strike Price, Column 2: Time to maturity, Column 3: Bid value of the option, Column 4: Ask value of the option
* S - Spot price

– **rmin.m** function to calculate the objective function value for risk calibration. Input:

* C - vector of European Call prices,
* S - Spot price,
* K - vector of Strike prices,
* T - vector of option maturities,
* r - vector of risk-free rates,
* sigma - vector of Implied Volatilities from market

• Heston - contains files regarding to the Heston model experiments:

– **HestonCallFft.m**: unction to calculate European Call using Bates model, through Fast Fourier Transform. Input parameters are:

* KAPPA - κ,
* THETA - θ,
* SIGMAv - σ_v,
* RHO - ρ,
* V0 - v_0,

    * r - risk-free rate,
    * q - dividend yield,
    * T - time to expiry,
    * S0 - Spot price,
    * K - strike price;
  - **rffix.m**: function is to calibrate the vector of from the data risk-free rates. Input:
    * Data - Matrix, containing data abour options, with the following entries: Column 1: Strike Price, Column 2: Time to maturity, Column 3: Bid value of the option, Column 4: Ask value of the option
    * S - Spot price
  - **rmin.m** function to calculate the objective function value for risk calibration. Input:
    * C - vector of European Call prices,
    * S - Spot price,
    * K - vector of Strike prices,
    * T - vector of option maturities,
    * r - vector of risk-free rates,
    * sigma - vector of Implied Volatilities from market

- MSV - contains files regarding to the MSV model experiments:
  - **sample_run_script.m** script designed to run experiments. Additional guidance is inside the script.
  - **MSV.m** is the option pricing function for the MSV model. Input:
    * c - vector of the parameters for MSV model
    * K - strike price vector
    * T - maturity times vector
    * S - the spot price
    * r - risk-free rate vector
  - The rest files are the same as in previous folders

- Hedging - folder contains files for hedging experiments.

- **Hedging.m** function has two outputs: Delta and Gamma values. This function is calculating numerical derivative for the given function with respect to Spot price. Inputs:
  - funchandle - function name to use,
  - S - current spot price,
  - n - change in the derivative
- **sample_run_script_hedging.m** sample script to run hedging experiment

- Simulation folder has two sub folders one for 'Asian' option pricing and one for 'Barrier' option pricing.

  Barrier folder:

  - **batesmontecarlo.m** : function to calculate 'Barrier' option price using Bates model.

    Input:

    - KAPPA - κ,
    - THETA - θ,
    - SIGMAv - σ_v,
    - RHO - ρ,
    - V0 - v_0,
    - LAMBDA - λ_p,
    - MUJ - μ_j,
    - SIGMAS - $sigma_J$,
    - r - risk-free rate,
    - T - time to expiry,
    - S - Spot price,
    - K - strike price;
    - PATHS - number of price paths;
    - n - number of simulations;
    - barrier - barrier value;

    Output:

    - Price - Option price value

* conf - confidence interval
* Time - time taken

– **MSVmontecarlo.m** : function to calculate 'Barrier' option price using MSV model.

Input:

* s0 - σ_0,
* s1 - σ_1,
* s2 - σ_2,
* lambda - λ,
* k,
* K - strike price;
* T - time to expiry,
* S - Spot price,
* r - risk-free rate,
* PATHS - number of price paths;
* n - number of simulations;
* barrier - barrier value;

Output:

* Price - Option price value,
* conf - confidence interval,
* Time - time taken,

– **hestonmontecarlo.m** : function to calculate 'Barrier' option price using Heston model.

Input:

* KAPPA - κ,
* THETA - θ,
* SIGMAv - σ_v,
* RHO - ρ,
* V0 - v_0,
* r - risk-free rate,
* T - time to expiry,

* S - Spot price,
* K - strike price;
* PATHS - number of price paths;
* n - number of simulations;
* barrier - barrier value;

Output:

* Price - Option price value,
* conf - confidence interval,
* Time - time taken,

– **sample_simulation_script.m**: script to run the experiment with above functions.

Asian folder:

– **batesmontecarlo_asian.m** : function to calculate 'Asian' option price using Bates model.

Input:

* KAPPA - κ,
* THETA - θ,
* SIGMAv - σ_v,
* RHO - ρ,
* V0 - v_0,
* LAMBDA - λ_p,
* MUJ - μ_j,
* SIGMAS - $sigma_J$,
* r - risk-free rate,
* T - time to expiry,
* S - Spot price,
* K - strike price;
* PATHS - number of price paths;
* n - number of simulations;

Output:

* Price - Option price value

* conf - confidence interval
* Time - time taken

– **MSV_asian.m**: function to calculate 'Asian' option price using MSV model.

Input:

* s0 - σ_0,
* s1 - σ_1,
* s2 - σ_2,
* lambda - λ,
* k,
* K - strike price;
* T - time to expiry,
* S - Spot price,
* r - risk-free rate,
* PATHS - number of price paths;
* n - number of simulations;

Output:

* Price - Option price value,
* conf - confidence interval,
* Time - time taken,

– **hestonmontecarlo_asian.m** : function to calculate 'Asian' option price using Heston model.

Input:

* KAPPA - κ,
* THETA - θ,
* SIGMAv - σ_v,
* RHO - ρ,
* V0 - v_0,
* r - risk-free rate,
* T - time to expiry,
* S - Spot price,

* K - strike price;
* PATHS - number of price paths;
* n - number of simulations;

Output:

* Price - Option price value,
* conf - confidence interval,
* Time - time taken,

– **sample_simulation_script.m**

Technical documentation for electricity commodity experiment

ElectricityCommodity folder has two sub folders:

- RVM folder:
  - **minimizeSeasonality.m** : Objective function for seasonality calibration. Input:
    * vec - vector of parameters
    * X - vector of the data
    * delta - timestep

    Output:
    * obj - value of the objective function (sum of squared residuals)
  - **resample.m**: This is a resampling function for a particle filter: Input:
    * particle - vector of particles
    * W - vector of corresponding particle weights

    Output:
    * particle - vector of updated particles
    * W - vector of updated corresponding particle weights
  - **futprice.m**: Function to calculate futures prices for the RVM model Input:

* delta - time step,
* t - current time step,
* T - vector of futures maturities,
* vec - vector of the model parameters,
* vecSeas - vector of seasonality parameters,
* xt - of value of underlying stochastic process,
* risk - risk parameters vector.

Output:

* F - vector of futures prices.

– **MLEndim.m**: Objective function for minimising RVM parameters. Input:

* nVec - initial parameters
* observationsY - observed futures prices
* observationsX - observed spot prices
* func - futures pricing function
* delta - timestep
* t - current time
* T - maturities vector
* vecSeas - seasonality parameters vector
* B - value of the jump integral
* observation - vector of current futures prices

Output:

* obj - value of the objective function

– **sample_experiment_RVM.m**: This script is designed to run an experiment with particle filter using the RVM model.

– **generatorR.m**: This is particle generator function: Input:

* vec - vector of parameters for thr model
* xt - value of the latent variable from the previous step
* delta - time step
* N - number of particles to generate

Output:

    * X - vector of particles

- **seasonality.m**: Seasonality function. Input:
    * c_1, c_2, c_3, c_4 - seasonality function parameters as in thesis
    * t - time

  Output:
    * seas - value of the seasonality function
- **minimize_risk.m**: Objective function for minimising risk parameters. Input:
    * fFut - futures price function,
    * delta - time step,
    * t - current time,
    * T - maturity vector,
    * vec - vector of parameters of the model,
    * vecSeas - vector of seasonality parameters,
    * xt - spot price value,
    * risk - vector of risk parameters,
    * observation - vector of current futures prices.

  Output:
    * obj - value of the objective function.
- **particle_filter.m**: This is particle filter algorithm for RVM model: Input:
    * func - futures price function for the filtration
    * vec - vector of parameters for the model
    * vecSeas - vector for seasonality components
    * observation - vector of observations
    * CM - Noise Matrix
    * delta - time step
    * t - current time
    * T - maturity times vector
    * S - the spot price
    * r - risk-free rate

* W - particle weight vector
* xt - value of the latent variable from the previous step
* risk - vector of old risk parameters
* st - step number

Output:

* risk - vector of updated risk parameters
* Noise - noise matrix
* X_filtered - value of the latent variable
* particle - particle vector
* W - particle weight vector
* RMSE - root mean square prediction error on this step
* MRAEi - mean absolute prediction error

- Two_factor_models folder:
  - **A_term.m** : Function to calculate the A term from the futures pricing formula. Input:
    * vec - vector of parameters for the model
    * risk - vector of risk parameters
    * tau - maturity time

    Output:
    * A - value of the A term
  - **minimize_risk.m** : Objective function for minimising risk parameters. Input:
    * fFut - futures price function,
    * delta - time step,
    * t - current time,
    * T - maturity vector,
    * vec - vector of parameters of the model,
    * vecSeas - vector of seasonality parameters,
    * xt - spot price value,
    * risk - vector of risk parameters,
    * observation - vector of current futures prices.

Output:

  * obj - value of the objective function.

- **excrisk.m** : Function to calculate risk parameters Input:

  * h - volatility risk
  * sigma - σ_1
  * beta - jump risk value
  * sj - jump size volatility (σ_J)
  * lambda - jump frequency (λ)

  Output:

  * R - vector or risk parameters

- **minimizeSeasonality.m** Objective function for seasonality calibration. Input:

  * vec - vector of parameters
  * X - vector of the data
  * delta - timestep

  Output:

  * obj - value of the objective function (sum of squared residuals)

- **particle_filter.m**: This is particle filter algorithm for RVM model: Input:

  * func - futures price function for the filtration
  * vec - vector of parameters for the model
  * vecSeas - vector for seasonality components
  * observation - vector of observations
  * CM - Noise Matrix
  * delta - time step
  * t - current time
  * T - maturity times vector
  * S - the spot price
  * r - risk-free rate
  * W - particle weight vector
  * xt - value of the latent variable from the previous step

* risk - vector of old risk parameters
* st - step number

Output:

* risk - vector of updated risk parameters
* Noise - noise matrix
* X_filtered - value of the latent variable
* particle - particle vector
* W - particle weight vector
* RMSE - root mean square prediction error on this step
* MRAEi - mean absolute prediction error

– **integral_jump.m**

– **MLEndim.m**

– **resample.m** This is a re sampling function for a particle filter: Input:

* particle - vector of particles
* W - vector of corresponding particle weights

Output:

* particle - vector of updated particles
* W - vector of updated corresponding particle weights

– **match_moments.m** Function for calibrating model parameters from the moments of data. Input:

* func - function to calculate moments
* Nmoments - number of moments used
* data - vector of spot prices
* delta - time step
* theta - vector of initial values of parameters
* Lambda - covariance matrix for the moments
* options - options for calibration
* calibration - method 'fmin' for fminsearch, 'lsq' for lsqnonlin

Output:

* theta - vector of moments

    * fval - value of the objective function after calibration

- **moments_of_data.m** Function for calculating moments from the data Input:
    * data
    * Nmoments - number of the moments to calculate

  Output:
    * M - vector of moments
- **sample_script_tf_with_jumps.m** This script is designed to run an experiment with particle filter using two-factor model with jumps.
- **momentsofOU2Fv2.m** Objective function for calculating analytical moments from the model. Input:
    * vec - vector of parameters for the model
    * delta - time step
    * L - number of the moments to calculate

  Output:
    * M - vector of the moments from the function
- **sample_script_tf_without_jumps.m** This script is designed to run an experiment with particle filter using Two-factor model without jumps.
- **minimizeMoments_lsq.m** Objective function for finding model parameters from the moments of data. Input:
    * M - vector of the moments from data
    * func - function that defines moments for given process
    * vec - vector of parameters for the model
    * xt - initial spot price
    * delta - time step

  Output:
    * M - vector of residuals
- **OUfutprice2F.m** Function to calculate futures prices for two-factor models Input:
    * t - current time step

* T - vector of futures maturities
* vec - vector of the model parameters
* vecSeas - vector of seasonality parameters
* xt - vector of values of underlying stochastic processes
* risk - risk parameters vector
* B - value of the jump integral

Output:

* F - futures price vector

– **seasonality.m**Seasonality function. Input:

* c_1, c_2, c_3, c_4 - seasonality function parameters as in thesis
* t - time

Output:

* seas - value of the seasonality function

– **minimizeMoments.m** Objective function for finding parameters from the moments of data using a covariance matrix. Input:

* M - vector of the moments from the data
* func - function that defines moments for given process
* vec - vector of parameters for the model
* Lambda - covariance matrix for the moments
* xt - initial spot price
* delta - time step
* L - number of the moments for calibration

Output:

* obj - value of the objective function

– **OUgenerator2F.m** Particle generator function for two-factor models Input:

* vec - vector of parameters for the model
* xt - initial spot price vector
* delta - timestep
* N - number of particles

Output:

* X - block matrix of generated particles for 2 factors 2xN

Bibliography

[1] Aihara, S., Bagchi, A., and Imreizeeq, E. (2011). Identification of electricity spot models by using convolution particle filter. *International Journal of Innovative Computing, Information and Control*, 7(1):61–72.

[2] Aiube, F., Baidya, T., and Huarsaya, E. (2008). Analysis of commodity prices with the particle filter. *Energy Economics*, 30(2):597–605.

[3] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.

[4] Albanese, C., Lo, H., and Tompaidis, S. (2012). A numerical algorithm for pricing electricity derivatives for jump-diffusion processes based on continuous time lattices. *European Journal of Operational Research*, 222(2):361 – 368.

[5] Alexander, C. (2004). Normal mixture diffusion with uncertain volatility: Modelling short-and long-term smile effects. *Journal of Banking & Finance*, 28(12):2957–2980.

[6] Arulampalam, M. S., Maskell, S., Gordon, N., and Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *Signal Processing, IEEE Transactions on*, 50(2):174–188.

[7] Babbs, S. H. and Nowman, K. B. (1999). Kalman filtering of generalized vasicek term structure models. *Journal of Financial and Quantitative Analysis*, 34(01):115–130.

[8] Bachelier, L., Davis, M., and Etheridge, A. (2011). *Louis Bachelier's theory of speculation: the origins of modern finance*, volume 13. Princeton University Press.

[9] Bakshi, G., Cao, C., and Chen, Z. (1997). Empirical performance of alternative option pricing models. *The Journal of Finance*, 52(5):2003–2049.

[10] Barndorff-Nielsen, O. and Shephard, N. (2002). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of Royal Statistical Society, Series B*, 63(2):253–280.

[11] Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options. *Review of Financial Studies*, 9(1):69–107.

[12] Black, F. (1976). The pricing of commodity contracts. *Journal of Financial Economics*, 3(1-2):167–179.

[13] Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81:637–654.

[14] Brennan, M. and Schwartz, E. (1985). Evaluating natural resource investments. *Journal of Business*, 58:135–157.

[15] Brigo, D. and Mercurio, F. (2006). *Interest rate models-theory and practice: with smile, inflation and credit*. Springerverlag Berlin Heidelberg.

[16] Brody, D. C., Hughston, L. P., and Mackie, E. (2012). General theory of geometric lévy models for dynamic asset pricing. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 468(2142):1778–1798.

[17] Carr, P., Geman, H., Madan, D. B., and Yor, M. (2003). Stochastic volatility for lévy processes. *Mathematical Finance*, 13(3):345–382.

[18] Cartea, A. and Figueroa, M. G. (2005). Pricing in electricity markets: a mean reverting jump diffusion model with seasonality. *Applied Mathematical Finance*, 12(4):313–335.

[19] Cortazar, G., Mila, C., and Severino, F. (2008). A multicommodity model of futures prices: Using futures prices of one commodity to estimate the stochastic process of another. *The Journal of Futures Markets*, 28(6):537–560.

[20] Cortazar, G. and Schwarz, E. (2003). Implementing a stochastic model for oil futures prices. *Energy Economics*, 25(3):215–238.

[21] Cox, J., Ingersoll Jr, J., and Ross, S. (1985). A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society*, 53(2):385–408.

[22] Date, P., Mamon, R., and Tenyakov, A. (2013). Filtering and forecasting commodity futures prices under an HMM framework. *Energy Economics*, 40:1001–1013.

[23] Date, P. and Ponomareva, K. (2011). Linear and non-linear filtering in mathematical finance: a review. *IMA Journal of Management Mathematics*, 22(3):195–211.

[24] Date, P. and Wang, C. (2009). Linear Gaussian affine term structure models with unobservable factors: Calibration and yield forecasting. *European Journal of Operational Research*, 195(1):156–166.

[25] Deaton, A. and Laroque, G. (1992). On the behaviour of commodity prices. *The Review of Economic Studies*, 59(1):1–23.

[26] Deng, S. (2000). Pricing electricity derivatives under alternative stochastic spot price models. In *System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on*, pages 10–pp. IEEE.

[27] Derman, E. and Kani, I. (1994). Riding on a smile. *Risk*, 7(2):277–284.

[28] Duffie, D. (1996). *Dynamic Asset Pricing Theory*. Princeton University Press.

[29] Duffie, D., Pan, J., and Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376.

[30] Dupire, B. (1994). Pricing with a smile. *Risk*, 7(1):18–20.

[31] Durbin, J., Koopman, S., and Atkinson, A. (2001). *Time series analysis by state space methods*, volume 15. Oxford University Press Oxford.

[32] Escribano Sáez, Á., Pena, J. I., and Villaplana, P. (2002). Modeling electricity prices: international evidence. In *EFMA 2002 London Meetings*.

[33] Eydeland, A. and Geman, H. (1999). *Energy Modelling and the Management of Uncertainty*, chapter Fundamentals of electricity derivatives, pages 35–43. Risk Publications, London, England.

[34] Fama, E. and French, K. (1987). Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage. *Journal of Business*, 60(1):55–73.

[35] Garman, M. B. (1976). A general theory of asset valuation under diffusion state processes. Technical report, University of California at Berkeley.

[36] Gatheral, J. (2006). *The volatility surface: a practitioner's guide*, volume 357. Wiley.

[37] Geman, H. and Roncoroni, A. (2003). A class of marked point processes for modelling electricity prices. Technical report, ESSEC Research Center, ESSEC Business School.

[38] Geman, H. and Roncoroni, A. (2006). Understanding the fine structure of electricity prices. *The Journal of Business*, 79(3):1225–1261.

[39] Gibson, R. and Schwartz, E. (1990). Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance*, 45(3):959–976.

[40] Girsanov, I. (1960). On transforming a certain class of stochastic processes by absolutely continuous substitution of measures. *Theory of Probability & Its Applications*, 5(3):285–301.

[41] Grimmett, G. and Stirzaker, D. (1992). *Probability and random processes*, volume 2. Oxford University Press.

[42] Harvey, A. (1991). *Forecasting, structural time series models and the Kalman filter*. Cambridge Univ Pr.

[43] Heath, D., Jarrow, R., and Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica: Journal of the Econometric Society*, 60(1):77–105.

[44] Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343.

[45] Higgs, H. and Worthington, A. C. (2005). Systematic features of high-frequency volatility in australian electricity markets: Intraday patterns, information arrival and calendar effects. *The Quarterly Journal of the IAEE's Energy Economics Education Foundation*, 26(4):23–42.

[46] Hotelling, H. (1931). The economics of exhaustible resources. *The Journal of Political Economy*, 39(2):137–175.

[47] Hubbert, M. K. (1956). *Nuclear energy and the fossil fuels*, volume 95. Shell Development Company, Exploration and Production Research Division Houston, TX.

[48] Huisman, R. and Mahieu, R. (2003). Regime jumps in electricity prices. *Energy Economics*, 25(5):425–434.

[49] Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42(2):281–300.

[50] Hyndman, C. (2007). Gaussian factor models - futures and forward prices. *IMA Journal of Management Mathematics*, 18(4):353–369.

[51] Javaheri, A. (2011). *Inside volatility arbitrage: the secrets of skewness*, volume 317. Wiley.

[52] Kalman, R. et al. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1):35–45.

[53] Kou, S. G. (2002). A jump-diffusion model for option pricing. *Management Science*, 48(8):1086–1101.

[54] Lautier, D. and Galli, A. (2004). Simple and extended Kalman filters: an application to term structures of commodity prices. *Applied Financial Economics*, 14(13):963–973.

[55] Mackey, M. (1989). Commodity price fluctuations: price dependent delays and nonlinearities as explanatory factors. *Journal of Economic Theory*, 48(2):497–509.

[56] Manoliu, M. and Tompaidis, S. (2002). Energy futures prices: term structure models with Kalman filter estimation. *Applied Mathematical Finance*, 9(1):21–43.

[57] Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

[58] Markowitz, H. (1959). *Portfolio selection: efficient diversification of investments*. Yale University Press.

[59] Marx, K. (1987). A contribution to the critique of political economy. In *Collected Works of Karl Marx and Frederick Engels*, volume 29. International Publishers: New York.

[60] Merton, R. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1):125–144.

[61] Mikhailov, S. and Nögel, U. (2004). Heston's stochastic volatility model implementation, calibration and some extensions. *The Best of Wilmott 1: Incorporating the Quantitative Finance Review, Volume 1*, Wilmott magazine:401–411.

[62] Mirantes, A., Poblacion, J., and Serna, G. (2012). The stochastic seasonal behaviour of natural gas prices. *European Financial Management*, 18(3):410–443.

[63] Monoyios, M. (2007). Optimal hedging and parameter uncertainty. *IMA Journal of Management Mathematics*, 18(4):331–351.

[64] Mst, D. and Keles, D. (2010). A survey of stochastic modelling approaches for liberalised electricity markets. *European Journal of Operational Research*, 207(2):543 – 556.

[65] Nelson, C. R. and Siegel, A. F. (1987). Parsimonious modeling of yield curves. *Journal of Business*, 60(4):473–489.

[66] Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics*, 63(1):3–50.

[67] Pascucci, A. (2011). *PDE and martingale methods in option pricing*, volume 2. Springer.

[68] Reuters (2009). Us henry hub average spot gas price sinks in 2009. article available at http://www.reuters.com/article/2009/12/30/energy-natgas-hubprice-idUSN3019624520091230.

[69] Ricardo, D. (1971). *Principles of Political Economy and Taxation*. Pelican Books, London.

[70] Samuelson, P. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, 6(2):41–49.

[71] Schwartz, E. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *Journal of Finance New-York*, 52(3):923–974.

[72] Smith, A. (1970). *Wealth of Nations*. Pelican Books: London.

[73] Sørensen, C. (2002). Modeling seasonality in agricultural commodity futures. *Journal of Futures Markets*, 22(5):393–426.

[74] Stein, E. and Stein, J. (1991). Stock price distributions with stochastic volatility: an analytic approach. *Review of Financial Studies*, 4(4):727–752.

[75] Swider, D. J. and Weber, C. (2007). Extended arma models for estimating price developments on day-ahead electricity markets. *Electric Power Systems Research*, 77(5):583–593.

[76] Tankov, P. (2004). *Financial modelling with jump processes*. CRC Press.

[77] Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177–188.

[78] Villaplana, P. (2003). Pricing power derivatives: A two-factor jump-diffusion approach. *Working Paper. Bussiness Economics Series 05, Departamento de Economiz de la Empresa, Universidad Carlos III de Mardid*, pages 03–18.

[79] Weron, R., Simonsen, I., and Wilman, P. (2004). Modeling highly volatile and seasonal markets: evidence from the nord pool electricity market. In *The application of econophysics*, pages 182–191. Springer.

[80] Wiggins, J. (1987). Option values under stochastic volatility: Theory and empirical estimates. *Journal of Financial Economics*, 19(2):351–372.

[81] Working, H. (1949). The theory of price of storage. *The American Economic Review*, 39(6):1254–1262.