Stochastic Models for Energy Markets

Statistics, Pricing and Model Risk

Dissertation

zur Erlangung des Doktorgrades
Dr. rer. nat.

der Fakultät für Mathematik
der Universität Duisburg-Essen

vorgelegt von

Anna Nazarova, M.Sc.

Essen
Februar 2014

Amtierender Dekan: Prof. Dr. Ulrich Görtz

1. Gutachter: Prof. Dr. Rüdiger Kiesel
2. Gutachter: Prof. Dr. Fred Espen Benth
Datum der Disputation: 13.05.2014

Ich versichere an Eides statt durch meine Unterschrift, dass ich die vorstehende Arbeit selbstständig und ohne fremde Hilfe angefertigt und alle Stellen, die ich wörtlich oder inhaltlich aus Veröffentlichungen entnommen habe, als solche kenntlich gemacht habe, mich auch keiner anderen als der angegebenen Literatur oder sonstiger Hilfsmittel bedient habe. Die Arbeit hat in dieser oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegen. Es hat kein vorangegangenes Promotionsverfahren gegeben.

Anna Nazarova
Essen, Februar 2014

Summary

In this thesis we develop new methods and procedures to complement and improve current modelling frameworks and to provide a deeper and better understanding of energy markets. We investigate various aspects of stochastic modelling of energy markets: we analyse statistical properties of power markets, study pricing methods for different financial energy-related instruments, design a new storage model and examine model risk.

In doing so we apply a wide range of methods from different branches of applied mathematics ranging from statistical and econometric techniques to a partial differential equations based approach and algorithms from numerical analysis. We modify and extend these methods to make them applicable to our problem setting.

The study reveals results of both theoretical and practical importance. In particular, there are the main findings of this thesis:

- A critical comparison of the properties and estimation procedures of three recently proposed and widely used stochastic power price models shows that none of the models outperforms each other, as all of them have some drawbacks. The more important issue when modelling power prices is that it is more efficient to use additive models (due to their analytical tractability) which present a power price as a sum of various stochastic process responsible for different price fluctuation magnitudes and mean-reversion forces.
- An integro-partial differential equation (integro-PDE) based method is implemented to find the power forward price dynamics for a regime-switching power price model which is a critical issue for hedging purposes.
- A new approach to storage value modelling is developed to complement current stochastic optimal control methods on finding an optimal storage policy. The main novelty is that the storage level process is represented as a bounded diffusion for which we are able to derive the transition probability density formula which in turn allows for a great variability of further applications to pricing and value storage.
- A detailed investigation of various sources of risks when modelling power price is applied to the example of a gas-fired power plant and finds that spike risk is by far the most important source of model risk.

Zusammenfassung

In dieser Arbeit entwickeln wir neue Methoden und Verfahren, um aktuelle Modellierungsverfahren zu ergänzen und zu verbessern und um so ein tieferes Verständnis von Energiemärkten zu gewinnen. Wir untersuchen verschiedene Aspekte der stochastischen Modellierung von Energiemärkten: wir analysieren stochastische Eigenschaften von Elektrizitätsmärkten, betrachten Bewertungsmethoden für verschiedene Energie-verwandte Finanzinstrumente, entwerfen ein neues Speichermodell und untersuchen Modellrisiko.

Im Zuge dessen wenden wir Methoden der verschiedenen Bereiche der angewandten Mathematik, von statistischen und ökonometrischen Techniken über einen auf partielle Differenzialgleichung basierenden Ansatz bis hin zu Algorithmen der numerischen Analysis, an. Wir modifizieren und erweitern diese Methoden, um sie auf unsere Problemstellung anwenden zu können.

Die Resultate der Arbeit sind theoretischer und praktischer Natur. Folgende Ergebnisse der Arbeit seien besonders hervorgehoben:

- Ein kritischer Vergleich der Eigenschaften und der Schätzverfahren von drei kürzlich veröffentlichten und weitverbreiteten stochastischen Elektrizitätspreismodellen zeigt, dass keines der Modelle eines der anderen übertrifft. Wichtiger bei der Modellierung von Elektrizitätspreisen ist, dass additive Modelle aufgrund ihrer analytischen Lenkbarkeit effizienter sind und die Elektrizitätspreise als Summe verschiedener stochastischer Prozesse verantwortlich für unterschiedliche Preisschwankungsausschläge und mean-reversion Kräfte darstellen.

- Eine auf integro-partielle Differenzialgleichung (integro-PDE) basierende Methode wird implementiert, um die Dynamiken des Elektrizitätsforwardpreises für ein regime-switching Elektrizitätspreismodell zu finden, die für Hedging von grundlegender Bedeutung sind.

- Ein neuer Ansatz der Speicherbewertung wird entwickelt, um aktuelle stochastische Methoden der optimalen Steuerung zu ergänzen und eine optimale Speichersteuerung zu finden. Die Hauptneuheit ist der Speicherstandsprozess, der als beschränkte Diffusion dargestellt wird. Hierzu können wir Formeln für die Übergangswahrscheinlichkeitsdichten herleiten, die eine große Variabilität weiterer Anwendungen in der Bepreisung und Speicherbewertung erlauben.

- Wir wenden eine detaillierte Untersuchung der verschiedenen Risikoquellen bei der Elektrizitätspreismodellierung auf das Beispiel eines Gaskraftwerkes an and finden, dass das Risiko von Preisspikes bei weitem die wichtigste Quelle des Modellrisikos ist.

Acknowledgements

The most rewarding achievement in my life, as I am approaching my thirtieth birthday, is the completion of my doctoral dissertation. During this period I am tremendously lucky to have the opportunity to work with my supervisors Prof. Rüdiger Kiesel and Prof. Fred Espen Benth.

I would like to express my immense gratitude to my first supervisor Prof. Rüdiger Kiesel who was very generous with his invaluable support and helpful advice. I am greatly thankful to him for his careful guidance, constant encouragement of my endeavours and countless edits of my thesis. I am profoundly indebted to my second supervisor Prof. Fred Espen Benth who was a great inspiration to me with his brilliant ideas throughout the research period both in person and on the phone.

I would also like to thank Prof. Nick Bingham for his vital help with corrections of the papers I have been involved in writing. I am very grateful to Prof. Mikhail Urusov for his time and numerous fruitful discussions which were crucial to my investigation.

My further appreciation and thanks goes to my colleagues Sascha Kollenberg for his beneficial remarks and corrections of Chapter 4 and Andrea von Avenarius for her kind help with translating the abstract. I would also like to thank the secretary of our chair Kirsten van der Koelen for her willingness to help with all the organisational matters.

Special thanks is reserved for my parents who were always supportive in letting me be who I have become and continue to progress. In addition, I am very thankful to my best friend Dr. Anton Reshetnikov for his endless patience and plentiful reliance during these times.

CONTENTS

List of Theorems, Propositions and Definitions

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Energy markets

Energy markets are classified as commodity markets that operate with various energy-related products. These commodity products are far different from the classical financial market products. The reason for the difference is the following: in addition to a rich spectrum of factors, there are some exceptional aspects that have a great impact on the energy markets formation and activity. Among these are mainly a stable growth of worldwide energy demand [International Energy Agency, 2013], a global climate policy, regional weather conditions, local balancing in supply and demand and, finally, storing and shipping difficulties.

To give an overview of energy-related commodities we extend a classification given in Eydeland & Wolyniec [2003]:

- Fuels: oil, gas, coal, and their derivatives and byproducts;
- Electricity (power);
- Weather, emissions, pulp and paper, and forced outage insurance;
- Renewables: solar, wind, rain, tides, waves, biomass and geothermal heat;

All these products possess original and unique characteristics that make them quite challenging to model and price. We list some of the most pronounced:

- nonstorability (power);
- limited predictability (weather, wind, rain);
- high costs (biomass);
- scarcity (fossil fuels);

- seasonality;
- mean-reversion;
- regime switching;
- extraordinary jumps (power);
- high volatility;
- state dependent regulatory constraints (emissions).

As energy markets have recently been deregulated the need for a thorough and careful analysis of their features remains. In particular, as they are quite young, fast growing and regularly continue to change and modify, it is very important to proceed with exploring these markets and providing a deeper insight into the variety of related problems.

1.2 Electricity as a principal component of energy markets

Among all energy-related products, electricity is the principal and most significant one since it is directly related to any other product and can be viewed as an outcome of a mixture of them. As stated in Kaminski [2005], this commodity possess some specific properties like extremely high volatility, mean-reversion, dependence volatility of the price level, strong seasonal behaviour, the tendency of prices to jump upwards and downwards, and significant differences in the behaviour of prices across different geographical markets. They explain this by various demand patterns, differences in level and structure of installed generation capacity, differences in the developments of the transmission network and power pool design. Eydeland & Wolyniec [2003] claim that the most influencing property of power markets is the need for real-time balancing of supply and demand provided that electricity cannot be stored. The first chapter in Kaminski [2005] emphasises that nonstorability is crucial, since it regulates the delivery at several points in time over the period. However, inventories cannot be used to smooth price fluctuations over time. This implies that the classical approach to storable commodity markets modelling with the cost-of-carry relationship cannot be helpful. Instead, one has to focus on capturing the stochastic behaviour of power prices.

Current literature mainly offers three groups of modelling approaches: pure stochastic models, hybrid models and equilibrium models. Equilibrium models focus on modelling supply and demand relationship and deriving a power price as a result of optimisation over a range of production constraints. Supatgiat *et al.* [2001] suggest an equilibrium model where the market clearing price is a result of Nash equilibrium. Further, Bessembinder & Lemmon [2002] present an equilibrium model for power forward prices and studied the equilibrium forward premium.

The main drawback of this approach is that it does not allow for a dynamical representation of power prices and, therefore, limits generating hedging strategies.

Hybrid models concentrate on the benefits of pure stochastic and equilibrium approaches. Pirrong & Jermakyan [1999] and Eydeland & Geman [1999] offer various underlying drivers when modelling power prices via transformation of the fundamental factors. Further, Barlow [2002] proposes a model with only one demand factor and a straightforward transformation of it to obtain a diffusion model for spot prices which can exhibit price spikes. These models include more information about the fundamental price drivers. The possible drawback is that it might require non-trivial estimation algorithms to fit the model to the real data. Another issue would be possibly difficult to capture non-linear relationship between the drivers and the price.

Finally, pure stochastic models solely centre on power price itself. The main goal here is to capture possibly all the structural properties of observed power prices. Among the first papers are Schwartz & Smith [2000] and Lucia & Schwartz [2002], where the authors offer a two-factor model capturing short-term mean-reversion effect and investigate its properties. Later Cartea & Figueroa [2005] extend it by adding a jump component, which is a critical point in power markets that exhibit extremely large jumps due to nonstorability. Geman & Roncoroni [2006] continue with a modification of this model by adding a state-dependent function in front of the jump component to allow the price to jump downwards in the case of relatively high level of power spot prices. These models have a limited ability in analytical derivation of the electricity derivatives. Next Benth *et al.* [2007] suggest to model a power price dynamics as a sum of non-Gaussian Ornstein–Uhlenbeck processes to capture the stylised features. One of the main advantages of this model is that it allows for analytical expressions for forwards and options. However, the main drawback of this pure stochastic approach is that it requires advanced techniques to estimate the model parameters and which entails the model risk.

1.3 Contribution of the thesis

In this thesis we explore various aspects of stochastic power price modelling. We attempt to fill some of the missing gaps in this research area by introducing new ideas on model risk and storage and extending previous analysis on model performance and pricing.

The starting point of our investigation is a critical analysis of widely used stochastic models for electricity spot price process. The first model, called the threshold model, is developed by Geman & Roncoroni [2006], and is an exponential Ornstein–Uhlenbeck process driven by a Brownian motion and a state-dependent compound Poisson process. It is designed to capture both statistical and pathwise properties of electricity spot prices. The second model, called the factor model, was proposed by Benth *et al.* [2007]. It is an additive linear model, where the price dynamics is a superposition of Ornstein–Uhlenbeck processes driven by subordinators to

ensure positivity of the prices. It separates the modelling of spikes and base components. The third model, called the jump-diffusion model, is proposed by Cartea & Figueroa [2005], and is a one-factor mean-reversion jump-diffusion model, adjusted to incorporate the most important characteristics of electricity prices. We calibrate all three models to German spot price data. We critically compare the properties and the estimation procedures of three models and discuss several shortcomings and possible improvements. Besides analysing the spot price behaviour, we compute forward prices (analytically for the jump-diffusion and the factor models and numerically for the threshold model) and risk premia for all three models for various German forward data and identify the key forward price drivers.

As the threshold model does not allow for an analytical expression for forward prices, the only way so far was to compute these via Monte Carlo technique. Albanese *et al.* [2008] also offer a numerical method for pricing derivatives on electricity prices within this model. The method is based on approximating the generator of the underlying process. Despite the fact that the authors claim that the method is accurate even in the case of processes with fast mean-reversion and jumps of large magnitude, it takes quite significant computational efforts to obtain desired results. In contrast to it, we suggest a new approach to this: basing on the generator of the process we obtain a partial-differential equation with an integral term (due to the jump component). This term is quite similar to PDE for the jump-diffusion models with one small but crucial difference: the state-dependent function h which takes either $+1$ or -1 values. We numerically solve this PDE with a finite difference method and obtain forward prices dynamics, which are very important in terms of hedging and managing the risks. When solving this equation numerically, we first have to cut the domain for x such that $x \in (x_{min}, x_{max})$. At second we have to truncate the integral term from $(-\infty, \infty)$ to (K_{min}, K_{max}). We are able to estimate these truncation error analytically. We study the influence of the function h on the forward dynamics. For a fair comparison, we benchmark our PDE-driven forward dynamics to the forward dynamics obtained for the jump-diffusion model, which is possible to do for some classes of distributions assumed for the jump size.

We also analyse a storage valuation problem. Power prices are closely related to storage for a number of reasons. Firstly, the share of hydro-driven generated power is increasing in many Scandinavian and European countries in terms of a general growth of renewable energy. Secondly, both fossil(gas)-driven power plant producers and/or storage owners face a problem of optimal managing the storage reservoir. Therefore, due to random nature of the prices (gas, power), optimal managing means searching for an optimal (in the sense of value maximisation) policy to inject (gas), to withdraw (gas or hydro) or to "do nothing" (if market prices are not attractive) over a set of constraints (managing costs). This type of problem belongs to a class of stochastic optimal control problems and is a main focus of present literature (Ahn *et al.* [2002], Chen & Forsyth [2007], Kjaer & Ronn [2008], Thompson *et al.* [2009], Carmona & Ludkovski [2010]) on storage value optimisation. We instead look at the problem differently and propose a new approach to storage problem which directly addresses to value modelling. The novelty

lies in the representation of the storage process as a bounded diffusion in the finite interval (l, u). Some specifications allow for an analytical formula for the transition probability density and, as a consequence, straightforward calculation of the storage value via various payoffs. The main benefit of this method is that it does not involve sophisticated algorithms of searching an optimal policy and that it can easily be implemented.

Finally, as was mentioned in the previous section, power price stochastic models are not easy to calibrate with the historical market prices. And even if there are some methods and procedures available around, the natural question is how valid and reliable our estimation is. In contrast to financial markets, energy model risk investigation area has not been discussed in literature at all to the best of our knowledge. We cover this topic which is of a great interest to practitioners. We assess the model risk inherent in the valuation procedure of fossil-driven power plants. To capture model risk we use risk-capturing functionals, a methodology recently established in a series of papers like Cont [2006] and Bannör & Scherer [2013]. As gas-fired power plants are seen as flexible and low-carbon sources of electricity which are important building blocks in terms of the switch to a low-carbon energy generation, we consider the model risk in this asset class in detail. Our findings reveal that spike risk is by far the most important source of model risk.

The thesis is based on four papers. The first one Benth *et al.* [2012] is already published paper. The second one Bannör *et al.* [2014] is submitted to Energy Economics. The last two are working papers and are in preparation to be submitted.

1.4 Structure of the thesis

The thesis is organised as follows. Every chapter starts with introduction and motivation sections, then proceeds with the research flow and concludes with discussion and perspective outlook. Since the thesis covers many topics from various areas of Probability and Stochastic Processes Theory, Financial Mathematics, Lévy processes and Statistics, additional ”background” chapter would be excessive. Instead, some of the useful definitions, facts and theorems are given in Appendices.

Chapter 2 begins with introduction of electricity price models and overview of used algorithms to calibrate the models to the historical data. Then it goes on with the critical comparison of the obtained results. Further for all three models assess their ability to pricing derivatives by computing forward prices and market risk premia.

Chapter 3 continues with pricing forwards prices. It firstly briefly introduces the models and then continues with numerical implementation of finite difference method. It also reports the estimates for the domain truncation error and for the integral truncation error. This chapter finalises with discussion of the resulting forward prices for two models.

Chapter 4 is focused on a storage value problem. It first of all presents theoretical development of the bounded stochastic processes. Then it continues with its application to storage modelling and explores various hydro- and gas-driven payoffs to construct the storage value. It also shows numerical examples to illustrate the resulting values. And eventually it discusses further application and extension possibilities.

Chapter 5 starts with general overview of the model risk and financial instruments we use to do so. It further continues with modelling assumptions and estimation algorithms. Next this chapter demonstrates the resulting model risk values for various risk sources. It then concludes and offers some outlook.

Finally, Chapter 6 concludes and provides some prospectives for future research.

CHAPTER 2

ELECTRICITY MODELLING

2.1 Introduction

The modelling of the dynamics of electricity spot prices is a delicate issue. The spot prices exhibit various characteristics (see Eydeland & Wolyniec [2003]): seasonality, spikes and mean-reversion. Depending on the market, we can observe daily, weekly, monthly or yearly seasonality. Sudden big changes in price or so-called spikes can be caused for example by unexpected weather change or outage of equipment. The intensity of spikes may also demonstrate both time dependency and randomness. Also, prices are mean-reverting at different speeds. As the deregulated electricity markets are still developing and fast growing, practitioners as well as academics have suggested several models to capture some or all of these features. Recently, three models have attracted considerable attention: a model proposed by Cartea & Figueroa [2005] (called the *jump-diffusion* model), a model proposed by Roncoroni [2002] and further developed by Geman & Roncoroni [2006] (called the *threshold* model), and a model derived by Benth *et al.* [2007] (called the *factor* model). We will provide a detailed comparison of the three models in terms of their empirical ability to fit spot price data and to price forwards. In order to do so we use data from the German Electricity Exchange, EEX, and investigate the performance of the models.

The jump-diffusion model can be seen as a one-factor mean-reverting jump-diffusion model close to the classical exponential Ornstein-Uhlenbeck process suggested by Schwartz [1997] and later applied to electricity markets by Lucia & Schwartz [2002]. As the two latter models do not incorporate jumps, the model proposed by Cartea & Figueroa [2005] is extended to account for jumps. The model is easy to calibrate and produces a straightforward formula to price forward contracts. Due to its simple and parsimonious structure the jump-diffusion model is quite extensively used among practitioners.

The threshold model can be seen as a one-factor mean-reverting jump-diffusion model close to the model of Cartea & Figueroa [2005] with two novel twists. Firstly, the authors introduce a state-dependent *sign* of the jump component, where high price levels induce negative jumps,

whereas in low price regimes the jumps are upwards. This feature prevents several spikes following each other. Secondly, the estimation process makes use of a threshold, which is set iteratively, so that the estimated parameters are calibrated to the empirical kurtosis. In order to price forwards with this model numerically, typically Monte Carlo, techniques have to be used. Albanese *et al.* [2008] propose a method based on approximating the generator of the underlying process and illustrate the speed and accuracy of the method by pricing European and Bermudan options. A lattice-based method for the discretisation of the threshold model that allows for the pricing of derivatives, including swing options, has been proposed in Geman & Kourouvakalis [2008].

The factor model is an additive (or linear) multi-factor model that separates the base and spike signals. By the base signal we mean the daily fluctuations of the price around the mean level due to small changes in supply and demand in the market; by the spike signal we mean the price jump of extreme size due to sudden imbalances of demand and supply. This structure allows for more flexibility in capturing the high speed of mean-reversion observed for spikes and the more slowly varying base signal. However, the estimation of the parameters in this model is challenging. Applying techniques from Meyer-Brandis & Tankov [2008] together with the prediction-based estimating functions technique of Sørensen [2000], we suggest an improved estimation procedure. Due to its specification, the model provides a simple and straightforward way to price forwards and options, see Benth *et al.* [2007] and for the pricing of spark spread and average options consult Benth & Kufakunesu [2009].

Clearly, a reliable and well-understood spot price model is important for risk management and pricing purposes. With our study we will provide a comparison of the data-fitting ability and pricing performance of the models. We discuss and clarify estimation procedures for the models. In particular, we observe that in the jump-diffusion and threshold models the mean-reversion is an average of the reversion of spikes and intra-spike behaviour, a property that has earlier been observed in jump-diffusion models of this kind. Also we demonstrate that parameter estimates in the threshold model are very sensitive to changes in the spike sizes. Furthermore, the state-dependent sign change seems to be of little importance for the German market data considered. However, it might be an issue for other data sets. We also find the ability of the model to capture the spot price risk questionable.

The factor model, on the other hand, achieves a more reasonable modelling of the spike behaviour. However, the base (or intra-spike) dynamics seems to be too regular in the sense that it produces a less volatile base signal compared to the data. This may be attributed to the use of subordinators, which are processes that in the most tractable cases are of finite activity.

The three models have been applied in various contexts for pricing purposes. The performance of the three models leads to the conclusion that they all require careful refinement in both specification and estimation in order to fully capture the stylised statistical and pathwise properties of electricity spot and forward price data.

We present our findings as follows. In the next three sections we introduce the models and discuss their theoretical properties. Further, in Section 2.5, we provide algorithms and estimate the parameters of the models using German electricity spot price data. In Section 2.6 we assess the calibration and discuss various possible improvements for modelling noise and jump size distribution. Section 2.7 is devoted to the computation of forward prices and study of the market risk premium. The final Section 2.8 in this chapter concludes.

2.2 The jump-diffusion model

Let $(\Omega, \mathbb{P}, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]})$ be a complete filtered probability space with $T < \infty$ a fixed time horizon. We denote the electricity spot price at time $0 \leq t \leq T$ by $S(t)$, and assume that it takes the form

$$S(t) = e^{\mu(t)} X(t)\,, \tag{2.1}$$

where $\mu(t)$ is a deterministic function modelling the seasonal trend, or mean variations, of the price evolution, and $X(t)$ is some stochastic process modelling the random fluctuations around this trend. In the three models, $X(t)$ will take very different forms, but the trend function $\mu(t)$ will stay the same.

In electricity markets spot prices may demonstrate various types of seasonality: daily, weekly, monthly, yearly or a combination of them. Cartea & Figueroa [2005] study historical spot data from England and Wales and suggest some intra-week seasonality, since the returns show correlation every 7 days. To explain this fact, recall that electricity is traded 7 days per week and the information contained in Friday prices has an impact on the Saturday, Sunday and Monday morning prices. In addition to this weekly seasonality effect, spot prices may vary by seasons, caused by changing supply and demand. In this case, we have to employ some periodic function to capture such a trend behaviour. For instance, in the Nord Pool market there is no inflow of water in the hydro reservoirs in the winter, at the same time the demand is high because of low temperatures, so the winter prices are higher than in summer. To model seasonalities, we follow Geman & Roncoroni [2006], who analyse three of the major U.S. power markets. The market conjuncture reveals yearly seasonality on a monthly basis and a combination of an affine function and two sine functions with a 12 and a 6 month period respectively is used to model this seasonality. We find that in the German EEX market this is also a reasonable choice. We therefore choose the following trend model

$$\mu(t) = \alpha + \beta t + \gamma \cos(\epsilon + 2\pi t) + \delta \cos(\zeta + 4\pi t)\,. \tag{2.2}$$

Here, the parameters $\alpha, \beta, \gamma, \delta$, ϵ and ζ are all constants. The first term is interpreted as fixed cost linked to the power production, while the second term drives the long-run linear trend in the total production cost. The remaining terms give periodicity by adding two maxima per year with

possibly different magnitude. The parameters can be estimated by fitting the trend function, for example by least squares. We continue by specifying the $X(t)$ process for the jump-diffusion model.

Here the deseasonalized logarithmic spot prices are modelled by

$$\mathrm{d}\ln X(t) = -\alpha \ln X(t)\ \mathrm{d}t + \sigma(t)\ \mathrm{d}W(t) + \ln J\ \mathrm{d}q(t)\,, \tag{2.3}$$

where α is the speed of mean-reversion, W is a Brownian motion, $\sigma(t)$ is a time-dependent volatility, J is a proportional random jump size and dq_t is a Poisson process of intensity l with

$$\mathrm{d}q_t = \begin{cases} 1 & \text{with probability} \quad l\ \mathrm{d}t \\ 0 & \text{with probability} \quad (1-l)\ \mathrm{d}t\,. \end{cases} \tag{2.4}$$

A typical assumption on the jump size distribution is $\ln J \sim \mathcal{N}(\mu_J, \sigma_J^2)$ and $\mathbb{E}[J] = 1$.

2.3 The threshold model

In this section we introduce the threshold model and provide a detailed discussion of its various components. Here the deseasonalized logarithmic spot prices are modelled as

$$\mathrm{d}\ln X(t) = -\theta_1 \ln X(t)\ \mathrm{d}t + \sigma\ \mathrm{d}W(t) + h(\ln(X(t-)))\ \mathrm{d}Q(t)\,, \tag{2.5}$$

where W is a Brownian motion, Q is a time-inhomogeneous compound Poisson process, i.e.

$$Q(t) = \sum_{i=1}^{N(t)} J_i\,. \tag{2.6}$$

$X(t-)$ denotes the left-limit as usual. $N(t)$ is a Poisson process with time-dependent jump intensity and counts the spikes up to time t. $J_1, J_2, \ldots$ model the magnitudes of the spikes and are assumed to be independent and identically distributed random variables. The constants θ_1 and σ are both positive. The function h attains two values, ± 1, indicating the direction of the jump. The Brownian component models the normal random variations of the electricity price around its mean, i.e., the base signal. The discontinuous price spikes are incorporated through the jump term $h(\ln X(t-))\,\mathrm{d}Q(t)$. The compound Poisson process Q has a time-dependent jump intensity to account for seasonal variations in the spike occurrence. Note that as in the jump-diffusion model, the threshold model has only one mean-reversion parameter, namely θ_1.

To review the threshold spike modelling approach (along the arguments given in Roncoroni [2002] and Geman & Roncoroni [2006]), we start with the spike intensity. Since spikes show

clustering and periodicity the intensity of $N(t)$, which models the spike intensity, is assumed to be deterministic function

$$\iota(t) = \theta_2 \times s(t)\,. \tag{2.7}$$

Here, θ_2 is interpreted as the expected number of spikes per time unit at a spike-clustering time. The function $s(t)$ represents the normalized and possible periodic jump intensity shape. One reasonable specification of $s(t)$ can be a sine function

$$s(t) = \left[\frac{2}{1 + |\sin[\pi(t-\tau)/k]|} - 1\right]^d, \tag{2.8}$$

where the positive constant k is the multiple of the peaking levels, beginning at time τ. For example, if k is equal to 0.5, then there are 2 peaking times per year, corresponding to two periods with spikes over the year. The exponent d is introduced to adjust the dispersion of jumps around the peaking times. In fact, this parameter is responsible for how short the periods of spike occurrences are. As we shall see in the sequel (see Figure 2.4), the intensity shape function $s(t)$ may exhibit convex or concave peaks with a given periodicity, and the choice of this function is motivated by the shape of the power stack function. We remark in passing that in Benth *et al.* [2007] the same form of intensity function as Equation (2.7) was used in an empirical example for Nord Pool electricity spot prices with $k = 1$. In the Nordic market, spikes occur in the winter period, thus the periodicity is one.

In their paper Geman & Roncoroni [2006] alternatively suggest a stochastic form of the spike intensity to increase the probability of spikes in case when prices are above some specified threshold $\overline{E}(t)$ different from $\mathcal{T}(t)$. The following form is used to capture this effect

$$\iota(t, E(t-)) = \theta_2 \times s(t) \times (1 + \max\{0, E(t-) - \overline{E}(t)\}), \tag{2.9}$$

where $E(t)$ denotes the logarithm of the price. The authors suggest to set this threshold $\overline{E}(t)$ smaller than the threshold $\mathcal{T}(t)$ to define the interval for prices where the spike activity will be higher. As soon as the price level falls below $\overline{E}(t)$, the stochastic intensity reduces to a deterministic intensity.

The spike sizes are modelled by the jump size distribution of the compound Poisson process, that is, by the J_i's. Geman & Roncoroni [2006] propose a truncated exponential distribution for the spike sizes J_i with density

$$p(x; \theta_3, \psi) = \frac{\theta_3 \exp(-\theta_3 x)}{1 - \exp(-\theta_3 \psi)},\ 0 \le x \le \psi\,. \tag{2.10}$$

The average jump size parameter is θ_3, and the maximal possible jump size is ψ. The latter implies an upper bound for the absolute value of price changes. For an empirical analysis of spot

price data series, such an upper bound corresponds to the implicit assumption that there will be no bigger price change in the future than that given by the bound. Consequently, as in Geman & Roncoroni [2006], the model does not generate jumps exceeding historically observed ones. This is a restrictive assumption in the sense that we limit ourselves to include only the observed big changes, which may not be adequate in the future. An alternative to this extreme is to allow potentially unbounded price changes, as proposed in Meyer-Brandis & Tankov [2008], where the authors use a Pareto distribution to model the spike sizes. In this case one can get outliers which will result in huge spikes. From the point of view that in most markets there is a maximum price for the spot which the market cannot exceed in the auction, this approach may also be questioned. However, the technical limit is typically rather high. Furthermore, Meyer-Brandis & Tankov [2008] study electricity spot prices on several European markets and defend their use of this extreme-value distribution based on the empirically observed power-law behaviour in the tails of the daily returns, along with excess kurtosis and positive skew.

Obviously, the chosen jump-size distribution strongly influences the empirical properties of the simulated price paths. Looking at the moments of the price paths resulting from these two specifications, the kurtosis for example will differ dramatically, as we shall see in the next section. It is quite natural that the specification with a truncated exponential jump size distribution has a kurtosis reasonably close to the observed one. For a Pareto specification, we predict future price changes which may be far larger than the historical observed ones, and thus the kurtosis will increase. This indicates that comparing empirical moments in order to assess the quality of a model may be misleading. To look beyond these extremes, we shall rely on an empirical analysis of observed jumps and consider alternative jump size distributions in between the two extreme choices of truncated exponential and Pareto.

In the threshold model the direction of spikes is given by an indicator function h taking values $+1$ and -1 depending on the current spot price level. A threshold is introduced to determine the sign of the spike, denoted by $\mathcal{T}$. Thus,

$$h(\ln X(t)) = \begin{cases} +1, \text{ if } \ln X(t) < \mathcal{T}(t), \\ -1, \text{ if } \ln X(t) \geq \mathcal{T}(t)\,. \end{cases} \tag{2.11}$$

Geman & Roncoroni [2006] introduce the h-function together with the smooth mean-reversion θ_1 to bring the prices to a normal range after being at a high level. They affirm that "a proper choice of the barrier $\mathcal{T}$ coupled with a high jump intensity can generate a sequence of upward jumps leading to high price levels, after which a discontinuous downward move together with the smooth mean-reversion brings prices down to a normal range".

One may believe that h models the mean-reversion of spikes, in the sense that if we first have an upward-pointing spike, the next jump will be pushed down due to the sign of h. However, this is not necessarily the case since it may take some time before the next jump actually occurs.

Indeed, in the high-spike intensity markets, when we first have a jump, at the next step we have either one more jump or not. Of course, if the compound Poisson process decides to jump twice, an upward spike will be followed by a downward jump due to the sign of h and the mean-reversion speed together. Such a high concentration of jumps in a period will lead to a rather strange sequence of up- and downward jumps and concentration of noise. To have a mean-reversion of a spike, the threshold model resorts to the θ_1-parameter. On the other hand, θ_1 also accounts for the mean-reversion of the base signal. A spike requires a fast mean-reversion, whereas the base signal is reverting more slowly. The mean-reversion estimate of θ_1 is higher than expected for a base signal, and somewhat slower than required to dampen a spike. Our empirical analysis of EEX data supports this view. If there are no consecutive jumps, the price path will wiggle around the new level, which has been reached after the first spike, due to the Brownian component and the mean-reversion, unless it is pushed down again. So the parameter h is not responsible for the mean-reversion of the spike process, but θ_1 takes care of this. Therefore, both cases guarantee that consecutive price values will not exceed the threshold level. At this point, we claim that h prevents two consecutive price values above the threshold, i.e. at least one downward price movement lies in between.

2.4 The factor model

Suppose that $X(t)$ is a stochastic process represented as a weighted sum of n independent non-Gaussian Ornstein-Uhlenbeck processes $Y_i(t)$, that is,

$$X(t) = \sum_{i=1}^{n} w_i Y_i(t)\,, \tag{2.12}$$

where each $Y_i(t)$ is defined as

$$\mathrm{d}Y_i(t) = -\lambda_i Y_i(t)\,\mathrm{d}t + \mathrm{d}L_i(t)\,,\ Y_i(0) = y_i, i = 1, \ldots, n\,. \tag{2.13}$$

The weight functions w_i and the mean-reversion coefficients λ_i are positive constants. $L_i(t)$ are assumed to be independent *càdlàg* pure-jump additive processes with increasing paths, i.e. so-called time-inhomogeneous *subordinators* (see Definition 6). Corresponding to each $L_i(t)$ we have a time-inhomogeneous Poisson random measure $N_i(\mathrm{d}t, \mathrm{d}z)$ with a deterministic predictable compensator $\nu_i(\mathrm{d}t, \mathrm{d}z)$. The compensated Poisson random measure is denoted by

$$\widetilde{N}_i(\mathrm{d}s, \mathrm{d}z) = N_i(\mathrm{d}s, \mathrm{d}z) - \nu_i(\mathrm{d}s, \mathrm{d}z)\,.$$

We have the representation

$$L_i(t) = \int_0^t \int_0^\infty z\, N_i(\mathrm{d}s, \mathrm{d}z)\,. \tag{2.14}$$

We shall choose compensator measures of the form $\nu_i(\,\mathrm{d}t,\,\mathrm{d}z) = \,\mathrm{d}t\,\widetilde{\nu}_i(\,\mathrm{d}z)$ for a Lévy measure $\widetilde{\nu}_i(\,\mathrm{d}z)$ (see Definition 5). This implies that $L_i(t)$ is a Lévy process (see Definition 4), or

$$\nu_i(\,\mathrm{d}t,\,\mathrm{d}z) = \rho_i(t)\;\mathrm{d}t\;f_i(\,\mathrm{d}z)\,,$$

where $\rho_i(t)$ is a deterministic function controlling the possibly time-varying jump intensity and $f_i(\,\mathrm{d}z)$ is a jump size distribution. In many markets spikes have a tendency to occur in certain periods of the year, and thus it is natural to let the jump intensity for these vary seasonally. Since the jump process has increasing paths, the jumps are only positive, and thus the Lévy measures $\widetilde{\nu}_i(dz)$ are supported on the positive real line. The spot price will be positive as well, since the processes $Y_i(t)$ will be positive by the definition of these jump processes.

The main idea of the factor model is to decompose the electricity spot price into the base and spike signals. This flexibility allows one to capture mean-reversion at different scales, but at the cost of a quite complicated estimation procedure. An example of a possible model specification using three OU processes is proposed in Benth *et al.* [2007]. The first OU process is assumed to have a stationary Gamma distribution and a constant volatility, responsible for small daily fluctuations around the mean trend. For the second OU process a compound Poisson process is used to capture larger price movements which revert faster to the mean. The third process drives the spikes, and has possibly a seasonally varying jump intensity.

In the additive structure of n OU processes one has n mean-reversion parameters $\lambda_1, \ldots, \lambda_n$. The larger λ_i is, the faster the process $Y_i(t)$ comes back to its mean-level. The autocorrelation function $\rho(k)$ for lag k of $X(t)$ is

$$\rho(k) = \widetilde{w}_1 \mathrm{e}^{-k\lambda_1} + \widetilde{w}_2 \mathrm{e}^{-k\lambda_2} + \cdots + \widetilde{w}_n \mathrm{e}^{-k\lambda_n}\,, \tag{2.15}$$

where $\widetilde{w}_i$ are positive weights summing up to 1. A comparison with the empirical autocorrelation function thus allows one to find the number of factors required and to estimate the mean-reversion from each of the factors. This approach was proposed in Barndorff-Nielsen & Shephard [2001] for their stochastic volatility model, which uses a structure similar to the factor model. A fast reversion in spikes will be observed as a strongly decaying slope in $\rho(h)$, whereas the more slowly reverting base signal is found as slower decaying exponentials.

2.5 Algorithms applied and estimation results

For our empirical analysis we use a data set of the Phelix Base electricity price index at the European Energy Exchange (EEX). The data series range from 13/07/2000 to 7/8/2008, where the weekends are excluded. In total, we have 2099 daily prices constituting the basis for our spot-price estimation. The reason for excluding weekends is mainly because Friday price infor-

mation is contained as a basis for Saturday, Sunday and Monday morning prices and thus the prices over the weekend are not directly comparable to those settled during the week.

2.5.1 Seasonality trend parameters estimation

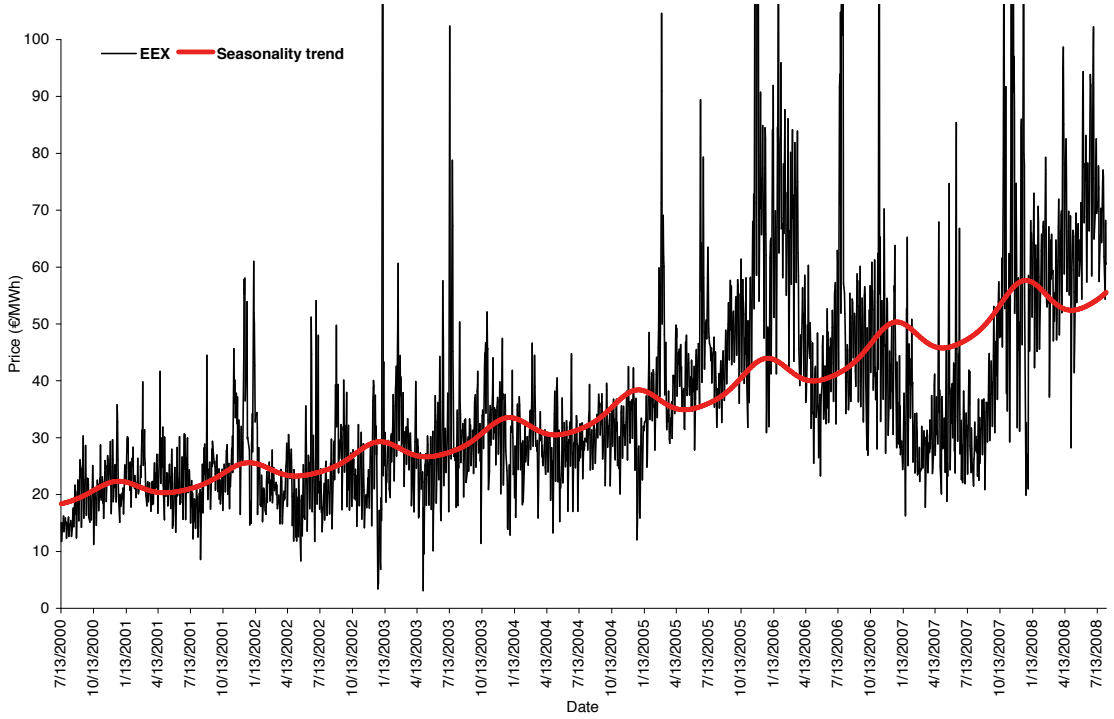


Figure 2.1: EEX price path with seasonality trend.

We start by estimating the seasonality function. The choice of this function is a very important issue, since the specified trend should explain the average market expectation of the price course for the next month, quarter or year. Moreover, deseasonalisation is the first step of the price estimation, so depending on the seasonality estimation, the deseasonalised data set may vary considerably. Cartea & Figueroa [2005] suggest fitting monthly averages of the historical data by a Fourier series of order 5. This is a questionable choice, since this function does not incorporate any trend component, which is necessary to model power dynamics. Also, it is not completely clear how to define the optimal order for a Fourier series. Moreover, the number of parameters for a chosen Fourier series fit with order of 5 is 12, which is twice as many as in the parametric case introduced in Geman & Roncoroni [2006] and Benth *et al.* [2007]. Here the authors use a function $\mu(t)$ described in the previous section. For a deseasonalisation procedure, we take the latter approach for all three models.

In order to adjust for the influence of large price outliers (such as spikes) in the seasonality estimation, we implemented a simple filtering procedure. This compares smoothed (by averaging

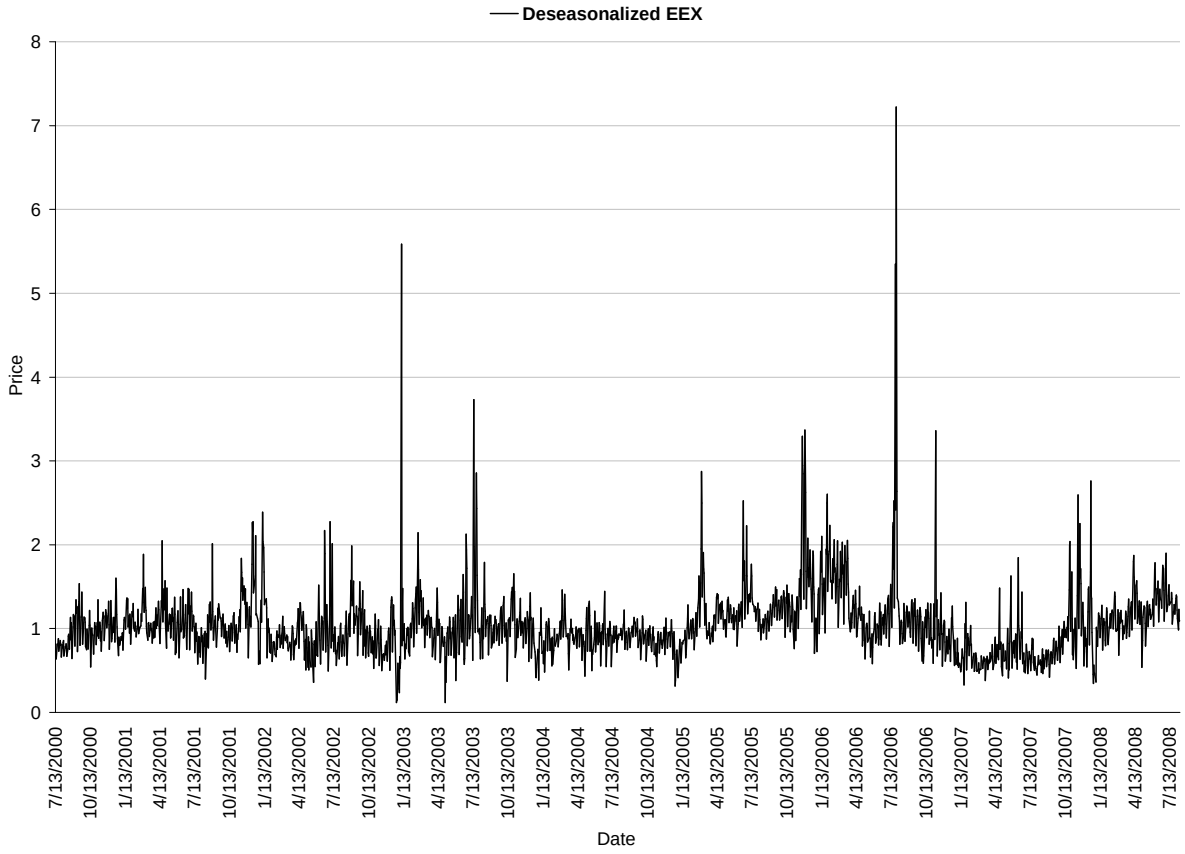


Figure 2.2: Deseasonalized EEX price.

Table 2.1: Seasonality function estimated on filtered data. $R^2 = 0.6961$.

Parameter	Estimate	Std. Error	t-value
α	2.9628	0.0092	320.8417
β	0.1354	0.0020	68.1191
γ	-0.0737	0.0064	-11.4226
δ	0.0117	0.0064	1.8446
ϵ	6.8662	0.0875	78.4585
ζ	0.7464	0.5436	1.3389

over sliding windows of 30 data points length) and observed log prices. If the absolute difference between the smoothed and real log price is bigger than a pre-defined level, here 0.5, we substitute the log price with the smoothed value. In this way we "filter" out the price outliers from the original price data series.

On the filtered data series, we estimate the parameters of the seasonality function $\mu(t)$ specified in Equation (5.5) using non-linear least squares method. The results are reported in Table 2.1. One can see that almost all estimates are significant: α , β , γ , ϵ are significant with $t_{0.99} = 2.326$, δ with $t_{0.95} = 1.645$ and ζ with $t_{0.9} = 1.282$. In Figure 2.1 we plot the estimated seasonality function $\mu(t)$ together with the EEX data. During the period of 2000–2006 the price path and seasonality are matched very well. Due to structural breaks the fit deteriorates between 2006–2008. The price path amplitude has changed in that period compared to previous years. We could capture this by splitting the data set into the two periods and then estimate the seasonality separately. In Figure 2.2 we show the deseasonalized price series.

We continue with the estimation of the three models based on the deseasonalized data, that is, the data obtained after dividing with the estimated $e^{\mu(t)}$-function. For the factor model, we work with these data, whereas for the jump-diffusion and threshold models we use the log-data.

2.5.2 Jump-diffusion model calibration

From the historical spot data we estimate the rolling historical volatility $\sigma(t)$ and its averaged value, the mean-reversion rate α, the frequency l and the standard deviation σ_j of the jumps.

Volatility

Cartea & Figueroa [2005] suggest the volatility to be time-dependent. The motivation is that markets do not show constant volatility, but some volatility structure. However, the authors do not provide any prescribed function or any stochastic alternative for the volatility. Instead, they compute a rolling historical volatility suggested in Eydeland & Wolyniec [2003], which in

Table 2.2: Estimates for the jump-diffusion model parameters.

Parameter	Estimate
α	0.2255
	(0.1938, 0.2584)
	$R^2 = 0.6373$
$[\sigma(t)]$	3.9025
σ_j	1.0996
l	5.67

fact is a deterministic result given a data path. If we make a plot of such a rolling historical volatility, we observe in Figure 2.3 that there appears some seasonal pattern as well as some stochastic element. Obviously, they both need to be incorporated into the model. Otherwise, such a "substitution" by historical volatility may result in unreasonable estimates.

Mean-reversion rate

To get a daily estimate for the mean-reversion rate α, Cartea & Figueroa [2005] suggest using linear regression. The idea of such approach is to rewrite the mean-reversion jump-diffusion process in the discrete version and represent the log price as

$$x_t = a_t + bx_{t-1} + c_t, \tag{2.16}$$

where a_t represents a function of $\mu(t)$, $b \equiv e^{-\alpha}$, c_t is the integral of the Brownian motion and the jump component between times $t-1$ and t.

Jump parameters

To estimate jump parameters, we need to identify jumps from the data. We use a simple technique based on the standard deviation of the returns (see Cartea & Figueroa [2005]). The iterative procedure filters out returns with absolute values greater than three times the standard deviation of the returns of the series at the current iteration. The process is repeated until no further outliers can be found. As a result we obtain a standard deviation of the jumps, σ_j, and a cumulative frequency of jumps, l. The latter is defined as the total number of filtered jumps divided by the annualised number of observations. We report the results in Table 2.2, where annualised estimates and average (denoted by $[\cdot]$) volatility are given appropriately with the 95% confidence interval values in the parenthesis.

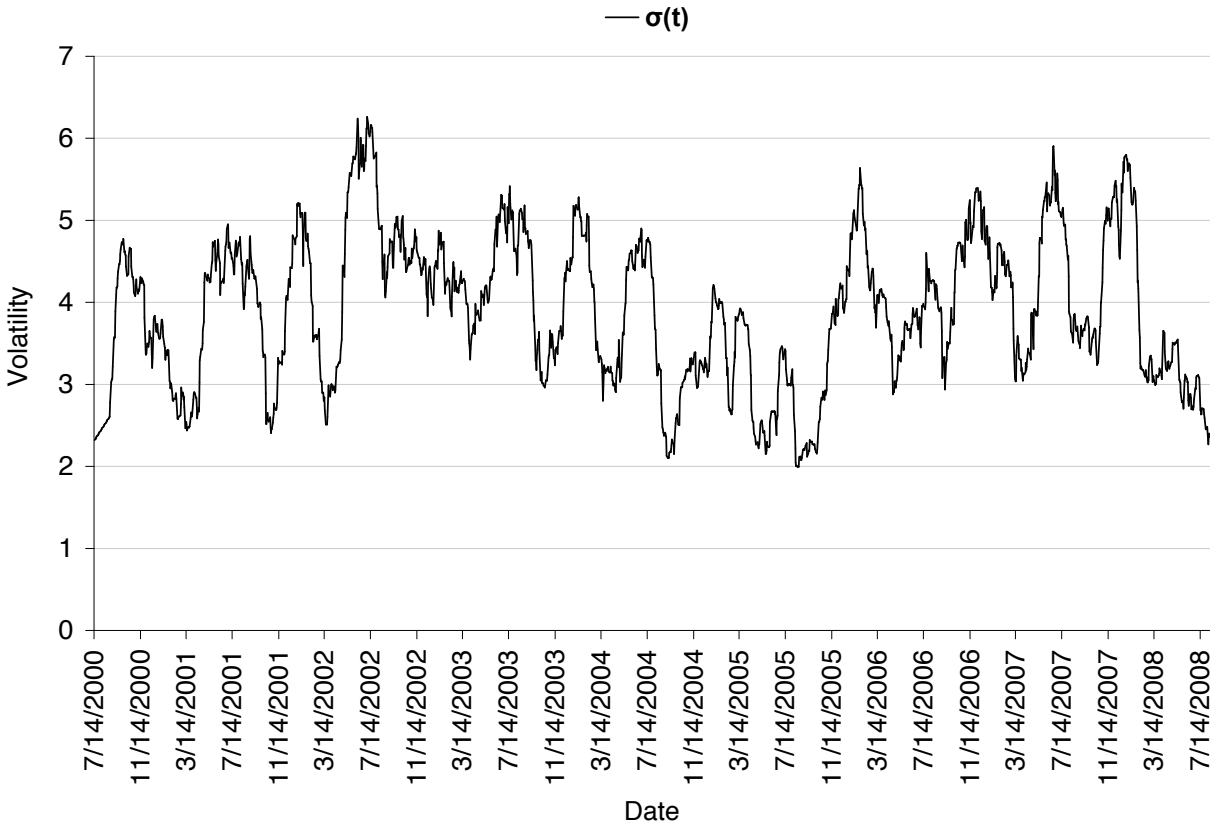


Figure 2.3: Time-dependent volatility for the jump-diffusion model.

2.5.3 Threshold model calibration

Geman & Roncoroni [2006] propose to split the calibration procedure in two steps. First, the so-called *structural elements* are to be estimated, then in the second step one estimates the *model parameters*. The structural elements are the spread Δ, the components of the intensity shape $\iota(t)$ and the maximum jump size ψ. The remaining are the model parameters.

Estimation of structural elements

First, one needs to find the spread Δ, which will limit the jump size of the model. The choice of Δ is a result of a balance between two competing effects: the larger Δ, the higher are the price levels which can be reached during the pressure period and the fewer spikes will occur; the smaller Δ, the sooner the downward jump effect toward normal levels takes place and more spikes will occur. Following Geman & Roncoroni [2006], we select Δ in such a way that the corresponding calibrated model generates paths whose average maximum values are equivalent to those observed in the market. In the case of the EEX data on log-scale, this results in a Δ spread given by 0.7.

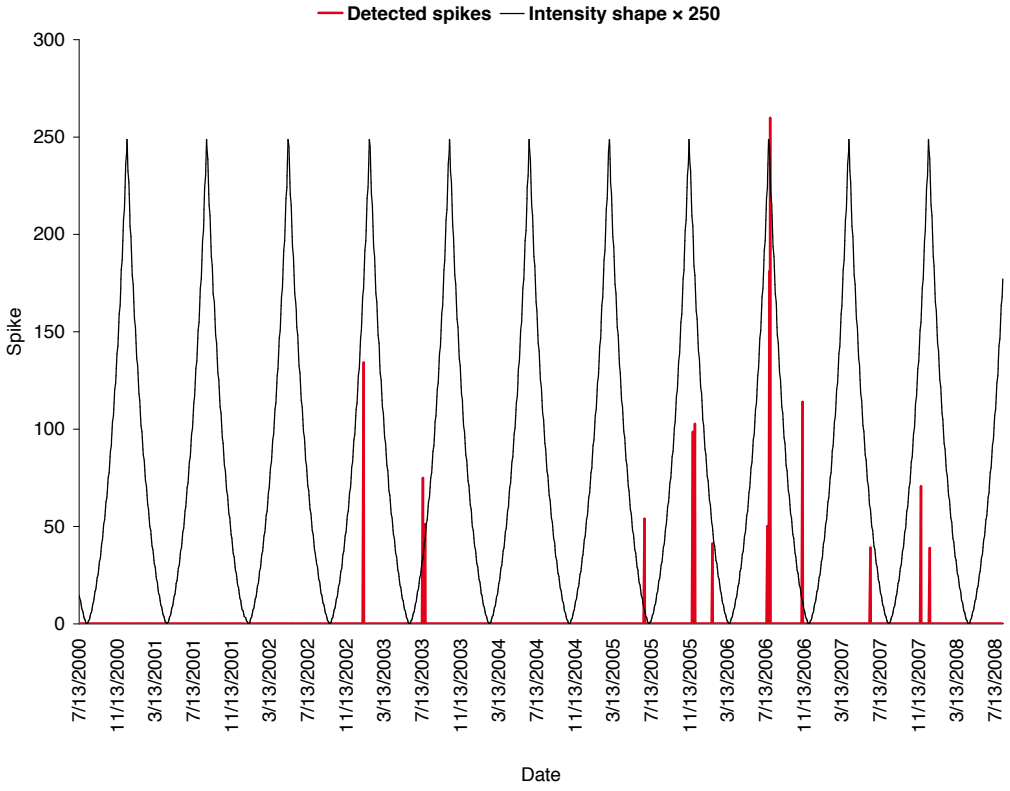


Figure 2.4: Calibrated intensity shape function and real values of detected spikes.

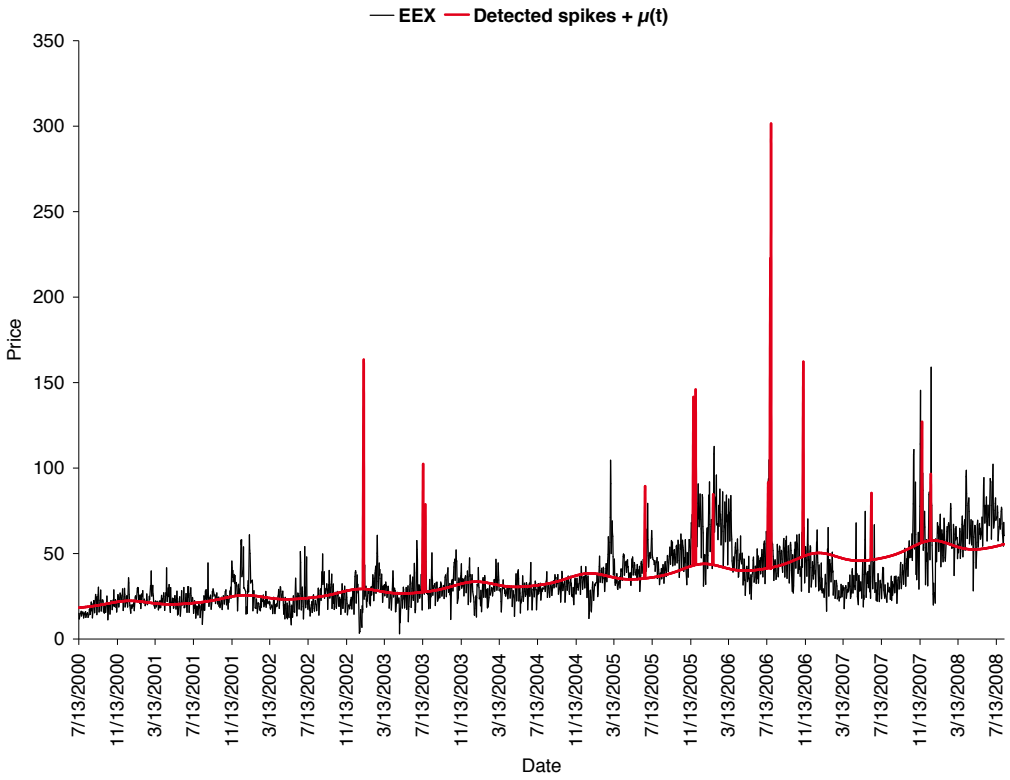


Figure 2.5: EEX price process together with calibrated seasonality and detected spikes.

Geman & Roncoroni [2006] suggest choosing the maximum jump size ψ as the observed maximum daily absolute variation in log-prices, which for our data set takes the value 2.2361. Note that this parameter can give some non-realistic results in the spike size, since it can be considered as an upper limit for the spikes. The intensity shape function $s(t)$ is estimated manually in such a way that the most salient spikes coincide with the intensity shape peaks. In addition, the intensity shape should capture clusters of spikes. This procedure leads to $k = 0.7$, $d = 0.75$ and $\tau = 0.42$. In Figure 2.4 we show the resulting calibrated shape function together with the most prominent spikes, extracted according to the specified threshold Γ. For comparison with the non-parametric spike intensity, the empirical spike intensity is plotted in Figure 2.6 based on all the detected spikes. (The comparable intensity for the factor model is in Figure 2.12.)

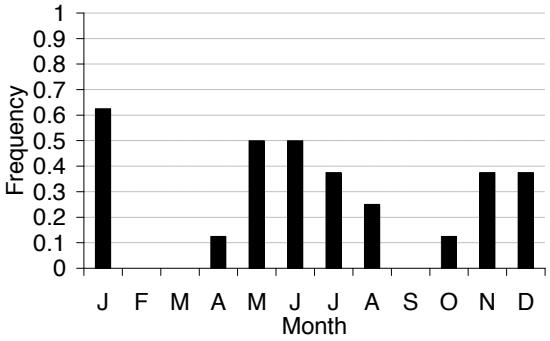


Figure 2.6: Historically based frequency of spike occurrence for the threshold model.

Geman & Roncoroni [2006] analysed the effect of the stochastic spike intensity function in Equation 2.9 on the ECAR market, and found no statistical evidence that this improved the fit. A stochastic intensity could potentially model a "spike reversion" given by a negative jump following a positive one. Figure 2.5 shows the EEX prices together with the calibrated trend $\mu(t)$ and detected spikes. Figure 2.7 depicts the price values for six typical weeks when spikes were detected (the spikes in red, the following prices in black). The general picture is that one can see an upward price spike, followed by a quick reversion back, given by one or more decreasing prices. One could attribute this to *first* having a negative spike, and next mean-reversion is dampening the prices further, *or* a sequence of negative spikes. In order to get this, one must have a low level for the stochastic spike intensity in order to create a negative spike with sufficiently high probability. But we also need the negative spike *size* to be of certain magnitude in order to push the prices sufficiently down. To have a sequence of negative spikes, we must have a very low threshold for the stochastic spike intensity, as well as the threshold for having negative sign of the jumps. This will most likely be in conflict with the estimation of positive price spikes that we naturally want to include in the model. The alternative to having negative spike(s) following a positive price spike is strong mean-reversion. Indeed, the mean-reversion speed does not have to be increased very much in order to have a decay like we see in Figure 2.7. This discussion suggests that the maximum-likelihood estimator will put more emphasis on a mean-reversion adjustment rather than the stochastic spike intensity, since after

all there are fairly few data for the spikes to rely the estimation on. As the variations in the EEX data are already quite big for the "non-spike" regime, coupled with strong reversion, we have the potential for large price decays without the help of negative spikes.

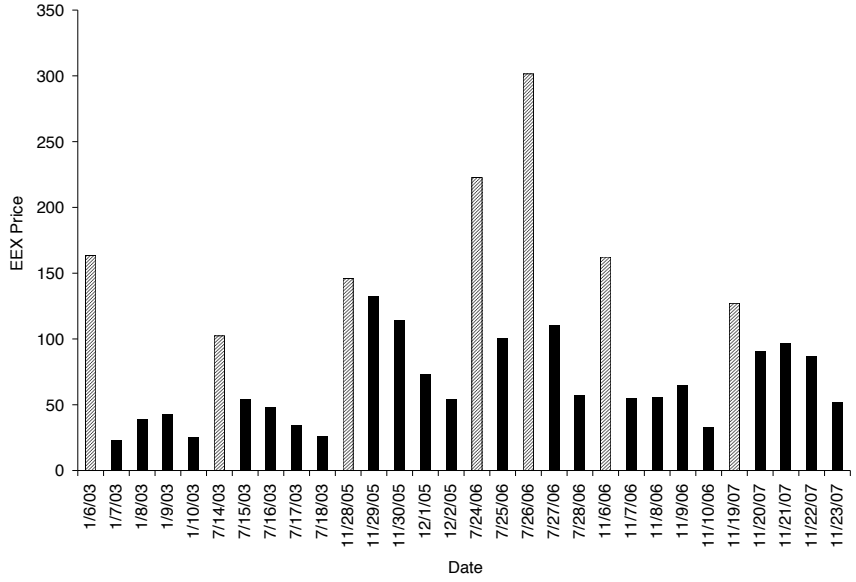


Figure 2.7: EEX price values for six selected weeks when spikes were detected. The spikes are in red, the following prices are in black.

Estimation of the model parameters

After selecting the structural parameters, we estimate the model parameters. First of all, a jump threshold Γ is set to filter out ΔE^d and E^c, i.e. the jump and continuous paths, respectively. Then, based on this, the smooth mean-reversion force θ_1, the maximal expected number of jumps θ_2, the reciprocal average jump size θ_3 and the Brownian local volatility σ are estimated. The fourth moment of the fully specified model is then computed, and compared to the empirical one. If the deviation is too large, the jump threshold Γ is either de- or increased, and the whole procedure is repeated. Geman & Roncoroni [2006] propose iterating the procedure of choosing the jump threshold Γ until the estimated model matches the fourth moment of the daily log-price return distribution.

The parameters $\Theta = (\theta_1, \theta_2, \theta_3)$ are to be estimated by approximate maximum likelihood. The log-likelihood function explicitly depends on Θ and the filtered data set $(\Delta E^d, E^c)$ and implicitly on the choice of Γ, which is used to obtain ΔE^d and E^c. The approximate logarithmic likelihood function is given by Geman & Roncoroni [2006] as

$$\mathcal{L}(\Theta \mid \Theta^0, E) \quad = \quad \sum_{i=0}^{n-1} \frac{(\mu(t_i) - E_i)\theta_1}{\sigma^2} \Delta E_i^c - \frac{\Delta t}{2} \sum_{i=0}^{n-1} \Big(\frac{(\mu - E_i)\theta_1}{\sigma} \Big)^2$$

$$-(\theta_2 - 1)\sum_{i=0}^{n-1} s(t_i)\Delta t + N(t)\ln\theta_2$$
$$+\sum_{i=0}^{n-1}\Big[-(\theta_3 - 1)\frac{\Delta E_i^d}{h(E_i)}\Big] + N(t)\ln\Big(\frac{1-e^{-\theta_3\psi}}{\theta_3(1-e^{-\psi})}\Big). \tag{2.17}$$

The first part represents a discretized version of the Doléan-Dade exponential for continuous processes. The remaining parts are responsible for the jump process. The log-likelihood function explicitly depends on θ_1, θ_2, θ_3 and the "Γ - filtered" data set. The likelihood function is maximized with respect to θ_1, θ_2 and θ_3 over a bounded parameter set Θ, taking some economically sound limiting values into account. The function is constructed in such a way that one can split it up into three independent parts and maximize them separately

$$L(\Theta \mid \Theta^0, E) = F_1(\theta_1) + F_2(\theta_2) + F_3(\theta_3). \tag{2.18}$$

Such a modification helps to facilitate the optimization algorithm and to increase the estimation correctness.

The next parameter to be estimated is the volatility of the continuous path. We use

$$\sigma = \sqrt{\frac{1}{T}\sum_{i=0}^{n-1}\Delta E^c(t_i)^2}\,. \tag{2.19}$$

This estimator for the volatility was applied in Roncoroni [2002], and is based on the quadratic variation of the continuous path; see Genon-Catalot & Jacod [1993] for details. Table 2.3 contains the estimates of the model parameters.

Table 2.3: Estimates of the threshold model parameters.

Element	Interpretation	EEX
θ_1	Smooth mean-reversion force	0.2480
θ_2	Max. expected number of jumps	14.5144
θ_3	Reciprocal average jump size	1.0584
σ	Brownian local volatility	3.8216
Γ	Jump threshold	0.6750

An estimate of the expected number of jumps during the period is provided by the integral of the intensity function over the whole period, resulting in $\mathbb{E}[N(1)] = 2.6885$. The number of filtered jumps is 39, a number which depends strongly on the selected jump threshold Γ.

To understand the estimated speed of mean-reversion θ_1 better, it is worthwhile to find the *half-life* of the mean-reversion. The concept of half-life takes its origin from physics and generally describes a period of time it takes for a substance undergoing decay to decrease by half. The half-life of an OU process is defined as the average time it takes before a price jump reverts back to

half of its original value (see Clewlow & Strickland [2000]). Mathematically it can be calculated as $(\ln 2)/\lambda$, where λ is a decay constant, i.e. the mean-reversion speed. Here the estimated θ_1 is yielding a half-life of 2.79 days, that is, it takes the process on average slightly less than 3 days to revert back to its mean. In the factor model, where we separated the spike process from the base signal, we found that the spikes had a half-life of around 2 days, whereas the base signal was initially estimated to have a half-life of around 3.5 days. We see that the threshold model has a half-life approximately the average of these two figures. This clearly demonstrates that the threshold model is not capable of allowing a mean-reversion which pushes spikes back fast enough on the one hand, and at the same time is sufficiently slow to push back the lower variations in the price path in quieter periods. In order to make up for the faster mean-reversion for the "base signal", one may expect a upward bias in the volatility estimate.

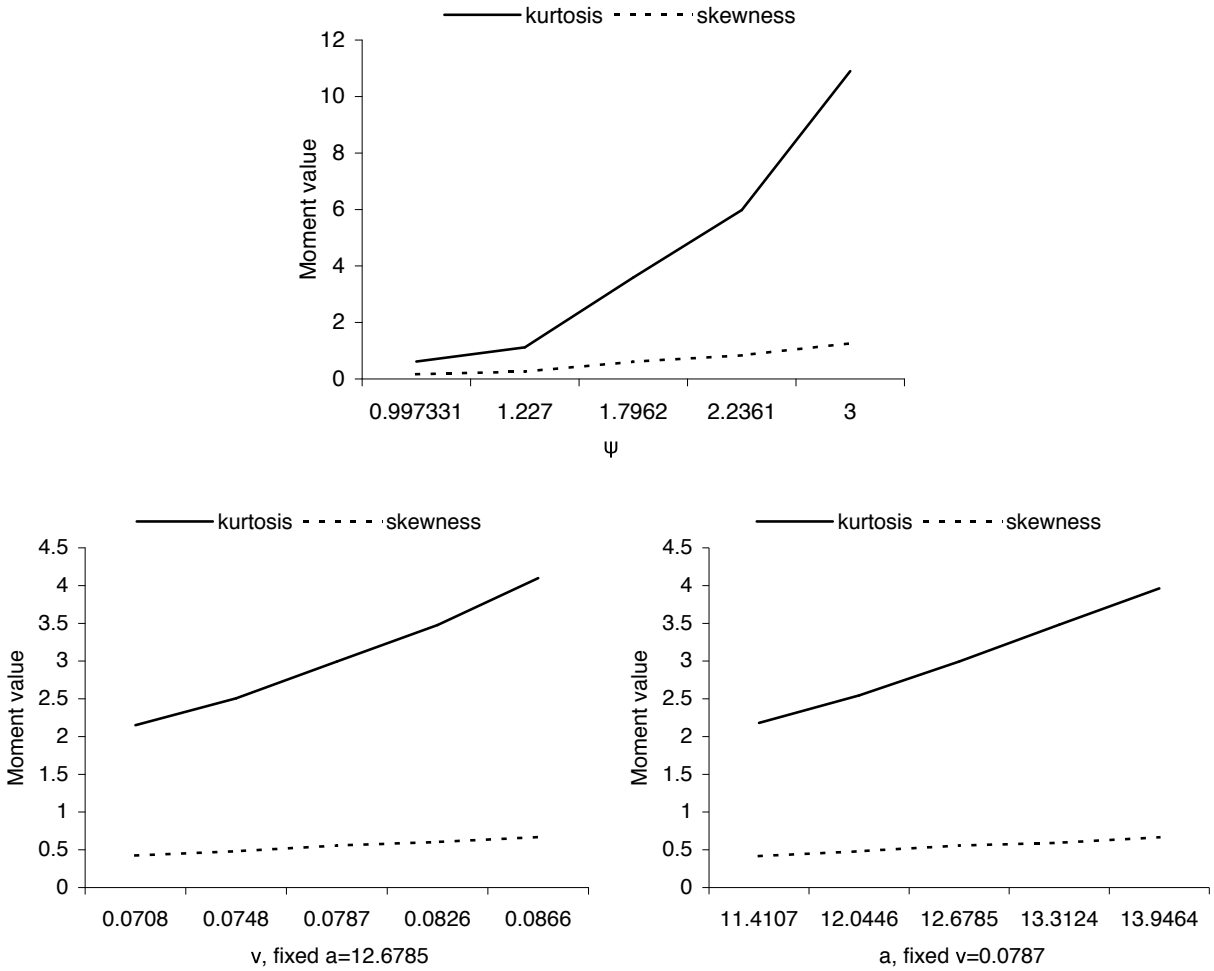


Figure 2.8: Top: Sensitivity check for the skewness and the kurtosis, when the parameter ψ takes various values. Bottom left: Sensitivity check for the skewness and the kurtosis, when the parameter ν takes various values, while the other parameter is at fixed value $\alpha = 12.6785$. Bottom right: Sensitivity check for the skewness and the kurtosis, when the parameter α takes various values, while the other parameter is at fixed value $\nu = 0.0787$.

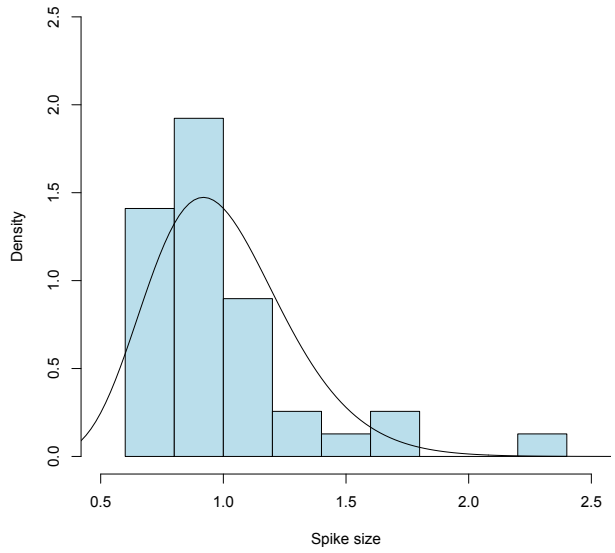


Figure 2.9: Fitted Gamma distribution versus the spike size histogram for the threshold model.

Based on the filtered jumps, one may look for alternative spike-size distributions which are more realistic than the truncated exponential. As mentioned earlier, by truncating the spike-size distribution, one essentially introduces an upper limit for possible jumps. By using the historical price changes future price behaviour will repeat the past. In fact, the estimate on the upper bound ψ may become very unstable, since the spikes in most electricity price series are of very different size, even by several magnitudes. In order to see this effect in the data, we based the estimate of ψ on different price changes. For example, rather than basing our truncation on $\psi = 2.2316$, corresponding to the three biggest price changes, we may choose the second biggest price change instead to obtain $\psi = 1.7962$. The effect on the moment properties of the path turns out to be dramatic. In Figure 2.8 we plot the skewness and the kurtosis of the price path as a function of ψ. The estimates of these two moments are based on a number of simulations. The simulations have been repeated until the change of the averaged moment value becomes less than 0.01%. We see that the range of the kurtosis varies dramatically with the choice of ψ. Thus, using historical price changes to truncate the spikes may lead to a very unreliable model which may seriously fail to capture the true distribution of spike sizes. We also plot the skewness and the kurtosis of the price path as a function of α and ν, when using a Gamma law for spikes sizes. By fixing one parameter and varying another, we can see their influence on the moment values. It is obvious that the range of the kurtosis and the skewness changes slightly with the parameters. This leads us to affirm that Gamma-distributed spikes yield more stable moments than the truncated exponential distribution for spikes. To cope with this defect, we have fitted a Gamma distribution for the spikes as we did with the factor model. In Figure 2.9 we have plotted the estimated Gamma distribution together with the empirical density of spikes. The maximum likelihood parameters were found to be $\alpha = 12.6785$ and $\nu = 0.0787$. Using a non-truncated distribution, we include the possibility of observing bigger jumps than the historically observed ones.

2.5.4 Factor model calibration

The first step in calibrating the factor model is to assess the number of factors required. We compare in the L^2 norm the empirical autocorrelation function (ACF) with the theoretical autocorrelation functions from Equation (2.15) for different numbers of factors. We obtain $n = 2$ as the optimal number of factors. The estimated speeds of mean-reversion and weights are reported in Table 2.4 and the estimated and empirical ACFs are in Figure 2.10.

Table 2.4: The fitted ACF with a sum of two exponentials.

λ_i	$\widetilde{w}_i$
0.0087	0.3547
0.3333	0.6453

We associate $\lambda_2 = 0.3333$ to the spikes having the fastest speed of mean-reversion. The base signal is associated to the factor Y_1 with estimated mean-reversion $\lambda_1 = 0.0087$. We find that the half-life of the base signal Y_1 is 79.6721 and for the spike 2.0794. Thus, on average the base signal needs nearly 80 days to come back to half its value while a spike needs only 2 days. After filtering out the spikes, we will re-estimate the speed of mean-reversion for the base signal.

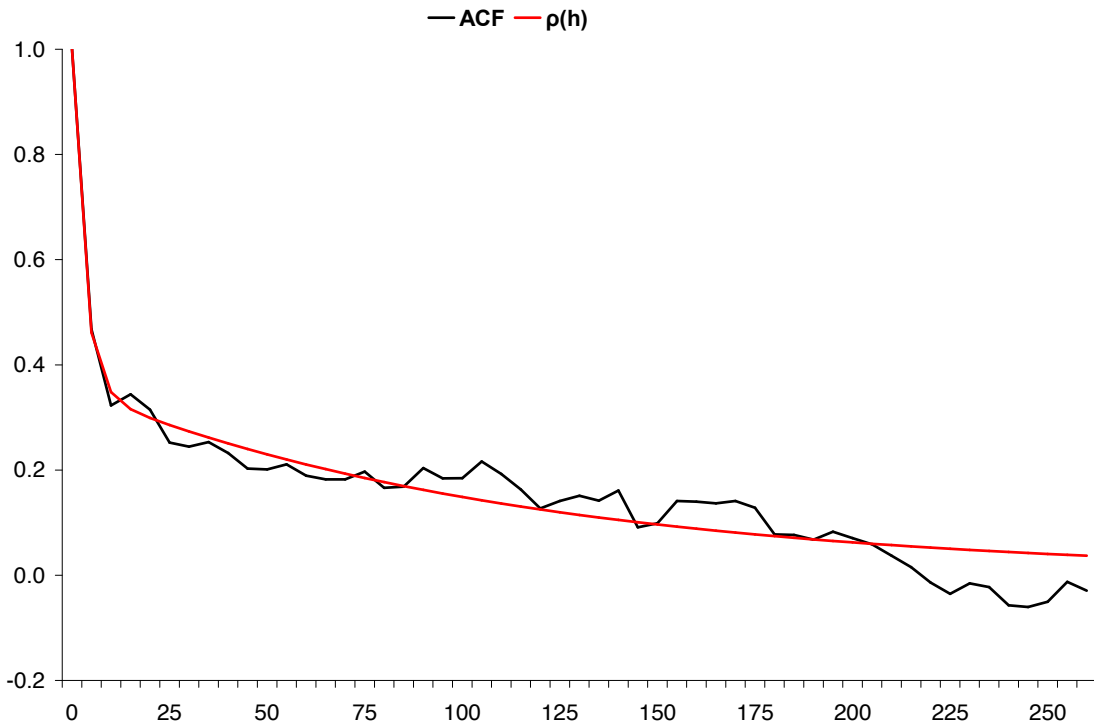


Figure 2.10: Empirical ACF for EEX series and weighted sum of two exponentials.

Filtering of spikes

The next step is to filter out the spikes. Meyer-Brandis & Tankov [2008] show that the hard-thresholding procedure taken from Extreme Value Theory to identify the spikes is a reliable technique in the context of return-distribution characteristics. In short, it filters out the spike process using methods from non-parametric statistics and provides as output both the base signal and the spike process. We refer to Meyer-Brandis & Tankov [2008] and Nazarova [2008] for a detailed explanation of the approach to the EEX data. In Figure 2.11 one sees the result of the hard-thresholding procedure: the spikes and the base signal on log-scale.

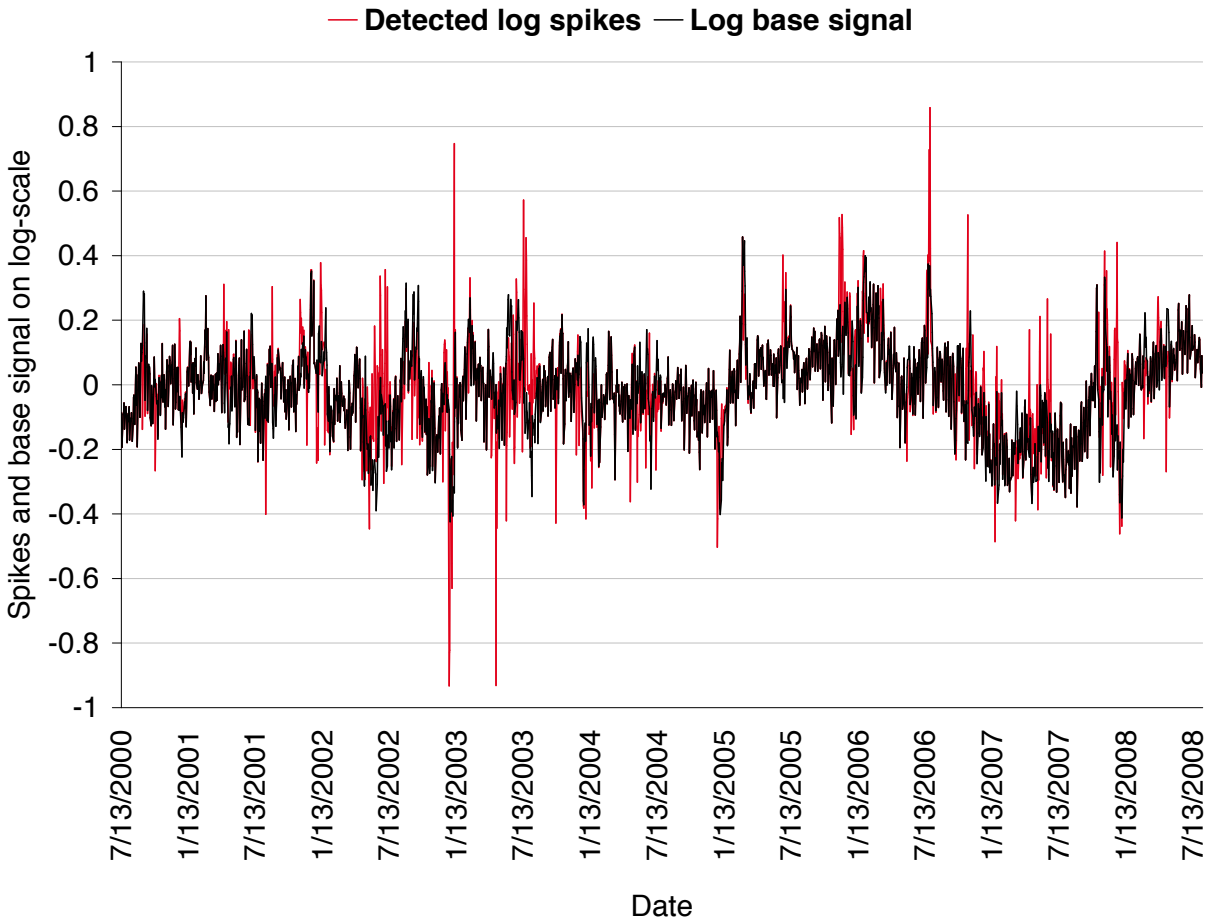


Figure 2.11: Detected log spikes and log base signal.

As found in Meyer-Brandis & Tankov [2008] the method of hard-thresholding is insensitive to the value of λ_1 in the range of 0.1 to 0.01 and relatively insensitive to the value of λ_2. Recall our estimate of $\lambda_1 = 0.0087$, being close to the desired range.

Table 2.5: Resulting estimates for OU process.

Element	Value
λ_1	0.2008
α	13.3009
ν	8.5689

Estimating the base signal

We continue to estimate the parameters of the model for the base signal. The issue here is to find the right subordinator process $L_1(t)$ which fits the observed filtered time series. The usual way is to propose a stationary distribution that fit the data, and find the so-called *background driving Lévy process* $L_1(t)$ such that $Y_1(t)$ has the same stationary distribution. A typical choice is the Gamma distribution, which leads to a compound Poisson process $L_1(t)$ with exponential jumps (see Benth *et al.* [2007] and Barndorff-Nielsen & Shephard [2001]). The reason for the choice of the Gamma distribution is the availability of an explicit analytical expression for the moments, otherwise we would need to use some numerical methods.

As suggested by Benth *et al.* [2007], we apply a method based on prediction-based estimating functions developed by Sørensen [2000] and Bibby *et al.* [2010] to calibrate the base signal model to data. The details of the method can be found in A.2.

Table 2.5 gives the parameter estimates from the implementation of the prediction-based estimating functions technique. The re-estimate of the speed of mean-reversion λ_1 implies a half-life of approximately 3.5 days, much faster than initially estimated by matching autocorrelation functions. In view of the very noisy behaviour in the base signal, this seems more likely than 80 days as initially found.

The method has its advantages and disadvantages. It is well-grounded from the theoretical point of view. However, for practical applications we face the problem that the Equation $G_n(\theta) = 0$ has no unique solution. Therefore, the algorithm may find different roots for different initial parameter values. However, if the initial parameters are close to the true one, the resulting estimated parameter values are correct in the sense that they match the moments. Bibby *et al.* [2010] suggested finding some optimal weights to improve the efficiency of the estimator, but in the case of multiple solutions this approach does not help significantly. Therefore, our calibration of the base signal was carried out in two steps: first a calibration "by hand" to identify likely intervals for the parameters values, then execution of the prediction-based estimating functions algorithm using these initial values.

Analysis of the spike process

The final step in our estimation procedure of the factor model is to calibrate the spike process. Since spikes are rather sparse compared to the total length of the data set there are few data points available for estimating the intensity and the jump-size distribution. To cope with this problem we consider various specifications and analyse their consequences.

We shall apply the popular shape function proposed by Geman & Roncoroni [2006], given in Equation (2.8), and estimate the parameters from the data at hand. An alternative approach would be to choose an intensity based directly on the observed distribution of spikes over the year. In Figure 2.12 we have plotted the historical frequency of the spike occurrence. To use the historical frequency has the advantage of an easy and fast adjustment as new market data become available. Furthermore, the calibration procedure is very simple compared to parametric approaches.

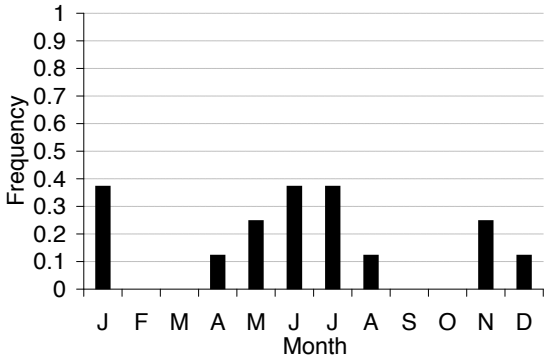


Figure 2.12: Historically based frequency of spike occurrence for the factor model.

Recall the intensity function $\iota(t,\theta)$ defined in Equation (2.7), where we included a dependency on the parameter vector $\theta = (\theta_2, k, d)$. The maximum likelihood estimator of θ is given by (see Meyer-Brandis & Tankov [2008])

$$\theta^* = \arg\max_{\theta} \Big(\sum_{i=1}^{N_T} \log \iota(\tau_i, \theta) - \int_0^T \iota(t,\theta)\ \mathrm{d}t \Big), \tag{2.20}$$

where τ_i are spike times. A time-dependent intensity function is the natural choice when there is some pronounced seasonality in spikes. The US markets analysed in Geman & Roncoroni [2006] demonstrate evident spike seasonality, while in the EEX market this is not so obvious. Out of the 30 biggest positive spikes, 16 occurred in summer, 7 in winter, the remaining 7 in spring and fall. We found a phase $\tau = 0.42$, with $d = 1.0359$, $k = 0.5$ and $\theta = 14.0163$ (as θ_2 in the threshold model). Figure 2.13 shows the intensity shape and the largest 30 spikes, detected by the hard-thresholding procedure. As can be seen from the picture, EEX data do not demonstrate such a pronounced seasonality in spikes as could be suggested from the intensity shape function.

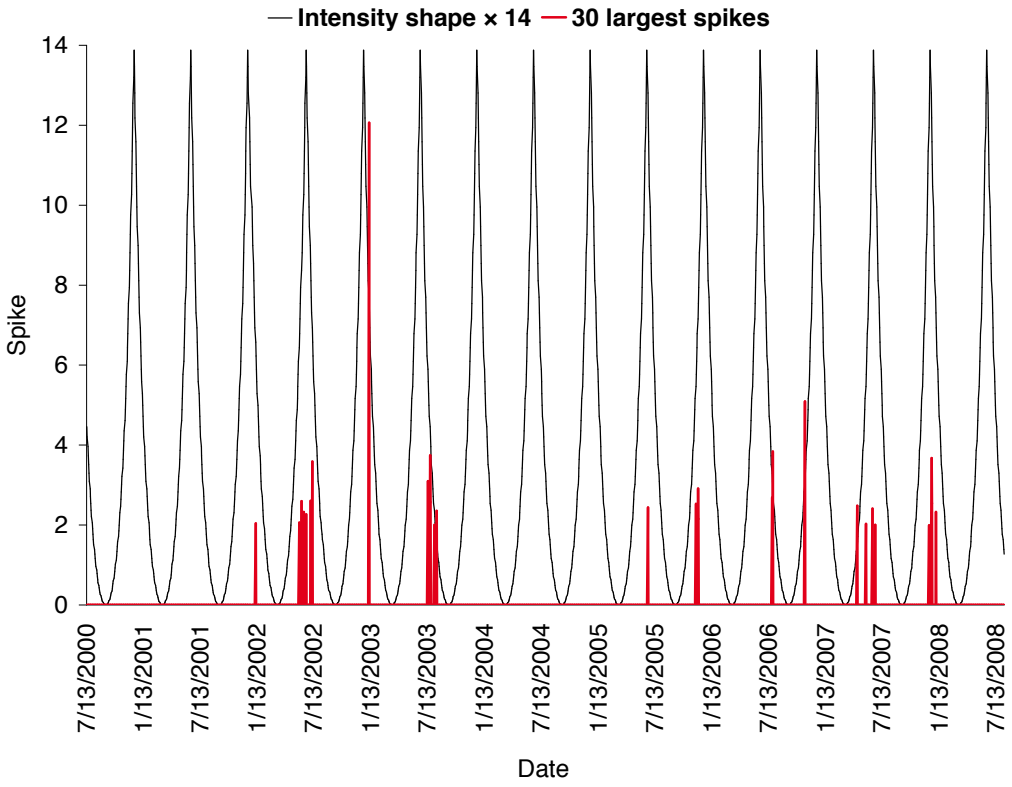


Figure 2.13: Calibrated intensity shape function and real values of the largest detected spikes.

The next problem is to estimate a jump-size distribution for the spikes. In Meyer-Brandis & Tankov [2008], a Pareto distribution was suggested for the spike sizes. The estimation procedure involves a threshold or scale parameter z_0, the smallest value the Pareto random variable may take, and a parameter α for the tail-fatness. These are estimated by means of fitting a straight line to the empirical cumulative distribution function (CDF) on log-log scale, i.e. a traditional Hill estimator, which is efficient when the underlying distribution is Pareto; see Drees *et al.* [2000] for details. We show the result in Figure 2.14. We find the estimates $z_0 = 0.3648$ and $\alpha = 2.5406$.

The Pareto distribution has very heavy tails, and may give unreasonably high values of the spot price. An alternative distribution for fitting the spike sizes may be the Gamma distribution. In Figure 2.15 we compare the fitted Gamma density with the empirical spike-size density. The maximum-likelihood estimates for the two parameters α and ν of the Gamma distribution are $\alpha = 6.2592$ and $\nu = 0.0942$.

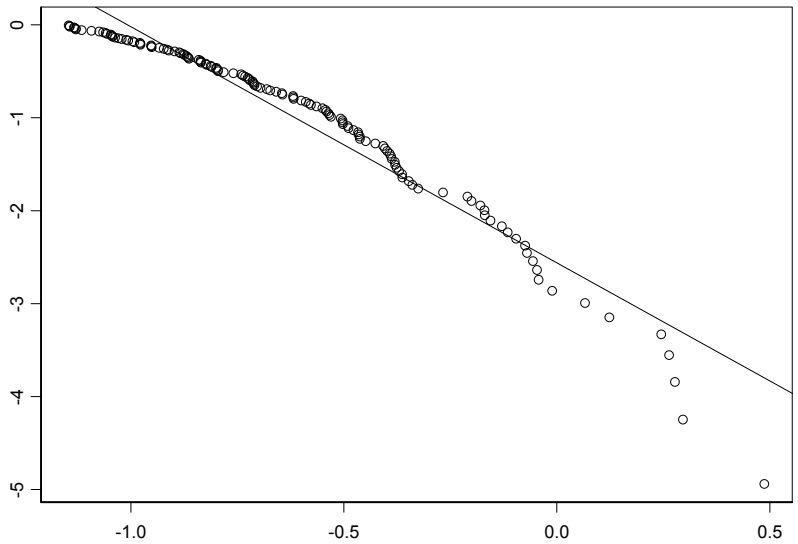


Figure 2.14: Empirical CDF of spike size on log scale.

2.6 Model comparison

In this section we assess our estimated models, and discuss their properties in the context of the EEX electricity spot price behaviour. In Figure 2.16 we have plotted typical simulated paths of the models, along with the observed EEX prices series. Visually the performance of all three models is quite satisfactory. One apparent difference is that the factor model seems to be less noisy in the intra-spike periods than the data. The jump-diffusion and the threshold models, on the other hand, are more noisy than the data, at least according to the experience from our simulation studies. The spike pattern looks better for the factor model in these simulations.

A standard, widely used model check is to compare model-based moments to the empirical ones. In our particular case it may not be reliable since the threshold model in fact is calibrated using the fourth moment as a target and thus should match at least the kurtosis almost perfectly. We report the first four moments of the returns in Table 2.6 for the jump-diffusion, the threshold and the factor models together with the empirical moments of the EEX data. The descriptive statistics are computed for the empirical versus simulated logarithmic price variations, i.e. log returns. The simulations have been repeated until the change of the averaged moment value becomes less than 0.01%.

Table 2.6: Empirical moments versus jump-diffusion, threshold and factor model moments.

Moment	Average	Std. Dev	Skewness	Kurtosis
EEX	0.0006	0.2985	0.4050	6.6179
Jump-diffusion model	0.0007	0.3191	0.8343	10.3935
Threshold model	0.0006	0.2935	0.8336	5.9783
Factor model	0.0006	0.1595	1.6749	10.5308

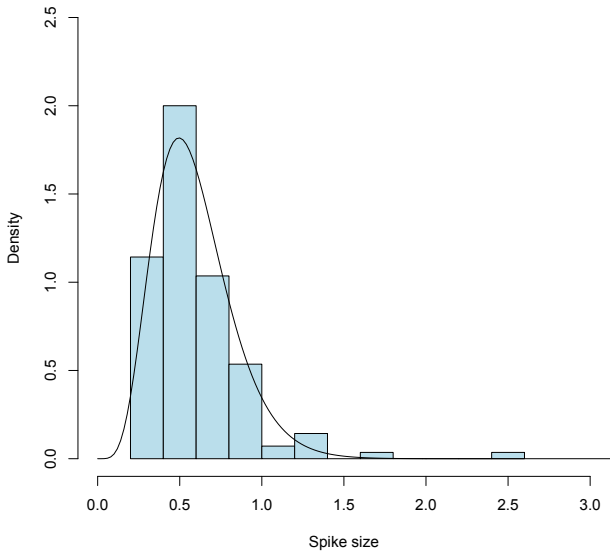


Figure 2.15: Fitted Gamma distribution versus the spike size histogram for the factor model.

The figures in the table indicate the best performance for the threshold model. It matches the first four moments very well, except possibly the skewness, where it overshoots. However, the close resemblance with the empirical moments is not really a surprise since these have been used as a measure in the estimation procedure. The jump-diffusion model, on the other hand, shows two apparent characteristics. It overestimates the skewness and kurtosis, which can be explained by the use of a Normal distribution for the spike sizes.

The factor model also demonstrates two very apparent characteristics. Firstly, it underestimates the standard deviation significantly, yielding a simulated volatility of the path only half the size of the empirical. This can be explained from using the hard-thresholding procedure. Meyer-Brandis & Tankov [2008] suggest choosing a threshold value for the deseasonalized log-returns such that the share of returns larger than the chosen threshold value does not exceed 5% of all returns. Then, the standard deviation of the remaining 95% is computed and called the target one, according to which we can separate spikes and the base signal. It is obvious that the total number and size of filtered spikes depend critically on the threshold value. In our case we obtained the target standard deviation equal to 0.1454, which looks quite consistent with the calibrated standard deviation. Of course, if we change the criterion in the hard-thresholding procedure, we will receive different results for the target standard deviation and the filtered spikes and the base signal, respectively. Decreasing the share of returns above the threshold to 1.76%, we obtain a larger target standard deviation, which may be more in line with the observed one. Possibly, one could think of an iterative procedure parallel to finding the Γ in the threshold model. The kurtosis of the estimated factor model is close to twice as big as the empirical. This is a result of using a Pareto law for the spike sizes, implying rather extreme jump sizes which obviously influences the kurtosis. Note that the high positive skewness of the factor model is a result of the large spikes.

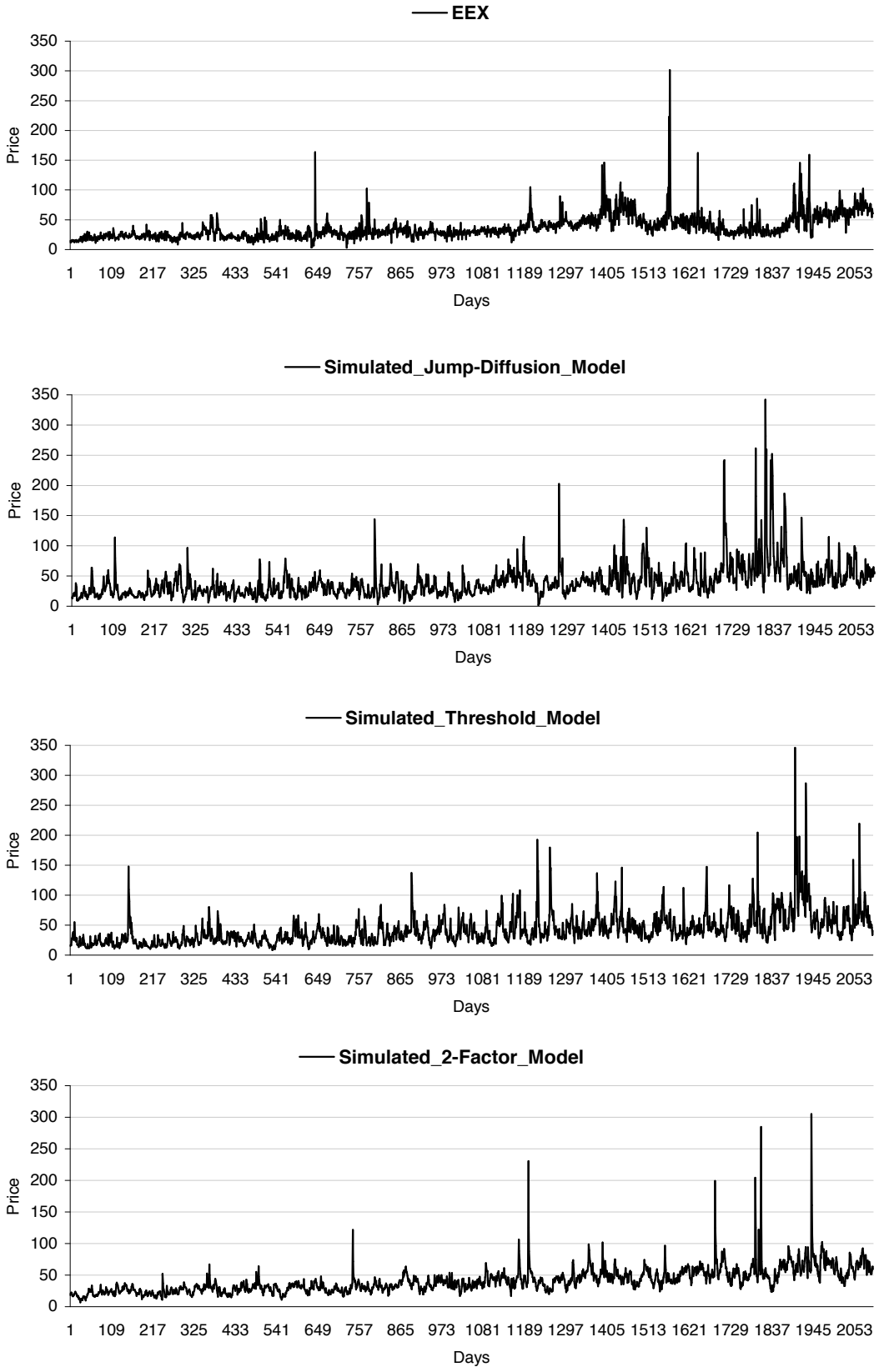


Figure 2.16: EEX, jump-diffusion, threshold and factor model simulated price paths.

Table 2.7: Base signal moments matching.

Moment	EEX base signal	Simulated
Average	1.0262	1.1241
Std. Dev.	0.3198	0.2637
Skewness	1.0177	0.4405
Kurtosis	2.4674	0.6756

The truncated exponential law used in the threshold model prevents the occurrence of any spikes bigger than the ones we have already observed, which clearly helps in getting a kurtosis close to the observed one. However, it is not clear whether matching the empirical kurtosis is a useful fitting criterion. The empirical kurtosis takes only the observed changes into account, and as we pointed out during the sensitivity analysis with respect to the truncation parameter in the jump size distribution for the threshold model, it may be very sensitive to changes in the data. For instance, the occurrence of new spikes bigger than the observed leads to a higher kurtosis. In practice, one should be careful since the empirical moments are backward looking. The Pareto, Gamma or Normal distribution may in fact give a much better picture of the true risk in the future.

As we could see in Figure 2.16, the noise in the intra-spike periods is too low in the factor model, an observation that is confirmed by the simulated standard deviation of the path. This may be attributed to the choice of the base signal model. Its relatively slow mean-reversion combined with a compound Possion process as driving noise will yield paths which may not look as random as the market. The compound Poisson process will only jump at Poisson distributed random times, and otherwise the base signal will decay exponentially. Although the stationary distribution of the model honours the statistical properties of the base signal, we may have paths which are too regular. A possible modification could be to include a Brownian-driven factor. This would possibly lead to negative prices, which may be a problem in some applications. However, in the market, and in particular the EEX market, negative prices occur from time to time. Alternatively, we may use a subordinator which is not given by a finite-activity process, but one with infinite activity such as the background driving Lévy process derived from supposing an inverse Gaussian stationary distribution of the base signal.

To further analyse the base signal properties, we look at the moments of the data after the spikes have been filtered out. In Table 2.7 we report the first four simulated moments of the factor model, together with the empirical ones. Note that here we consider the moments of the process, but not its log-returns, as we do while comparing models' performance. It is very obvious from these numbers that the base signal model does not sufficiently explain the variations in the data. Turning to the jump-diffusion or to the threshold model, we can mimic the base signal by considering the dynamics without the jump component. Since this is a mean-reverting model on log-scale, the variance in stationarity is given by $\sigma^2/2\alpha$ or $\sigma^2/2\theta_1$, and inserting the estimates this gives us a stationary standard deviation of 5.8111 or 5.4263 respectively. The figures are far

above what is reasonable to expect by looking at the empirical standard deviation. Hence, we see that due to the too fast speed of mean-reversion for the base signal in the jump-diffusion and the threshold models, it compensates by gearing up the volatility to unrealistic levels. Combined with a too slow mean-reversion for the spikes, one may question whether the pathwise properties of the spot price are honoured in the threshold model. We remark that in Oyebanji [2007] an even higher σ was estimated for EEX data.

To better understand the effect of having a different spike size distribution, we apply the Kolmogorov-Smirnov test to check the goodness of fit of the different spike size distributions. We test the H_0 hypothesis of the suggested distribution with estimated parameters at the 5% confidence level. The results are reported in Table 2.8.

Table 2.8: Kolmogorov-Smirnov test results for the considered spike distributions.

Distribution	Test statistics	Decision on H_0
Normal (jump-diffusion model)	2.1855	reject H_0
Trunc. exp (threshold model)	3.5032	reject H_0
Pareto (factor model)	1.7748	reject H_0
Gamma (threshold model)	0.9375	no reason to reject H_0
Gamma (factor model)	1.1556	no reason to reject H_0

From the results we can observe that the data present sufficient evidence to contradict our hypothesis that the spike sizes follow Normal, truncated exponential or Pareto distributions in case of the jump-diffusion, the threshold or the factor models respectively. However, there is no reason to reject the hypothesis on the choice of Gamma law for spike-size distribution for the threshold and the factor models. Therefore, we can conclude that the results of the Kolmogorov-Smirnov test support the modification of the spike size distribution. We do not provide the results of the test for the Gamma distribution for the jump-diffusion model. We analyse the jump-diffusion model to have a fair comparison of two complex models with one parsimonious and to check the effect of the model specification on the results it gives. Therefore, we omit the modifications for the jump-diffusion model here.

Table 2.9: Comparative descriptive statistics results for the log-returns of the jump-diffusion, threshold and factor models.

Moment	Average	Std. Dev	Skewness	Kurtosis
EEX spot	0.0006	0.2985	0.4050	6.6179
Jump-diffusion model (Normal)	0.0007	0.3191	0.8343	10.3935
Threshold model (trunc. exp)	0.0006	0.2935	0.8336	5.9783
Factor model (Pareto)	0.0006	0.1595	1.6749	10.5308
Modified threshold model (gamma)	0.0006	0.2822	0.5566	2.9946
Modified factor model (gamma)	0.0006	0.1465	1.2414	5.7399

Coming back to comparing the simulated moments of the three models with the empirical ones based on the Gamma distribution for the jump sizes, we report the result in Table 2.9. We ob-

serve that the kurtosis of the threshold model drops significantly, and is now underestimating the empirical kurtosis. The reason is that although the Gamma distribution allows for unbounded jumps, it is concentrated more on the smaller spikes sizes, while the truncated exponential is more stretched and gives a higher probability to bigger spikes. The factor model is more in line with the empirical kurtosis when using the Gamma distribution for spike sizes. The standard deviation is still insufficient for the factor model, whereas the threshold matches very well even for Gamma-distributed spikes. But in this comparison, which is fairer since we use the same model for spikes, the factor model seems to be better at capturing the leptokurtic effects of the model. Of course, being faithful to the estimation procedure of Geman & Roncoroni [2006], the iterative approach combined with estimating parameters so that the fourth moment is matched would also yield a perfect fit to the empirical kurtosis for the case of a Gamma distribution. However, as we can see here, the factor model can obtain a significant improvement by simply changing the jump distribution if this is desirable, whereas the threshold model requires a complete re-estimation although we only introduce a new distribution for the identified spikes being more in line with the observed jump sizes. Re-estimating the complete model as in Geman & Roncoroni [2006] would lead to a different estimation function (see Equation (2.17)) and different parameter estimates, showing that the speed of mean-reversion and volatility, for instance, are indeed very sensitive to the choice of the spike-size distribution. Therefore, the model cannot easily be adjusted to new assumptions.

2.7 Application to derivative pricing

In this section we apply the three spot models to pricing of forward contracts. To assess the performance of the models we compare theoretical prices with the observed ones, and compute the implied risk premium. We have available analytical forward prices for the jump-diffusion and factor models, while we use a simulation-based Monte Carlo pricing method for the threshold model.

The price at time t of the forward contract with maturity T is defined as the expected value of the spot price at expiry date under an equivalent pricing measure $\mathbb{Q}$

$$F^{\mathbb{Q}}(t,T) = \mathbb{E}_t^{\mathbb{Q}}\left[S(T) \,|\, \mathcal{F}_t\right]. \tag{2.21}$$

Since the electricity market is an incomplete market, there exist many pricing measures $\mathbb{Q}$. To pin down a price $F(t,T)$, one needs to select *one* such measure, and this would typically be done in practice by restricting the space of measures $\mathbb{Q}$ to a parametric class, for example given by the Girsanov and Esscher transforms (see Benth *et al.* [2008a] for the details). Next, by minimizing the distance between theoretical and observed prices, one could find a pricing measure.

For the three spot models under inquiry, one has different classes of pricing measures $\mathbb{Q}$ avail-

able. The jump-diffusion and threshold models would naturally involve a change of measure for Brownian motion and compound Poisson processes, whereas the factor model does not involve any Girsanov transformations, but a change of measure with respect to subordinators. In order to have a fair comparison of the forward pricing ability of the three spot models, we focus on the implied risk premium. This does not involve any change of measure $\mathbb{Q}$, and therefore avoids introducing properties of the models which are depending on the selection of pricing measure. The empirical risk premium is computed as the difference between the predicted spot price and the observed market price (see Benth *et al.* [2008a])

$$RP(t) = F^{observed}(t,T) - F^{\mathbb{P}}(t,T). \tag{2.22}$$

Here, $F^{\mathbb{P}}(t,T)$ is the so-called *predicted* spot price at delivery T, computed as in (2.21) with $\mathbb{Q} = \mathbb{P}$. A careful analysis of the empirical risk premium is not only valuable for understanding the spot model performance, but it is also the fundament for proposing classes of pricing measures $\mathbb{Q}$. Moreover, it has obvious applications to risk management.

In electricity markets the forward contracts deliver power (or the money equivalent to power) over a specified period. These periods may typically be a month, a quarter or a year. In the empirical studies to come, we shall focus on monthly delivery periods, and the risk premium is extended in a natural way as simply taking the average of the right-hand side of (2.22) over the delivery period. That is, the risk premium for a forward contract delivering electricity over the period $[T_1, T_2]$ is

$$RP(t) = F^{observed}(t,T_1,T_2) - F^{\mathbb{P}}(t,T_1,T_2), \tag{2.23}$$

where

$$F^{\mathbb{P}}(t,T_1,T_2) = \frac{1}{T_2 - T_1}\int_{T_1}^{T_2} F^{\mathbb{P}}(t,T)\ \mathrm{d}T.$$

2.7.1 Pricing formulas

For the jump-diffusion model, we can compute the predicted spot price semi-analytically

Proposition 1 (Forward price for the jump-diffusion model)**.** *The predicted spot price of the jump-diffusion model is given by*

$$\begin{aligned} F^{\mathbb{P}}(t,T) &= (X(t))^{e^{-\alpha(T-t)}} \exp\left(\mu(T) + \int_t^T \frac{1}{2}\sigma^2(s)e^{-2\alpha(T-s)}\,\mathrm{d}s \right. \\ &\quad \left. + \int_t^T \left[e^{-\frac{\sigma_J^2}{2}e^{-\alpha(T-s)} + \frac{\sigma_J^2}{2}e^{-2\alpha(T-s)}}\right] l\ \mathrm{d}s - l(T-t)\right) \end{aligned} \tag{2.24}$$

Proof. We refer to Cartea & Figueroa [2005] for a detailed proof. □

The formula for $F^{\mathbb{P}}(t,T)$ involves the rolling historical volatility $\sigma(s)$. In our investigations, we have set this to a constant, defined as the average of the rolling historical volatility over

the period with data, see Table 2.2. Furthermore, to derive $F^{\mathbb{P}}(t, T_1, T_2)$ for a delivery period $[T_1, T_2]$, we take the average of the computed forward prices $F^{\mathbb{P}}(t, T)$ for each day T in the delivery period.

One of the beneficial properties of the factor model is that it provides us with analytical forward prices for contracts with a delivery period, $F^{\mathbb{P}}(t, T_1, T_2)$. In the next proposition we state this price for the case of $Y_1(t)$ and $Y_2(t)$ identified and being driven by compound Poisson processes with exponential jump distributions. Note that this means that both are stationary Gamma distributed. The resulting predicted spot price can be computed following the lines in Benth *et al.* [2007].

Proposition 2 (Forward price for the factor model)**.** *Suppose for the factor model that $Y_1(t)$ is stationary Gamma distributed $\Gamma(\nu, \alpha)$, and $Y_2(t)$ are driven by a compound Poisson process $L_2(t)$ with exponential jump size distribution with parameter γ, and jump intensity l. Then, the predicted spot price of the factor model is*

$$\begin{aligned} F^{\mathbb{P}}(t, T_1, T_2) &= \frac{1}{T_2 - T_1} \Big[\big(Y_1(t) - \frac{\nu}{\alpha}\big) \int_{T_1}^{T_2} e^{\mu(u) - \lambda_1(u-t)} \, \mathrm{d}u \\ &\quad + \big(Y_2(t) - \frac{l}{\gamma}\big) \int_{T_1}^{T_2} e^{\mu(u) - \lambda_2(u-t)} \, \mathrm{d}u \\ &\quad + \big(\frac{\nu}{\alpha} + \frac{l}{\gamma}\big) \int_{T_1}^{T_2} e^{\mu(u)} \, \mathrm{d}u \Big]. \end{aligned} \tag{2.25}$$

Proof. The result follows from a straightforward calculation. We start by plugging Formulas 2.1, 2.12 and 2.13 together. Hence, after commuting integration and conditional expectation, and using the fact that $Y_1(t)$ and $Y_2(t)$ are $\mathcal{F}_t$-measurable, we find

$$\begin{aligned} F^{\mathbb{P}}(t, T_1, T_2) &= \mathbb{E}^{\mathbb{P}} \Big[\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} S(u) \, \mathrm{d}u \,|\, \mathcal{F}_t \Big] \\ &= \mathbb{E}_t^{\mathbb{P}} \Big[\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} e^{\mu(u)} (Y_1(u) + Y_2(u)) \, \mathrm{d}u \,|\, \mathcal{F}_t \Big] \\ &= \mathbb{E}_t^{\mathbb{P}} \Big[\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} e^{\mu(u)} \Big(Y_1(t) e^{-\lambda_1(u-t)} + \int_t^u e^{-\lambda_1(u-s)} \, \mathrm{d}L_1(s) \\ &\quad + Y_2(t) e^{-\lambda_2(u-t)} \int_t^u e^{-\lambda_2(u-s)} \, dL_2(s) \Big) \, \mathrm{d}u \,|\, \mathcal{F}_t \Big] \\ &= \frac{1}{T_2 - T_1} \int_{T_1}^{T_2} e^{\mu(u)} \Big(Y_1(t) e^{-\lambda_1(u-t)} + \mathbb{E}_t^{\mathbb{P}} \Big[\int_t^u e^{-\lambda_1(u-s)} \, \mathrm{d}L_1(s) \,|\, \mathcal{F}_t \Big] \\ &\quad + Y_2(t) e^{-\lambda_2(u-t)} + \mathbb{E}_t^{\mathbb{P}} \Big[\int_t^u e^{-\lambda_2(u-s)} \, \mathrm{d}L_2(s) \,|\, \mathcal{F}_t \Big] \Big) \, \mathrm{d}u \end{aligned}$$

By the independent increment property of Lévy processes, we get

$$
\begin{aligned}
F^{\mathbb{P}}(t,T_1,T_2) &= \frac{1}{T_2-T_1}\int_{T_1}^{T_2} e^{\mu(u)}\Big(Y_1(t)e^{-\lambda_1(u-t)} + \int_t^u \phi_1^{'}(0)e^{-\lambda_1(u-s)}\,\mathrm{d}s \\
&\quad + Y_2(t)e^{-\lambda_2(u-t)} + \int_t^u \phi_2^{'}(0)e^{-\lambda_2(u-s)}\,ds\Big)\,\mathrm{d}u \\
&= \frac{1}{T_2-T_1}\Big[\big(Y_1(t) - \phi_1^{'}(0)\,\frac{1}{\lambda_1}\big)\int_{T_1}^{T_2} e^{\mu(u)-\lambda_1(u-t)}\,\mathrm{d}u \\
&\quad + \big(Y_2(t) - \phi_2^{'}(0)\,\frac{1}{\lambda_2}\big)\int_{T_1}^{T_2} e^{\mu(u)-\lambda_2(u-t)}\,\mathrm{d}u \\
&\quad + \big(\phi_1^{'}(0)\,\frac{1}{\lambda_1} + \phi_2^{'}(0)\,\frac{1}{\lambda_2}\big)\int_{T_1}^{T_2} e^{\mu(u)}\,\mathrm{d}u\Big] \\
&= \frac{1}{T_2-T_1}\Big[Y_1(t)\int_{T_1}^{T_2} e^{\mu(u)-\lambda_1(u-t)}\,\mathrm{d}u + Y_2(t)\int_{T_1}^{T_2} e^{\mu(u)-\lambda_2(u-t)}\,\mathrm{d}u \\
&\quad + \int_{T_1}^{T_2} e^{\mu(u)}\Big(\frac{\nu}{\alpha}\,(1-e^{-\lambda_1(u-t)}) + \frac{l}{\gamma}\,(1-e^{-\lambda_2(u-t)})\Big)\,\mathrm{d}u\Big] \\
&= \frac{1}{T_2-T_1}\Big[\big(Y_1(t) - \frac{\nu}{\alpha}\big)\int_{T_1}^{T_2} e^{\mu(u)-\lambda_1(u-t)}\,\mathrm{d}u \\
&\quad + \big(Y_2(t) - \frac{l}{\gamma}\big)\int_{T_1}^{T_2} e^{\mu(u)-\lambda_2(u-t)}\,\mathrm{d}u \\
&\quad + \big(\frac{\nu}{\alpha} + \frac{l}{\gamma}\big)\int_{T_1}^{T_2} e^{\mu(u)}\,\mathrm{d}u\Big] \\
&\quad . \qquad (2.26)
\end{aligned}
$$

In the derivation procedure we use the fact that $\mathbb{E}[L] = \phi^{'}(0)$, where the latter is the derivative of the log-moment generating function of the process L, more precisely $\phi(x) = \ln \mathbb{E}[e^{xL(1)}]$. For the case of our processes, $\phi_1^{'}(0) = \frac{\nu\lambda_1}{\alpha}$ and $\phi_1^{'}(0) = \frac{l\lambda_2}{\gamma}$ respectively. Hence, the proposition follows. □

Due to the state-dependent sign function h in the jump term of the threshold model, it does not allow for any analytical forward prices. Therefore, we apply a Monte Carlo simulation method to price forwards $F^{\mathbb{P}}(t,T)$. We start with the simulation of the spot process algorithm of which can be found in the original paper of Geman & Roncoroni [2006]. We further compute a price $F^{\mathbb{P}}(t,T)$. To define an optimal number of simulations we use the method of control variates and take the jump-diffusion model as a benchmark. By minimising the total sum of squared differences between the simulated and the analytical forward prices for the jump-diffusion model, we define the necessary number of simulations. The Monte Carlo technique we use is properly described in Glasserman [2004] and Fusai & Roncoroni [2008]. To obtain $F^{\mathbb{P}}(t,T_1,T_2)$, we compute the forward prices for every particular day of the delivery period and then average the obtained results over this period.

2.7.2 Empirical analysis

We work with the following data sets of daily electricity forward prices collected from the EEX data sets of the summer and winter terms of different years:

- Forwards with delivery in July of 2004, 2005, 2006, 2007 and 2008, observed in the preceding month June
- Forwards with delivery in December of 2004, 2005, 2006 and 2007, observed in the preceding month November

Given the observed forward prices $F(t, T_1, T_2)$, with the delivery period $[T_1, T_2]$ being June or December, and t ranging over the working days in the month prior to delivery, we compute the implied risk premium based on the predicted spot prices derived according to the algorithms and formulas described above. In the following discussion, we focus our attention on the results of June 2007, November 2007 and June 2008. The results of other observation periods can be found in A.3.

Figures 2.17, 2.18 and 2.19 show the predicted spot price (blue curve) together with the observed forward dynamics (green curve). In addition, we have included the seasonal function (seasonality trend $\mu(t)$, Equation 5.5) over the delivery period as a reference level (red curve). In general, we observe that the average seasonal function sets the level of forward prices in the market, except for July 2008 where it seems to be a large deviation in observed prices away from the seasonal level. The implied risk premia are not converging to zero, which is an obvious implication from the delivery period feature of the electricity forward contracts. Also, the market forward prices are more volatile than the predicted spot prices, except maybe in the period before delivery starts.

The shape of the risk premia looks very similar for all three models. However, there are big differences in the values. For the July 2007 contracts, the jump-diffusion and threshold models assign a negative risk premium (see Figures 2.17a and 2.18a), whereas the factor model implies a positive premium. In fact, the premium is always positive for the factor model (except of one instance, see Figure A.-1b), with occasional very large values (see Figure 2.19c). There is a clear tendency of a decreasing risk premium for the factor model as time approaches delivery, while the two other models show evidence of an increasing risk premium for this period. A negative risk premium is in line with the theory of normal backwardation, where producers accept a reduction in price in order to reduce their price risk. However, there exist both theoretical and empirical evidence for a positive premium, explained as the retailers hedging their short-term spike risk (see Benth *et al.* [2008b]). Although we consider forward prices in the days prior to delivery, i.e., being in the short-end of the forward market, the monthly delivery period should average out this risk. From an economical point of view it seems reasonable to expect a negative premium even close to start of delivery. With this perspective in mind, the factor model does a

poor job compared to the two others. This may be attributed to the fact that the factor model did not capture the variations in the base component of the spot price very well. By using a compound Poisson process to describe the base variations, we obtain much less variations that what is obtained by a Brownian component, say, which is present in both of the two other models. Although the mean-reversion of these two are too slow for the base variations of the spot, it seemingly gives an advantage for forward pricing purposes.

Inspecting the risk premia for the jump-diffusion and threshold models more carefully, we observe some unreasonable features as well. For example, the jump-diffusion model has a sign change in the risk premium for the December 2007 contract. As indicated above, one may expect a positive premium in the short end of the market, meaning in the days prior to delivery. For the December 2007 contract, we have a positive premium when we are far from delivery, which is at stake with this (see Figures in A.3 for other such examples).

Comparing the predicted spot price path with the observed forward prices, it seems that both the threshold and jump-diffusion models are closer to explaining the market than the factor model. The difference between the two models are not too big, which is a reflection of the low frequency of spike occurrences and therefore a similar path behaviour. However, the differences are still significant, so the impact of the function h is apparent. The function h will switch the direction of a jump for exceedingly high or low prices, and thereby increase the variations. Interestingly, the factor model seems to converge towards the predicted spot price when we approach delivery, whereas the two others drive apart (this is of course also reflected in a risk premium with increasing absolute value for the jump-diffusion and threshold models, and whereas decreasing to zero for the factor model). Let us discuss this in closer detail with a view towards a potential class of measure changes $\mathbb{Q}$.

Let us simplify the discussion and consider a toy model for the spot price given by a Brownian motion driven by Ornstein-Uhlenbeck process, i.e.

$$\mathrm{d}S(t) = -\alpha S(t)\ \mathrm{d}t + \sigma\ \mathrm{d}B(t).$$

A natural measure change is a constant Girsanov transform,

$$\mathrm{d}W(t) = -\frac{\theta}{\sigma}\ \mathrm{d}t + \ \mathrm{d}B(t)$$

with θ a constant, called the market price of risk. From the Girsanov theorem it follows that there exists a probability $\mathbb{Q} = \mathbb{Q}_\theta$ such that W is a Brownian motion under this probability. A direct computation shows that

$$F^{\mathbb{Q}_\theta}(t,T) = X(t)\, e^{-\alpha(T-t)} + \frac{\theta}{\alpha}\big(1 - e^{-\alpha(T-t)}\big) + \sigma \int_t^T e^{-\alpha(T-s)}\ \mathrm{d}W(s).$$

Therefore, the theoretical risk premium for contracts with delivery over $[T_1, T_2]$ becomes

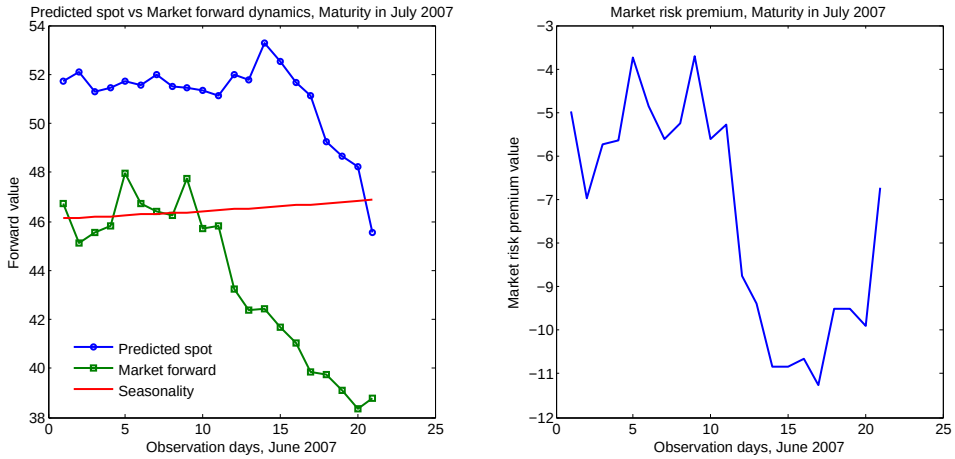
$$R(t, T_1, T_2) = \frac{\theta}{\alpha}\big(1 - \widehat{\alpha}(T_2 - T_1)\, e^{-\alpha(T_1 - t)}\big)$$

with

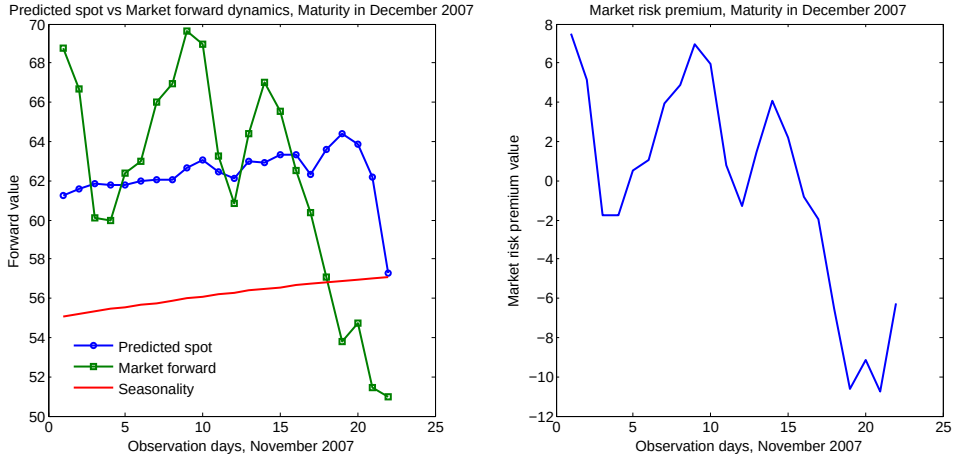
$$\widehat{\alpha}(u) = \frac{1 - e^{-\alpha u}}{\alpha u}.$$

Hence, as $T_1 - t$ increases, the risk premium increases to θ/α. On the other hand, as we approach start of delivery, i.e., $T_1 - t \to 0$, the risk premium decreases towards $\theta(1 - \widehat{\alpha}(T_2 - T_1))/\alpha$. Inspecting Figure 2.19, this is exactly the behaviour we observe in the case of the July and December 2007 contracts for the factor model. In the jump diffusion and threshold models we do not observe a similar pattern, but a much more complex structure of the pricing measure $\mathbb{Q}$. Admittedly, the factor model does not have any Brownian component, but we can do the exact same analysis for Lévy driven Ornstein-Uhlenbeck processes using the Esscher transform (see Benth *et al.* [2008a]). Concerning the two other spot models, the mean-reversion feature will create a similar behaviour when changing measure using a constant market price of risk, although the models are formulated on an exponential form. From the figures, we observe a risk premium contradicting this change of measure. In our opinion, it is a sign of quality that a model allows for an easy explanation of the risk premium, in this case a simple constant change, explaining the market price of risk easily. Note that the market price of risk will be positive for the factor model in the cases we discuss. The July 2008 contract has also a decreasing risk premium towards start of delivery, but a much more complex nature before and does not allow for this simple explanation.

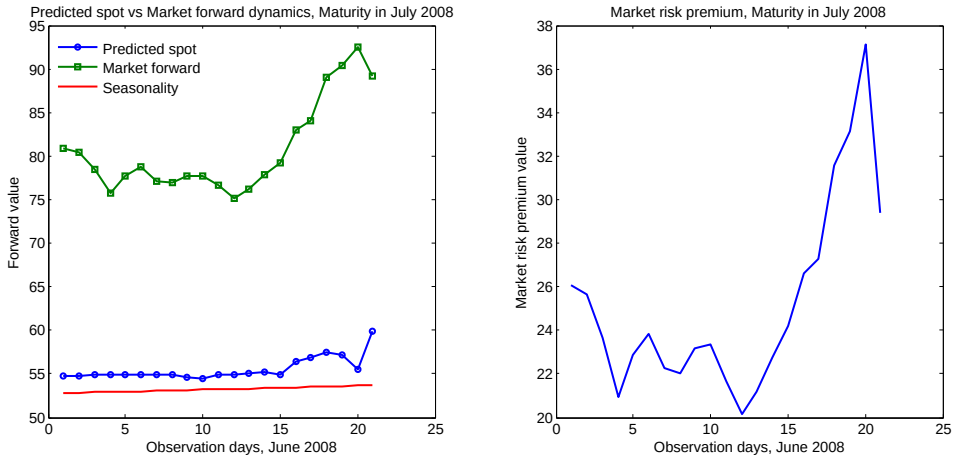
In Figure 2.20 we plot the descriptive statistics of the log-returns of the predicted spot prices for the observed periods. One can see that the jump-diffusion and threshold model produce similar results for the returns that match the observed EEX forward returns quite well. We can also find that the factor model gives the predicted spot dynamics whose mean of the log-returns is larger than the observed. This can be explained by the fact that several days before the maturity starts the predicted spot price becomes more sensitive and starts increasing to capture the volatility risk. We observe similar but smaller effect for the jump-diffusion model. The second moment of the observed EEX forwards is more in line with the jump-diffusion and factor models compared to the threshold model, since the latter produces noisier price dynamics. Also, as earlier stated, the market prices look more volatile than the predicted spots. However, this is not the case when considering all the contracts at hand, where we in fact see more variations in predicted spots. This is observed in the data from the early years 2004 and 2005.



(a) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2007.

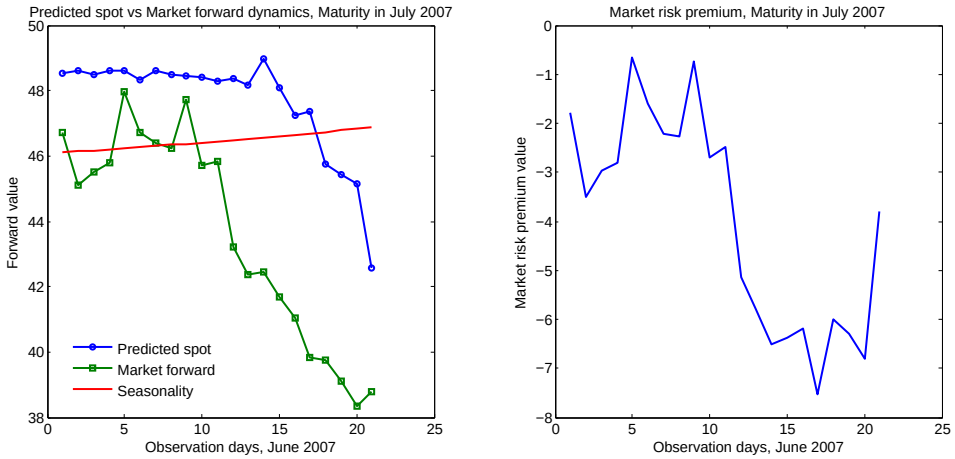


(b) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2007.

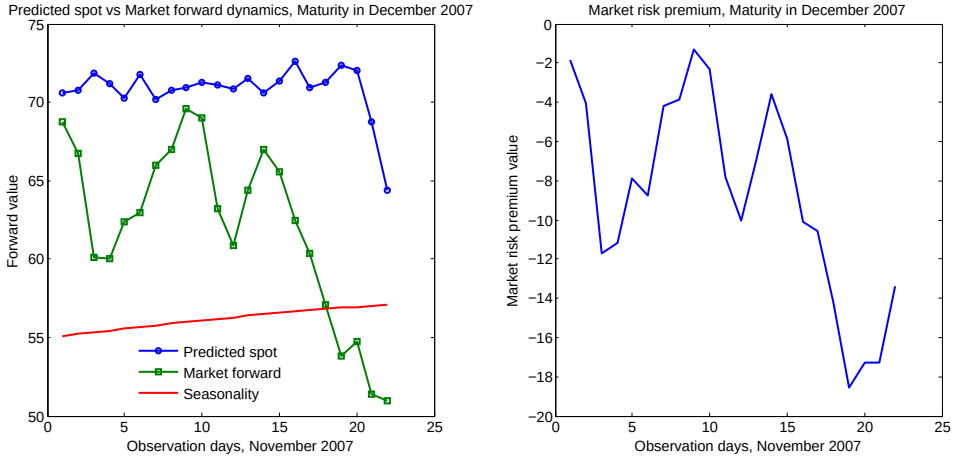


(c) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2008.

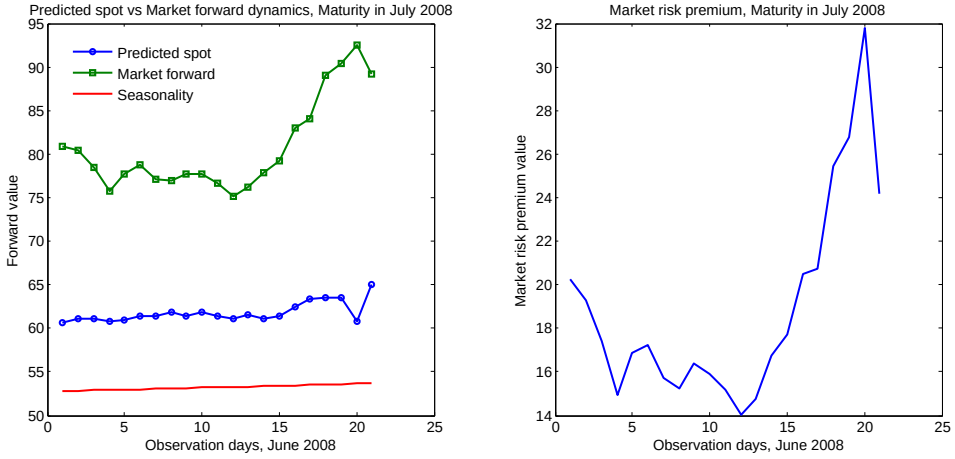
Figure 2.17: The predicted spot, observed forward dynamics and market risk premium for the jump-diffusion model.



(a) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2007.

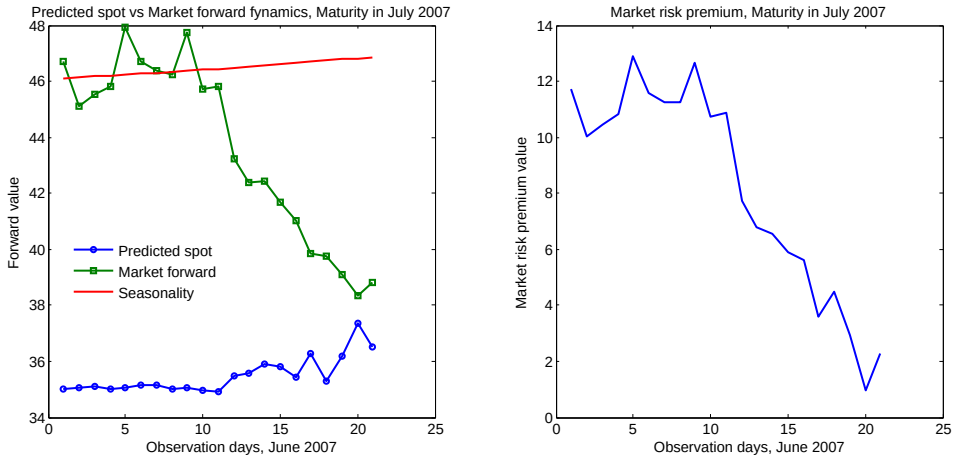


(b) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2007.

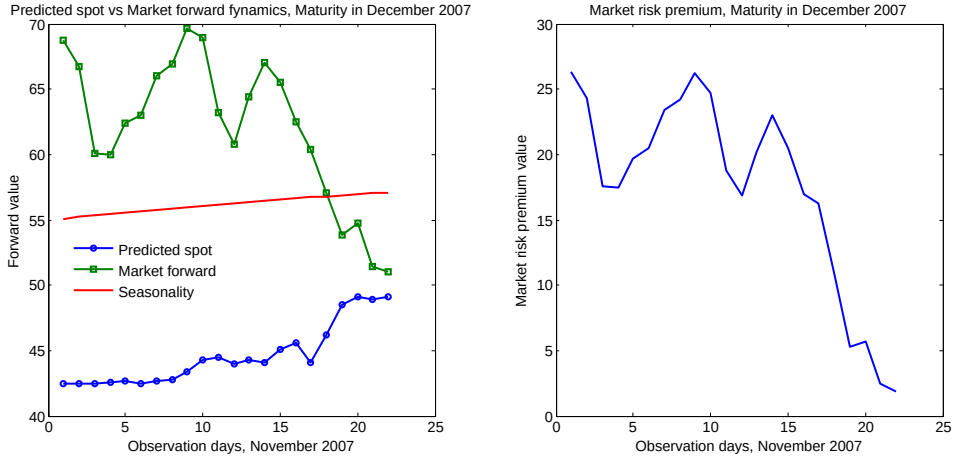


(c) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2008.

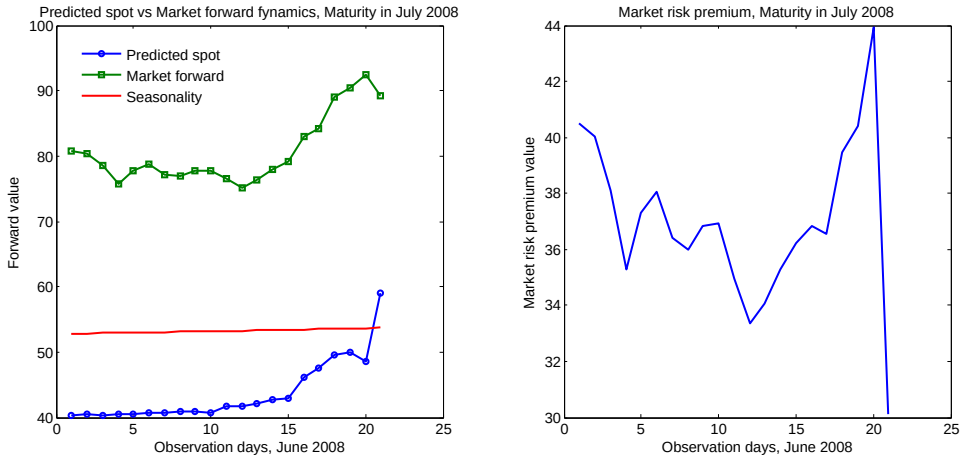
Figure 2.18: The predicted spot, observed forward dynamics and market risk premium for the threshold model.



(a) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2007.

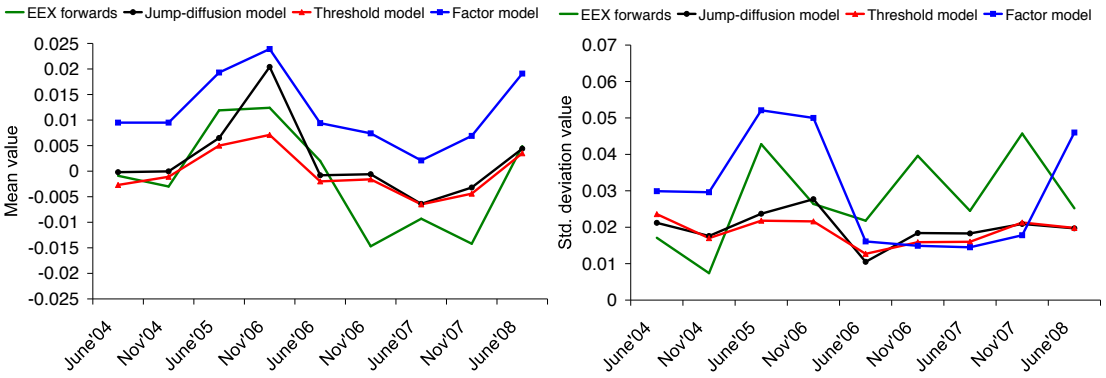


(b) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2007.



(c) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2008.

Figure 2.19: The predicted spot, observed forward dynamics and market risk premium for the factor model.



(a) Mean of the log-returns.

(b) Standard deviation of the log-returns.

Figure 2.20: Comparative descriptive statistics results for the log-returns of the jump-diffusion, threshold and factor models.

2.8 Conclusion

We have analysed and discussed the empirical performance of three continuous-time electricity spot price models that have received considerable attention from academics and practitioners recently. The mean-reversion parameter both for the jump-diffusion and the threshold models is not able to distinguish between spikes and base signal leading to a too slow mean-reversion for the spikes and a too fast mean-reversion for the base signal. For the base signal the models try to compensate for this by a very high volatility. So, the pathwise properties of the EEX price dynamics are not captured well by the jump-diffusion and the threshold models. We find that the factor model captures the fast mean-reversion of spikes and the slow mean-reversion of the base signal very well. It therefore allows for an excellent modelling of the path behaviour of the mean of the prices. However, the variability of the paths are not captured appropriately. The factor model underestimates the noise in the base signal, a fact that we attribute to the choice of an OU process with a subordinator. Such a selection produces too little variation and thus leads to an underestimation of the standard deviation of the base signal.

One further comment on the performance of the models is their analytical tractability, i.e. for pricing power derivatives. Here the factor and the jump-diffusion models are advantageous. In the case of the jump-diffusion model, by assuming that the jumps J are drawn from a Normal distribution and by requiring that $\mathbb{E}[J] = 1$ we are able to derive the forward price in closed form. If we switch to another jump-size distribution, then the model will lose its analytical tractability and we have to search for a numerical solution. However, in case of the factor model one may explicitly calculate all the probabilistic properties of the prices in terms of characteristic functions. Furthermore, due to its additive linear structure, electricity forward contract prices are obtained analytically. Electricity forward contracts have the distinctive feature of delivering the underlying commodity, spot power, over a period of time rather than at a fixed time. This implies that the price is defined as a conditional expectation (possibly risk-adjusted) of the

integral of the spot. Under the factor model an explicit calculation of this condition expectation is feasible, and we obtain the implied forward price dynamics. Thus the factor model allows one to study price determination in the forward market and risk premia; see Benth *et al.* [2007] for details. Due to the state-dependent sign of the spike process in the threshold model, we cannot obtain analytical expressions for the characteristic function of the prices process or calculate forward prices explicitly (even with fixed maturity and no delivery period). Although efficient numerical and simulation-based Monte Carlo evaluations are available, we consider the lack of analytical treatability of the threshold model as a major drawback in valuation and risk-management applications.

CHAPTER 3

PRICING POWER FORWARDS IN A REGIME-SWITCHING MODEL WITH AN INTEGRO-PDE METHOD

3.1 Motivation and introduction

In Geman & Roncoroni [2006] a Markov model for the spot price dynamics of electricity is proposed. This model is a standard Ornstein-Uhlenbeck process, with a non-standard jump term. The jumps are designed to model the spikes frequently observed in electricity spot prices.

This threshold model is a Markov jump-diffusion model, however, not feasible for explicit pricing of forward contracts due to its specification. A forward contract is an agreement where the buyer purchases any specified commodity at an agreed time to an agreed price. The agreed price, commonly known as the *forward price* is the price such that the current value of the contract is zero. With a pricing measure it can be viewed as the best predicted spot time at time of the transaction. Mathematically it is expressed as the conditional expected spot price (possibly under a risk-neutral probability).

We focus in this chapter on the derivation of the forward price dynamics. Unlike many other models, the threshold model does not allow for explicit calculation of the forward price, and numerical methods are called for. Since we want to find the dynamics, the Monte Carlo method is very cumbersome, and we are going to analyse a PDE-based approach. Since the threshold model involves jumps, we are led to integro-PDEs, and numerical methods for such. We want to compare the resulting forward prices with those of similar models.

The chapter is organised as follows. Section 3.2 introduces mathematical preliminaries and gives an overview of the models we use to compute forward prices. Next, in Section 3.3, we explain the finite difference scheme method used to solve the integro-PDE. In this section we also provide estimates for the truncation errors of the domain and of the integral term. Then, in

Section 3.4, we provide results and compare forward prices for the considered models. Finally, Section 3.5 concludes and gives an outlook for further research.

3.2 Mathematical formulation

3.2.1 Electricity price modelling

Background on Lévy processes and useful theorems is given in Appendix A.1. As in Chapter 2, we consider a Lévy-type process for modelling the power price. Let $(\Omega, \mathbb{P}, \mathcal{F}, \{\mathcal{F}_t\})$ be a complete filtered probability space. Let T defines the time horizon and $S(t)$ be the spot price of electricity (power) defined as

$$S(t) = \exp(\mu(t) + Y(t)), \tag{3.1}$$

where $\mu(t)$ is some deterministic seasonality function and $Y(t)$ is some stochastic process. The classical jump-diffusion approach to model the process $Y(t)$, as for example considered in Cartea & Figueroa [2005], is stated as

$$dY(t) = -\alpha Y(t)\ \mathrm{d}t + \sigma\ \mathrm{d}W_t + \mathrm{d}Q_t, \tag{3.2}$$

where W_t is a Brownian motion, Q_t is a compound Poisson process, α is the speed of mean-reversion and σ is the volatility. The processes W_t and Q_t are assumed to be mutually independent.

Alternatively, Geman & Roncoroni [2006] propose to model this stochastic component differently

$$dX(t) = -\alpha X(t)\ \mathrm{d}t + \sigma\ \mathrm{d}W_t + h(X(t))\ \mathrm{d}Q_t, \tag{3.3}$$

where all the notations and assumptions are the same as for the process $Y(t)$, and $h(x)$ is a state-dependent function which is -1 for large values of X (defined by some threshold $\mathfrak{T}$) and 1 otherwise. Despite the "regime-switching" term $h(x)$, this process holds the Markov property in a single state variable, for a proof see Roncoroni [2002]. The authors claim that the process X_t is a special semimartingale. This model is referred here as the threshold model.

The difference of the two models lies in the change of sign of the h-function. This function ensures that the price may jump downwards in the case of high spot prices. Note that when we

use a minus in front of the jumps in the jump-diffusion model, we want to use it in the "high price" regime for the threshold model. By the "high price" regime we mean here that the price is far above its mean level, which can happen when a spike or big jump occurred.

We also notice here that the process $L_t := \sigma W_t + Q_t$ is a Lévy process in a contrast to the process $L_t^h := \sigma W_t + h(L_t^h) Q_t$ which does not satisfy the properties given in Definition 4.

3.2.2 Forward modelling

We know that the forward price $F(t,T)$ at time t, for a contract with a delivery at time $T \geq t$, is

$$F(t,T) = \mathbb{E}^{\mathbb{Q}}[S(T) \,|\, \mathcal{F}_t], \tag{3.4}$$

which is a martingale under an equivalent martingale measure $\mathbb{Q}$. So we call the measure $\mathbb{Q}$ a pricing measure, as it is a probability that takes into account all the risks associated with the change in the price (spikes can, for example, happen due to sudden weather change or unexpected outage of equipment). The choice of $\mathbb{Q}$ can be done via a canonical Girsanov (drift part) and Esscher (jump part) transformations. Alternatively, one could say that the process S_t is already under the measure $\mathbb{Q}$ and one could argue that the market will charge an additional risk premium by changing/adjusting the mean level (for the details see Benth *et al.* [2012]. This would mean the $\mathbb{Q} = \mathbb{P}$ with the latter being the real-world pricing measure. From now on we apply this assumption.

Denote by $f(t,x) := \mathbb{E}\,[e^{X(T)} \,|\, X(t) = x]$. By the Markovian property we can write the forward price explicitly as a function of $X(t)$ as

$$F(t,T,X(t)) = e^{\mu(T)}\, \mathbb{E}\,[e^{X(T)} \,|\, X(t)] = e^{\mu(T)}\, f(t,x).$$

Our aim here is to derive efficient routines to calculate the function $f(t,x)$ based on the associated integro-PDE in terms of the threshold model and to study the impact of the function $h(x)$ on the forward prices. We will also compare obtained forwards with ones from the jump-diffusion model given in Equation (3.2).

3.2.3 Forward price of the classical jump-diffusion model

This model allows for explicit forward price formula for some class of distributions assumed for a jump size. Denote by $g(t,y) := \mathbb{E}\,[e^{Y(T)} \,|\, Y(t) = y]$. Then the forward price for the jump-diffusion model is

$$G(t,T,Y(t)) = e^{\mu(T)}\,\mathbb{E}\left[e^{Y(T)}\,|\,Y(t)\right] = e^{\mu(T)}\,g(t,y), \tag{3.5}$$

One can calculate the function $g(t,y)$ analytically by appealing to the moment generating function of the compound Poisson process Q_t. We start with the dynamics of the logarithm of the price S_t

$$\begin{aligned} \mathrm{d}\ln S_t &= \mu'(t)\,\mathrm{d}t - \alpha\big(\ln S_t - \mu(t)\big)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t + \mathrm{d}Q_t \\ &= \alpha\big(\hat{\mu}(t) - \ln S_t\big)\,\mathrm{d}t + \sigma\,\mathrm{d}W_t + \mathrm{d}Q_t, \end{aligned} \tag{3.6}$$

where $\hat{\mu}(t) := \frac{1}{\alpha}\mu'(t) + \mu(t)$. Let us now apply Ito's lemma to $(e^{\alpha t}\,\ln S_t)$ to obtain

$$\mathrm{d}(e^{\alpha t}\ln S_t) = \alpha e^{\alpha t}\hat{\mu}(t)\,\mathrm{d}t + e^{\alpha t}\sigma\,\mathrm{d}W_t + e^{\alpha t}\,\mathrm{d}Q_t. \tag{3.7}$$

After integrating from t to T and replacing terms, we have

$$\ln S_T = \mu(T) + Y(t)e^{-\alpha(T-t)} + \sigma\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}W_s + \int_t^T e^{-\alpha(T-s)}\,\mathrm{d}Q_s, \tag{3.8}$$

then the price S_T becomes

$$S_T = e^{\mu(T)+Y(T)} = e^{\mu(T)+Y(t)e^{-\alpha(T-t)}+\sigma\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}W_s+\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}Q_s}. \tag{3.9}$$

Now coming back to Equation (3.5) we have that function $g(t,y)$ becomes

$$\begin{aligned} g(t,y) &:= \mathbb{E}\left[\exp(Y(T))\,|\,Y(t)=y\right] \\ &= \mathbb{E}\left[e^{Y(t)e^{-\alpha(T-t)}}\,e^{\sigma\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}W_s}\,e^{\int_t^T Je^{-\alpha(T-s)}\,\mathrm{d}N_s}\,|\,Y(t)=y\right] \\ &= e^{ye^{-\alpha(T-t)}}\exp\Big(\frac{\sigma^2}{4\alpha}\big(1-e^{-2\alpha(T-t)}\big)\Big)\exp\Big(\lambda\int_t^T(\mathbb{E}[e^{Je^{-\alpha(T-s)}}]-1)\,\mathrm{d}s\Big), \end{aligned} \tag{3.10}$$

where for the last equality we use mutual independence of W_t and Q_t, the fact the $\sigma\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}W_s$ is a normally distributed random variable and the Lévy-Khinchin representation for the compound Poisson process Q_t.

Normal distribution for jumps

Now following Cartea & Figueroa [2005], we assume a Normal distribution for the jump size J, i.e. $J \sim \mathcal{N}(m_1, m_2)$ with mean m_1 and standard deviation m_2. This allows us to compute

$$\mathbb{E}[e^{Je^{-\alpha(T-s)}}] = \exp\left(m_1 e^{-\alpha(T-s)} + \frac{m_2^2}{2} e^{-2\alpha(T-s)}\right). \tag{3.11}$$

Then the forward price $g(t, y)$ when the jump size J follows Normal distribution is given as

$$g(t,y) = e^{ye^{-\alpha(T-t)}} \exp\left(\frac{\sigma^2}{4\alpha}\left(1 - e^{-2\alpha(T-t)}\right)\right) \exp\left(\lambda \int_t^T e^{m_1 e^{-\alpha(T-s)} + \frac{m_2^2}{2} e^{-2\alpha(T-s)}} \,\mathrm{d}s - \lambda(T-t)\right). \tag{3.12}$$

Laplace distribution for jumps

Another option for the jump size distribution is a Laplace distribution. Besides explanatory advantages such as capturing the heavy tails spike nature, we have the following useful property: when $J \sim \mathcal{L}aplace(m_1, m_2)$ with m_1 – the location parameter and $m_2 > 0$ – the scale parameter, then $e^{cJ} \sim \mathcal{L}og\mathcal{L}aplace(cm_1, cm_2)$ with some constant c. A very detailed investigation of the LogLaplace distribution can be found in a book of Kozubowski & Podgorski [2003]. We can use their formula for the expected value of a random variable e^{cJ} and obtain

$$\mathbb{E}[e^{cJ}] = \frac{\delta}{1 - c^2 m_2^2}, \tag{3.13}$$

where $\delta := e^{cm_1}$. Then this allows us to compute the expected value

$$\mathbb{E}[e^{Je^{-\alpha(T-s)}}] = \frac{\exp(e^{-\alpha(T-s)} m_1)}{1 - e^{-2\alpha(T-s)} m_2^2}. \tag{3.14}$$

So the the forward price $g(t, y)$ when the jump size J follows Laplace distribution is given as

$$g(t,y) = e^{ye^{-\alpha(T-t)}} \exp\left(\frac{\sigma^2}{4\alpha}\left(1 - e^{-2\alpha(T-t)}\right)\right) \exp\left(\lambda \int_t^T \left(\frac{\exp(e^{-\alpha(T-s)} m_1)}{1 - e^{-2\alpha(T-s)} m_2^2}\right) \mathrm{d}s - \lambda(T-t)\right). \tag{3.15}$$

3.2.4 Forward price of the threshold model

Suppose $X(t)$ solves SDE in Equation (3.3). Let

$$f(t,x) := \mathbb{E}\left[e^{X(T)} \mid X(t) = x\right] = \mathbb{E}\left[\Phi(X(T)) \mid X(t) = x\right] \tag{3.16}$$

be the expected value of the payoff $\Phi(x) = e^x$ at maturity time $T > t$, given that $X(t) = x$. Then f solves

$$f_t + \mathcal{L}f = 0, \quad \text{for} \quad t < T, \quad \text{with} \quad f(T,x) = \Phi(x) \quad \text{at} \quad t = T, \tag{3.17}$$

where $\mathcal{L}$ is the generator of the process $X(t)$ given as

$$\mathcal{L}f = -\alpha x f_x + \frac{\sigma^2}{2} f_{xx} + \lambda h(x) \mathbb{E}\Big[f(x+J,t) - f(x,t)\Big]. \tag{3.18}$$

The proof is the following: for any function $\phi(t,x)$, Ito's formula is

$$\begin{aligned}
& \mathrm{d}\phi(s, X(s)) \\
= \;& \sigma\phi_x(s,X(s))\,\mathrm{d}W_s + \big(\phi_s(s,X(s)) - \alpha x\phi_x(s,X(s)) + \frac{\sigma^2}{2}\phi_{xx}(s,X(s))\big)\,\mathrm{d}s \\
+ \;& h(x)\left[\phi(s,X(s)+J) - \phi(s,X(s))\right]\,\mathrm{d}N_s \\
= \;& \sigma\phi_x(s,X(s))\,\mathrm{d}W_s + \big(\phi_s(s,X(s)) - \alpha x\phi_x(s,X(s)) + \frac{\sigma^2}{2}\phi_{xx}(s,X(s))\big)\,\mathrm{d}s \\
+ \;& h(x)\left[\phi(s,X(s)+J) - \phi(s,X(s))\right]\,\Big(\mathrm{d}\tilde{N}_s + \lambda\,\mathrm{d}s\Big) \\
= \;& \sigma\phi_x(s,X(s))\,\mathrm{d}W_s + h(x)\left[\phi(s,X(s)+J) - \phi(s,X(s))\right]\,\mathrm{d}\tilde{N}_s \\
+ \;& \big(\phi_s(s,X(s)) - \alpha x\phi_x(s,X(s)) + \frac{\sigma^2}{2}\phi_{xx}(s,X(s)) + \lambda\, h(x)\left[\phi(s,X(s)+J) - \phi(s,X(s))\right]\big)\,\mathrm{d}s,
\end{aligned} \tag{3.19}$$

where we used the fact that the compensated Poisson process $\tilde{N}_t = N_t - \lambda t$ and $\tilde{N}_t$ is a martingale. Now when choosing $\phi = f$, the solution of Equation (3.17), we have

$$\begin{aligned}
f(T,X(T)) - f(t,x) \;=\; & \int_t^T (\sigma f_x)(s,X(s))\,\mathrm{d}W_s + \int_t^T h(X(s))\left[f(s,X(s)+J) - f(s,X(s))\right]\mathrm{d}\tilde{N}_s \\
+ \; & \int_t^T \Big\{\big(f_s - \alpha x f_x + \frac{\sigma^2}{2} f_{xx}\big)(s,X(s)) + \lambda\, h(X(s))\left[f(s,X(s)+J) - f(s,X(s))\right]\Big\}\,\mathrm{d}s
\end{aligned}$$

Now taking the expected values from both sides, dividing by $T-t$ and using the fact that $\tilde{N}_t$ is a martingale yield

$$\begin{aligned} & \frac{\mathbb{E}[f(T,X(T))]-f(t,x)}{T-t} \\ = \quad & \frac{\mathbb{E}\Big[\int_t^T \Big\{\big(f_s-\alpha x f_x+\frac{\sigma^2}{2}f_{xx}\big)(s,X(s))+\lambda\, h(X(s))\,\big(f(s,X(s)+J)-f(s,X(s))\big)\Big\}\,\mathrm{d}s\Big]}{T-t}, \end{aligned}$$

then taking the limit with $T-t\to 0$ and using

$$\mathbb{E}[f(T,X(T))]-f(t,x)=0$$

gives

$$0=f_t-\alpha x f_x+\frac{\sigma^2}{2}f_{xx}+\lambda h(x)\mathbb{E}[f(t,x+J)-f(t,x)], \tag{3.20}$$

which is exactly our claim.

In the next section we will implement the numerical scheme that solves this partial differential equation with an integral term to investigate the property of the function $h(x)$ on the forward price. Before to continue with the numerical investigation, we do a heuristic calculation for illustration. Let us approximate the term of the expected jump in Equation (3.18) as with a 1-step Taylor expansion

$$\begin{aligned} \mathbb{E}[f(t,x+J)-f(t,x)] &= \int_{-\infty}^{\infty}\big(f(t,x+y)-f(t,x)\big)\,f_J(y)\,\mathrm{d}y \\ &\approxeq \int_{-\infty}^{\infty}\frac{\big(f(t,x+y)-f(t,x)\big)}{\mathrm{d}x}\,y\,f_J(y)\,\mathrm{d}y \\ &= f_x\underbrace{\int_{-\infty}^{\infty} y\,f_J(y)\,\mathrm{d}y}_{\approxeq C}. \end{aligned} \tag{3.21}$$

with $f_J(y)$ is a probability density function of the random jump size J and $C:=\mathbb{E}(J)$. We mean here that the integral part in this equation behaves like a gradient term, approximately giving a rise to a second order differential operator as the right-hand side of Equation (3.22). So then Equation (3.20) can be re-written as

$$f_t\approxeq(-\alpha x+\lambda h(x)C)\,f_x+\frac{\sigma^2}{2}f_{xx}, \tag{3.22}$$

which clearly indicates that the term in front of f_x is of discontinuous nature. It also points out the curvature due to $h(x)$ function. This "discontinuous" curvature can clearly be seen in Figure 3.1.

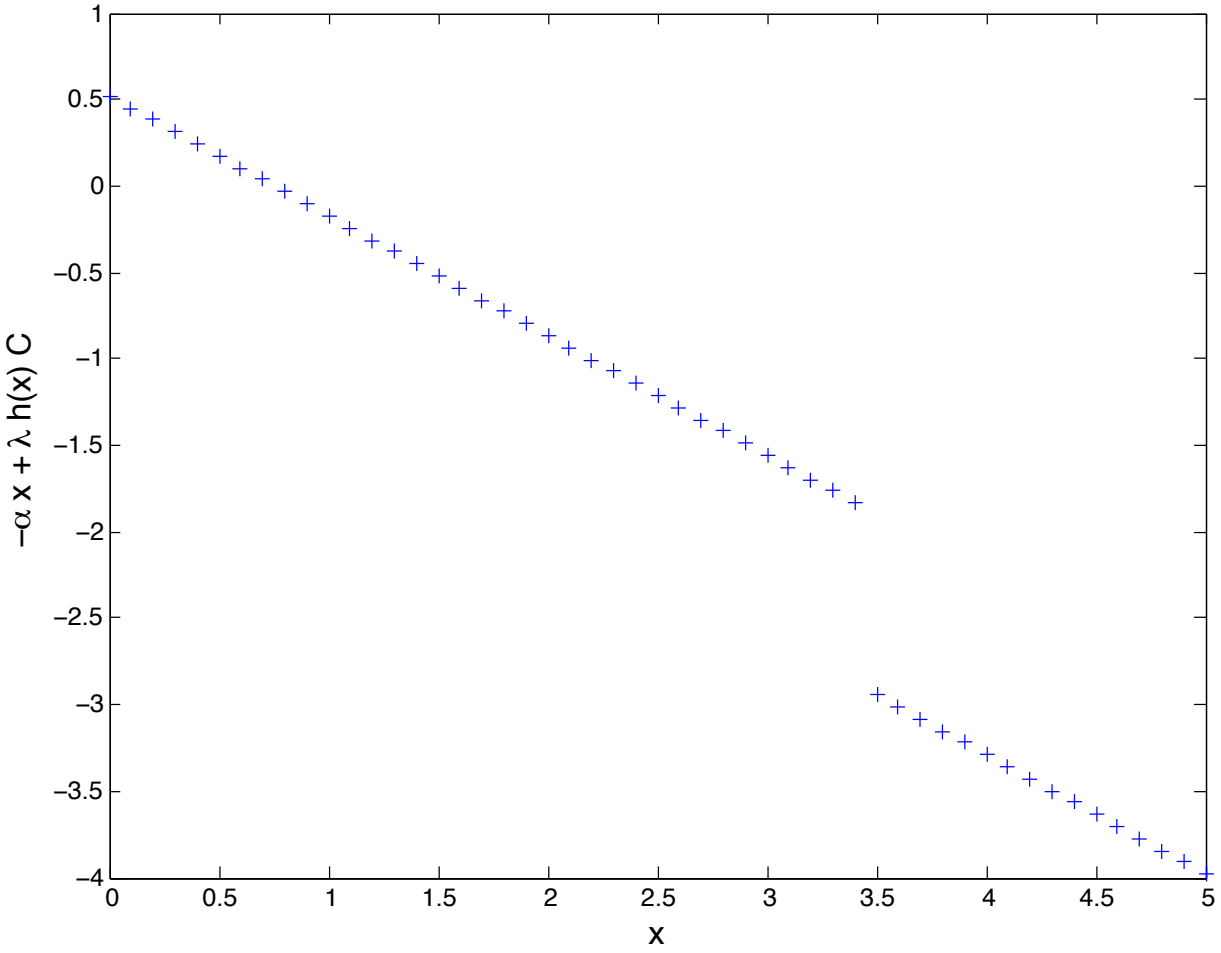


Figure 3.1: Discontinuity due to the function $h(x)$. Parameters: $C = 10$, $\alpha = 0.69$ (more than half a day), $\sigma = 2.59$, $\lambda = 13.5$ spikes per year, $\mathcal{T} = 3.5$.

3.3 Numerical implementation

This section consists of three parts. Firstly, we discuss the method of finite differences which we apply to solve the integro-PDE given above to obtain forward prices. When solving this equation we have to restrict our domain for the values x. It results in the truncation error which we discuss in the the second part of this section. And finally, we have to cut the integral term and investigate this truncation error as well.

3.3.1 Method

In this section we implement the finite difference method to solve equation (3.17). Excellent overviews of numerical methods in application to finance are given in the books of Cont & Tankov [2004] and Fusai & Roncoroni [2008].

This scheme is an approximation for this equation and is based on replacing derivatives by finite differences in the equation. The integral term responsible for the jumps is approximated by the Riemann sums. Since we have time and space variables we have to discretise the time and spacial domains.

Let us first rewrite Equation (3.17) for the forward price $f(t,x)$ we are going to numerically solve

$$f_t = \alpha x f_x - \frac{\sigma^2}{2} f_{xx} + \lambda h(x) f - \lambda h(x) \int_{-\infty}^{\infty} f(t, x+y)\, f_Y(y)\, \mathrm{d}y, \quad f(T,x) = e^x. \tag{3.23}$$

Then let us make the following replacement: $T - t = \tau$ is the time to maturity, which allows us to move backward in time when solving the equation numerically. Then the integro-PDE becomes

$$f_\tau = -\alpha x f_x + \frac{\sigma^2}{2} f_{xx} - \lambda h(x) f + \lambda h(x) \int_{-\infty}^{\infty} f(T-\tau, x+y)\, f_Y(y)\, \mathrm{d}y, \quad f(\tau,x) = e^x. \tag{3.24}$$

Since there is enormous variety of sources on the finite difference method precisely applied to solving financial mathematics problems, we will not focus on the details. Instead, we provide the exact scheme we use here to solve Equation (3.24).

We start with a time domain for $\tau \in [0,T]$ and discretise it with $\Delta\tau = \frac{\tau}{N}$ with N being the number of time steps. Then we continue with a space domain $[x_{min}, x_{max}]$ and $\Delta x = \frac{x_{max} - x_{min}}{M}$ with M being the number of space steps. We define $x_i = x_{min} + (i-1)\Delta x$ and $\tau_n = N\Delta\tau$ with $n = 1, \ldots, N+1$ and $i = 1, \ldots, M+1$. We also introduce some values $K_{min} < 0$ and $K_{max} > 0$ responsible for the interval for the jump size given in the integral term of the equation. Let $\{f_i^n\}$ be the solution on our discretised grid on the interval $[x_{min}, x_{max}]$. The time and space derivatives become

$$\frac{\partial f}{\partial \tau} \approxeq \frac{f^{n+1} - f^n}{\Delta \tau}, \tag{3.25}$$

$$\frac{\partial f}{\partial x} \approxeq \frac{f_{i+1}^{n+1} - f_i^{n+1}}{\Delta x}, \tag{3.26}$$

$$\frac{\partial^2 f}{\partial x^2} \approxeq \frac{f_{i+1}^{n+1} - 2f_i^{n+1} + f_{i-1}^{n+1}}{(\Delta x)^2}. \tag{3.27}$$

Here we use a so-called explicit scheme which allows to find the value at time $n+1$ knowing the value at time n. Then we approximate the integral term. We do this via trapezoidal quadrature rule with the same grid resolution Δx. As it is stated in Cont & Tankov [2004], due to computational complexity when there is a jump term, it is more convenient to use an implicit scheme for the integral part which allows to find the value at time $n+1$ knowing the value at n, namely

$$\begin{aligned}
\int_{-\infty}^{\infty} f(T-\tau, x_i + y)\, f_Y(y)\, \mathrm{d}y &\approxeq \lim_{\substack{K_{min}\to-\infty \\ K_{max}\to\infty}} \int_{K_{min}}^{K_{max}} f(T-\tau, x_i+y)\, f_Y(y)\ \mathrm{d}y \\
&\approxeq \sum_{j=K_l}^{K_u} f^n(x_i + j\Delta x) \int_{(j-1/2)\Delta x}^{(j+1/2)\Delta x} f_Y(y)\, \mathrm{d}y \\
&= \sum_{j=K_l}^{K_u} f_{i+j}^n \int_{(j-1/2)\Delta x}^{(j+1/2)\Delta x} f_Y(y)\, \mathrm{d}y,
\end{aligned} \tag{3.28}$$

where $[K_{min}, K_{max}] \in [(K_l - 1/2)\Delta x, (K_l + 1/2)\Delta x]$.

So, the total explicit-implicit scheme then becomes

$$\frac{f^{n+1} - f^n}{\Delta t} = Df^{n+1} + Jf^n, \tag{3.29}$$

where

$$(Df^{n+1})_i = \alpha x_i \frac{f_{i+1}^{n+1} - f_i^{n+1}}{\Delta x} - \frac{\sigma^2}{2} \frac{f_{i+1}^{n+1} - 2f_i^{n+1} + f_{i-1}^{n+1}}{(\Delta x)^2}, \tag{3.30}$$

$$(Jf^n)_i = -\lambda h(x_i) \sum_{j=K_l}^{K_u} f_{i+j}^n \int_{(j-1/2)\Delta x}^{(j+1/2)\Delta x} f_Y(y)\, \mathrm{d}y. \tag{3.31}$$

Equation (3.29) can be rewritten as

$$(I - \Delta\tau D)f^{n+1} = (I + \Delta\tau J)f^n, \tag{3.32}$$

where I is an identity matrix, $(I-\Delta\tau D)$ and $(I+\Delta\tau J)$ are tridiagonal matrices. Since we know the final condition, i.e. $f(\tau, x) = e^x$ we move backward in time when solving this scheme, i.e. knowing the value of f^n we search for the value of f^{n+1}.

As the reader can see, there are two things we should agree on when solving this equation: when $i = 1$ we need to know the term f_{i-1}^{n+1} which is out of the domain of x. When $i = M + 1$ we need to know the term f_{i+1}^{n+1} which is also out of the domain of x. There are several possible solutions to this obstacle. One of them is to represent the second derivative differently as

$$\frac{f_{i-2}^{n+1} - 2f_{i-1}^{n+1} + f_i^{n+1}}{(\Delta x)^2}.$$

Another optionality is to assume that our final condition at $\tau = 0$ is extended not only for the domain (x_{min}, x_{max}) but for all the values x we need out of the domain.

However, in our case we can go for the third option and find these values explicitly since we have an exact solution for the forward price for the jump-diffusion model considered above, at least for some distributions. We mean here the following: when $x > x_{max}$ it implies that the function $h(x) = -1$, which implies that the function $f(t, x)$ is "almost" equal to $g(t, y)$ for $x = y$: the difference is the minus sign in front of the jump component in Equation (3.2). In the next section we show that if the jump size distribution is symmetric, then $f(t, x)$ is equal to $g(t, y)$ for $x = y$. When $x < x_{min}$ it implies that the function $h(x) = 1$, which gives forward price $f(t, x)$ is exactly equal to $g(t, y)$.

We also need to stress the issues linked to the consistence with the continuous equation and stability (i.e. our obtained solution does not blow up when $\Delta\tau$ and Δx goes to 0). The choice of those should be done carefully. The stability condition stated in Cont & Tankov [2004] and used here is $\Delta\tau \leq \inf\{\frac{1}{\lambda}, \frac{(\Delta x)^2}{\sigma^2}\}$.

3.3.2 Domain truncation error

When we solve numerically the partial differential equation with an integral part, we have to define the domain for the x values. Defining this interval for $x \in (x_{min}, x_{max})$ means that we have to specify some boundary conditions at $x = x_{min}$ and $x = x_{max}$. On top of that we have specify the boundary values for our function for the integral term. In our case we can specify these boundary values explicitly. When $x \leq x_{min}$, then the function $h(x) = 1$ and then the forward price $f(t, x) = g(t, y)$ for $x = y$. When $x \geq x_{max}$, the function $h(x) = -1$, then the forward price $f(t, x) = g^-(t, y)$, i.e. with a minus sign in front of the jump component in Equation (3.2), namely

$$dY^{-}(t) = -\alpha Y(t)\ \mathrm{d}t + \sigma\ \mathrm{d}W_t - \ \mathrm{d}Q_t. \tag{3.33}$$

Let us now compute the forward price $g^{-}(t,y)$ associated with this process $Y^{-}(t)$ assuming that the starting value y is the same as for the function $g(t,y)$

$$g^{-}(t,y) = e^{ye^{-\alpha(T-t)}} \exp\Big(\frac{\sigma^2}{4\alpha}\big(1-e^{-2\alpha(T-t)}\big)\Big) \exp\Big(\lambda\int_t^T (\mathbb{E}[e^{-Je^{-\alpha(T-s)}}]-1)\,\mathrm{d}s\Big). \tag{3.34}$$

We can see that the only difference between $g(t,y)$ and $g^{-}(t,y)$ lies in the expected values of the exponent of the jump component, let us compare these terms keeping in mind that $e^{-\alpha(T-s)} \in (0,1]$ for $\alpha > 0$ and $T-s \geq 0$

$$\begin{aligned}
& \mathbb{E}[e^{Je^{-\alpha(T-s)}} - e^{-Je^{-\alpha(T-s)}}] \\
= & \int_{-\infty}^{+\infty} \Big(e^{xe^{-\alpha(T-s)}} - e^{-xe^{-\alpha(T-s)}}\Big)\ f_J(x)\ \mathrm{d}x \\
= & \int_{-\infty}^{0} \Big(e^{xe^{-\alpha(T-s)}} - e^{-xe^{-\alpha(T-s)}}\Big)\ f_J(x)\ \mathrm{d}x + \int_{0}^{+\infty} \Big(e^{xe^{-\alpha(T-s)}} - e^{-xe^{-\alpha(T-s)}}\Big)\ f_J(x)\ \mathrm{d}x \\
= & \int_{0}^{+\infty} \Big(e^{-xe^{-\alpha(T-s)}} - e^{xe^{-\alpha(T-s)}}\Big)\ f_J(-x)\ \mathrm{d}x + \int_{0}^{+\infty} \Big(e^{xe^{-\alpha(T-s)}} - e^{-xe^{-\alpha(T-s)}}\Big)\ f_J(x)\ \mathrm{d}x \\
= & \begin{cases} 0, & \text{if distribution is symmetric} \\ \int_0^{+\infty} \Big(e^{xe^{-\alpha(T-s)}} - e^{-xe^{-\alpha(T-s)}}\Big)\ (f_J(x) - f_J(-x))\ \mathrm{d}x, & \text{otherwise} \end{cases} \\
= & \begin{cases} 0, & \text{if} \quad f_J(x) = f_J(-x), \quad \text{distribution is symmetric;} \\ C_{>0}, & \text{if} \quad f_J(x) > f_J(-x), \quad \text{upward jumps are more likely;} \\ C_{<0}, & \text{if} \quad f_J(x) < f_J(-x), \quad \text{downward jumps are more likely.} \end{cases}
\end{aligned} \tag{3.35}$$

So, we observe that

$$\begin{cases} g(t,y) = g^{-}(t,y), & \text{distribution is symmetric;} \\ g(t,y) > g^{-}(t,y), & \text{upward jumps are more likely;} \\ g(t,y) < g^{-}(t,y), & \text{downward jumps are more likely;} \end{cases}$$

Let us then denote the solution to our boundary problem given in Equation (3.17) for $x = y$ as

$$f_B(t,x) = \begin{cases} \hat{f}_B(t,x), & \text{if } x \in (x_{min}, x_{max}) \\ g(t,y), & \text{if } x \leq x_{min} \\ g^-(t,y), & \text{if } x \geq x_{max} \end{cases} \tag{3.36}$$

where $\hat{f}_B(t,x)$ solves Equation (3.17) on a bounded domain (x_{min}, x_{max}), solution $g(t,y)$ is given in Equation (3.15) and solution $g^-(t,y)$ is given in Equation (3.34). Now we are ready to calculate the domain truncation error with the following proposition. In the case of absence of the function $h(x)$ an analogous proposition is given in Cont & Voltchkova [2005].

Proposition 3 (Domain truncation error)**.** *Assume that*

- $\nu(\mathrm{d}x)$ *is a Lévy measure such that for* $\epsilon > 0$, $\int_{|x|>1} \left(e^{\epsilon|x|} - 1\right) \nu(\mathrm{d}x) < \infty$;
- $x_{min} = -x_{max}$.

Let $f(t,x)$ *be the solution of our problem in Equation* (3.17) *on the unbounded region and* $f_B(t,x)$ *be the solution defined in Equation* (3.36)*. Then for* $x = y$

$$|f(t,x) - f_B(t,x)| \leq 2\,C\,e^{-\epsilon\,(x_{max}-|x|)} \begin{cases} g(t,y), & \textit{if} \quad f_J(x) \geq f_J(-x), \\ g^-(t,y), & \textit{if} \quad f_J(x) < f_J(-x), \end{cases}$$

where C *is some constant which does not depend on* x_{max}*,* $f_J(\cdot)$ *is the probability density of the jump size distribution.*

Proof. For $t < T$ denote by

$$Z(T) := \sigma \int_t^T e^{-\alpha(T-s)}\,\mathrm{d}W_s + \int_t^T e^{-\alpha(T-s)} h(X(s))\,\mathrm{d}Q_s,$$

then $X(t,T) = xe^{-\alpha(T-t)} + Z(T)$. Denote also by $M_T^x := \sup_{t\leq s\leq T} |X(t,s)|$, $\underline{M}_T^x := \inf_{t\leq s\leq T} X(t,s)$ and $\overline{M}_T^x := \sup_{t\leq s\leq T} X(t,s)$. Then we have for $x = y$ and a shorthand notation $f(t,x) = f$, $g(t,y) = g$ and $g^-(t,y) = g^-$

$$\begin{aligned} & |f(t,x) - f_B(t,x)| \\ = \;& \left|\left(f - \hat{f}_B\right) \mathbf{1}_{M_T^x < x_{max}}\right| \\ + \;& \left|(f-g)\,\mathbf{1}_{\underline{M}_T^x < x_{min}}\right| + \left|\left(f - g^-\right) \mathbf{1}_{\overline{M}_T^x > x_{max}}\right| \\ \leq \;& \left|(f-g)\,\mathbf{1}_{\underline{M}_T^x < x_{min}}\right| + \left|\left(f - g^-\right) \mathbf{1}_{\overline{M}_T^x > x_{max}}\right| \end{aligned}$$

$$
\begin{aligned}
&= |f-g|\ \mathbf{1}_{\underline{M}_T^x<x_{min}} + \left|f-g^-\right|\ \mathbf{1}_{\overline{M}_T^x>x_{max}} \\
&\leq \left(\left|g-g^-\right|\ \mathbf{1}_{h(x)=1} + \left|g^- - g^-\right|\ \mathbf{1}_{h(x)=-1} + \left|g^- - g\right|\right)\ \mathbf{1}_{\underline{M}_T^x<x_{min}} \\
&+ \left(|g-g|\ \mathbf{1}_{h(x)=1} + \left|g^- - g\right|\ \mathbf{1}_{h(x)=-1} + \left|g - g^-\right|\right)\ \mathbf{1}_{\overline{M}_T^x>x_{max}} \\
&\leq 2\left|g-g^-\right|\ \left(\mathbf{1}_{\underline{M}_T^x<x_{min}} + \mathbf{1}_{\overline{M}_T^x>x_{max}}\right) \\
&= 2\left|g-g^-\right|\ \mathbb{P}\left(M_T^x \geq x_{max}\right) \\
&\leq \begin{cases} 2\,g(t,y)\,\mathbb{P}\left(M_T^x \geq x_{max}\right), & \text{if} \quad f_J(x) \geq f_J(-x), \\ 2\,g^-(t,y)\,\mathbb{P}\left(M_T^x \geq x_{max}\right), & \text{if} \quad f_J(x) < f_J(-x). \end{cases}
\end{aligned}
\tag{3.37}
$$

Since the function $h(x)$ is bounded by -1 and 1, we observe that

$$
\begin{aligned}
Z(T) &= \sigma \int_t^T e^{-\alpha(T-s)}\,\mathrm{d}W_s + \int_t^T e^{-\alpha(T-s)} h(X(s))\ \mathrm{d}Q_s \\
&\leq \sigma \int_t^T e^{-\alpha(T-s)}\,\mathrm{d}W_s + \int_t^T e^{-\alpha(T-s)}\ \mathrm{d}\tilde{Q}_s \\
&=: \tilde{Z}^T,
\end{aligned}
\tag{3.38}
$$

where $\tilde{Q}_t$ is a compound Poisson process and $\tilde{Z}^T$ is a Lévy process with a measure $\nu(\,\mathrm{d}x)$.

$$
\begin{aligned}
M_T^x &= \sup_{t\leq s\leq T}\left|xe^{-\alpha(s-t)} + Z(s)\right| \\
&\leq \sup_{t\leq s\leq T}\left|xe^{-\alpha(s-t)}\right| + \sup_{t\leq s\leq T}|Z(s)| \\
&\leq \sup_{t\leq s\leq T}\left|xe^{-\alpha(s-t)}\right| + \sup_{t\leq s\leq T}\left|\tilde{Z}(s)\right| \\
&\leq |x| + \sup_{t\leq s\leq T}\left|\tilde{Z}(s)\right| \\
&=: M_T^{\tilde{x}}.
\end{aligned}
\tag{3.39}
$$

We note that

$$
\mathbb{P}\left(M_T^x \geq x_{max}\right) \leq \mathbb{P}\left(M_T^{\tilde{x}} \geq x_{max}\right). \tag{3.40}
$$

Now with the help of Theorem 4 we find that

$$
C := \mathbb{E}e^{\epsilon\, M_T^{\tilde{0}}} < \infty. \tag{3.41}
$$

Applying further Chebyshev's inequality gives

$$\mathbb{P}\left(M_T^{\tilde{0}} \geq x_{max}\right) \leq Ce^{-\epsilon\, x_{max}}. \tag{3.42}$$

Now we find that

$$\begin{aligned}
\mathbb{P}\left(M_T^{\tilde{x}} \geq x_{max}\right) &= \mathbb{P}\left(|x| + \sup_{t\leq s\leq T}\left|\tilde{Z}(s)\right| \geq x_{max}\right) \\
&= \mathbb{P}\left(\sup_{t\leq s\leq T}\left|\tilde{Z}(s)\right| \geq x_{max} - |x|\right) \\
&= \mathbb{P}\left(M_T^{\tilde{0}} \geq x_{max} - |x|\right) \\
&\leq Ce^{-\epsilon\,(x_{max}-|x|)}.
\end{aligned} \tag{3.43}$$

□

3.3.3 Jump size domain truncation error

Here we discuss the integral term truncation error which we obtain when we cut the interval for the jump size as it was done in Equation (3.28). The process Q_t, which is responsible for the jump component in Equation (3.3), is a compound Poisson process with a jump measure $\nu(\,\mathrm{d}x)$ that measures the expected number of jumps per unit time whose size belong to a set $A \in \mathcal{B}(\mathbb{R})$.

Now let us introduce a new compound Poisson process Q_t^K with a new jump measure $\nu^K := \nu(\,\mathrm{d}x)\,1_{x\in[K_{min},K_{max}]}$. Then since the function $h(x)$ is bounded by -1 and 1 we observe the following

$$\begin{aligned}
X(T) &= xe^{-\alpha(T-t)} + \sigma\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}W_s + \int_t^T e^{-\alpha(T-s)}h(X(s))\,\mathrm{d}Q_s \\
&= xe^{-\alpha(T-t)} + \sigma\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}W_s + \sum_{t<s\leq T} e^{-\alpha(T-s)}h(X(s))\Delta Q(s) \\
&= xe^{-\alpha(T-t)} + \sigma\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}W_s + \sum_{t<s\leq T} e^{-\alpha(T-s)}h(X(s))\sum_{i=N(t)+1}^{N(s)} J_i \\
&\leq xe^{-\alpha(T-t)} + \sigma\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}W_s + \sum_{t<s\leq T} e^{-\alpha(T-s)}\sum_{i=N(t)+1}^{N(s)} |J_i|
\end{aligned}$$

$$
\begin{aligned}
&= xe^{-\alpha(T-t)} + \sigma \int_t^T e^{-\alpha(T-s)}\,\mathrm{d}W_s + \sum_{t<s\leq T} e^{-\alpha(T-s)}\Delta\tilde{Q}(s) \\
&= xe^{-\alpha(T-t)} + \sigma \int_t^T e^{-\alpha(T-s)}\,\mathrm{d}W_s + \int_t^T e^{-\alpha(T-s)}\,\mathrm{d}\tilde{Q}_s, \qquad (3.44)
\end{aligned}
$$

where in the third equality we used a definition of a compound Poisson process given in Section 11.3.1 in Shreve [2004]. Analogously, the process $X^K(T)$ (with the jump component formed by the process Q_t^K) with $X^K(t) = x$ at time t is

$$
\begin{aligned}
X^K(T) &= xe^{-\alpha(T-t)} + \sigma \int_t^T e^{-\alpha(T-s)}\,\mathrm{d}W_s + \int_t^T e^{-\alpha(T-s)}h(X^K(s))\,\mathrm{d}Q_s^K \\
&\leq xe^{-\alpha(T-t)} + \sigma \int_t^T e^{-\alpha(T-s)}\,\mathrm{d}W_s + \int_t^T e^{-\alpha(T-s)}\,\mathrm{d}\tilde{Q}_s^K, \qquad (3.45)
\end{aligned}
$$

where J^K is a random variable responsible for the jump size and that falls into the interval $I = [K_{min}, K_{max}]$.

Let us now introduce a solution $f_I(t,x)$ which solves the problem given in Equation (3.17) but with the truncated jump size domain $[K_{min}, K_{max}]$ and a solution $f(t,x)$ with an unbounded jump size domain $(-\infty, +\infty)$. We give an error estimate associated to this truncation. This error estimate derivation is similar to the proof of Proposition 4.2 in Cont & Voltchkova [2005]. The main difference here is the presence of function $h(x)$ due to which our process $X(t)$ is not a Lévy process.

Proposition 4 (Jump size domain truncation error)**.** *Assume that a jump measure $\nu(\mathrm{d}x)$ for a compound Poisson process satisfies the following*

- $\int_{\mathbb{R}} \nu(\mathrm{d}x) < \infty$*;*
- $K_{max} > 1$ *and* $K_{min} = -K_{max}$*;*
- *for* $\xi_1, \xi_2 > 0$*,* $\int_{-\infty}^{-1} |x|\, e^{\xi_1|x|}\,\nu(\mathrm{d}x) < \infty$ *and* $\int_1^{\infty} |x|\, e^{\xi_2|x|}\,\nu(\mathrm{d}x) < \infty$*;*
- *for* $\epsilon > 0$*,* $\int_{\mathbb{R}} \left(e^{\epsilon|x|} - 1\right)\,\nu(\mathrm{d}x) < \infty$ *(for the details see Cont & Tankov [2004], Proposition 3.8,).*

Then

$$
\begin{aligned}
& |f(t,x) - f_I(t,x)| \\
\leq \;& e^{C_1^{t,T} + \frac{1}{2} C_1^{t,T}} \, e^{(T-t)\left(C_3\, e^{-\beta_1 |K_{min}|} + C_4\, e^{-\beta_2 |K_{max}|}\right)} \left(e^{(T-t)\left(C_5\, e^{-\beta_1 |K_{min}|} + C_6\, e^{-\beta_2 |K_{max}|}\right)} - 1 \right. \\
+ \;& \left. 2\left(1 - e^{-\alpha(T-t)}\right)\left(C_7\, e^{-\beta_1 |K_{min}|} + C_8\, e^{-\beta_2 |K_{max}|}\right)\right),
\end{aligned}
\tag{3.46}
$$

where $C_1^{t,T} = xe^{-\alpha(T-t)}$, $C_2^{t,T} = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha(T-t)})$, *and some constants* $\beta_1, \beta_2, C_3, C_4, C_5, C_6, C_7, C_8 > 0$.

Proof. We recall a few useful facts and give some notations:

- $|e^x - 1| = (e^x - 1) + 2\,(1 - e^x)^+ \leq (e^x - 1) + 2\,|x|$;
- by $U_T := xe^{-\alpha(T-t)} + \int_t^T \sigma^{-\alpha(T-s)}\, \mathrm{d}W_s$ we denote a process that follows Gaussian distribution with mean $C_1^{t,T}$ and variance $C_2^{t,T}$;
- by $\tilde{Q}_t$ we denote a compound Poisson process with a measure $\tilde{\nu}(\,\mathrm{d}x) := \nu(\,\mathrm{d}x)\, 1_{|x|}$;
- by $\tilde{Q}_t^K$ we denote a compound Poisson process with a measure $\tilde{\nu}^K(\,\mathrm{d}x) := \nu(\,\mathrm{d}x)\, 1_{|x| \in [K_{min}, K_{max}]}$;
- by $D_t^K := \tilde{Q}_t - \tilde{Q}_t^K$ we denote a compound Poisson process with a measure $\hat{\nu}^K(\,\mathrm{d}x) = \tilde{\nu}(\,\mathrm{d}x) - \tilde{\nu}^K(\,\mathrm{d}x) = \nu(\,\mathrm{d}x)\, 1_{|x| \notin [K_{min}, K_{max}]}$;

Then we have

$$
\begin{aligned}
& |f(t,x) - f_I(t,x)| \\
= \;& \left| \mathbb{E}\left[e^{U_T + \int_t^T e^{-\alpha(T-s)} h(X(s))\, \mathrm{d}Q_s} - e^{U_T + \int_t^T e^{-\alpha(T-s)} h(X^K(s))\, \mathrm{d}Q_s^K} \right] \right| \\
\leq \;& e^{C_1^{t,T} + \frac{1}{2} C_2^{t,T}} \left| \mathbb{E}\left[e^{\int_t^T e^{-\alpha(T-s)}\, \mathrm{d}\tilde{Q}_s} - e^{\int_t^T e^{-\alpha(T-s)}\, \mathrm{d}\tilde{Q}_s^K} \right] \right| \\
= \;& e^{C_1^{t,T} + \frac{1}{2} C_2^{t,T}} \left| \mathbb{E}\left[e^{\int_t^T e^{-\alpha(T-s)}\, \mathrm{d}\tilde{Q}_s^K} \left(e^{\int_t^T e^{-\alpha(T-s)}\, \mathrm{d}D_s^K} - 1 \right) \right] \right| \\
\leq \;& e^{C_1^{t,T} + \frac{1}{2} C_2^{t,T}}\, \mathbb{E}\left[e^{\int_t^T e^{-\alpha(T-s)}\, \mathrm{d}\tilde{Q}_s^K} \right]\, \mathbb{E}\left[\left| e^{\int_t^T e^{-\alpha(T-s)}\, \mathrm{d}D_s^K} - 1 \right| \right] \\
\leq \;& e^{C_1^{t,T} + \frac{1}{2} C_2^{t,T}}\, \mathbb{E}\left[e^{\int_t^T e^{-\alpha(T-s)}\, \mathrm{d}\tilde{Q}_s^K} \right] \left(\mathbb{E}\left[e^{\int_t^T e^{-\alpha(T-s)}\, \mathrm{d}D_s^K} - 1 \right] + 2\mathbb{E}\left[\left| \int_t^T e^{-\alpha(T-s)}\, \mathrm{d}D_s^K \right| \right] \right).
\end{aligned}
\tag{3.47}
$$

Now let us show that every term is bounded by some constant. Since the first term is easily computable, we start with the second term

$$
\begin{aligned}
& \mathbb{E}\left[e^{\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}\tilde{Q}_s^K}\right] \\
= \;& e^{\int_t^T \int_{\mathbb{R}}\left(e^{e^{-\alpha(T-s)}|x|}-1\right)\tilde{\nu}^K(\,\mathrm{d}x)\,\mathrm{d}s} \\
= \;& e^{\int_t^T \int_{K_{min}}^0\left(e^{e^{-\alpha(T-s)}(-x)}-1\right)\nu(\,\mathrm{d}x)\,\mathrm{d}s+\int_t^T \int_0^{K_{max}}\left(e^{e^{-\alpha(T-s)}x}-1\right)\nu(\,\mathrm{d}x)\,\mathrm{d}s} \\
\leq \;& e^{\int_t^T \int_{K_{min}}^0\left(e^{-x}-1\right)\nu(\,\mathrm{d}x)\,\mathrm{d}s+\int_t^T \int_0^{K_{max}}(e^{x}-1)\,\nu(\,\mathrm{d}x)\,\mathrm{d}s} \\
= \;& e^{\int_t^T \int_{\mathbb{R}}\left(e^{|x|}-1\right)\tilde{\nu}^K(\,\mathrm{d}x)\,\mathrm{d}s} \\
= \;& e^{(T-t)\left(e^{-\beta_1|K_{min}|}\int_{K_{min}}^0\left(e^{|x|+\beta_1|K_{min}|}-e^{\beta_1|K_{min}|}\right)\nu(\,\mathrm{d}x)+e^{-\beta_2|K_{max}|}\int_0^{K_{max}}\left(e^{|x|+\beta_2|K_{max}|}-e^{\beta_2|K_{max}|}\right)\nu(\,\mathrm{d}x)\right)} \\
\leq \;& e^{(T-t)\left(e^{-\beta_1|K_{min}|}\int_{K_{min}}^0\left(e^{|x|+\beta_1|K_{min}|}-1\right)\nu(\,\mathrm{d}x)+e^{-\beta_2|K_{max}|}\int_0^{K_{max}}\left(e^{|x|+\beta_2|K_{max}|}-1\right)\nu(\,\mathrm{d}x)\right)} \\
\leq \;& e^{(T-t)\left(C_3\,e^{-\beta_1|K_{min}|}+C_4\,e^{-\beta_2|K_{max}|}\right)}.
\end{aligned}
\tag{3.48}
$$

Analogously, we consider the first term in the sum in the brackets of Equation (3.47)

$$
\begin{aligned}
& \mathbb{E}\left[e^{\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}D_s^K}-1\right] \\
= \;& e^{\int_t^T \int_{\mathbb{R}}\left(e^{e^{-\alpha(T-s)}|x|}-1\right)\hat{\nu}^K(\,\mathrm{d}x)\,\mathrm{d}s}-1 \\
= \;& e^{\int_t^T \int_{-\infty}^{K_{min}}\left(e^{e^{-\alpha(T-s)}(-x)}-1\right)\nu(\,\mathrm{d}x)\,\mathrm{d}s+\int_t^T \int_{K_{max}}^{+\infty}\left(e^{e^{-\alpha(T-s)}x}-1\right)\nu(\,\mathrm{d}x)\,\mathrm{d}s}-1 \\
\leq \;& e^{\int_t^T \int_{-\infty}^{K_{min}}\left(e^{-x}-1\right)\nu(\,\mathrm{d}x)\,\mathrm{d}s+\int_t^T \int_{K_{max}}^{+\infty}(e^{x}-1)\,\nu(\,\mathrm{d}x)\,\mathrm{d}s}-1 \\
= \;& e^{\int_t^T \int_{\mathbb{R}}\left(e^{|x|}-1\right)\hat{\nu}^K(\,\mathrm{d}x)\,\mathrm{d}s}-1 \\
\leq \;& e^{(T-t)\left(e^{-\beta_1|K_{min}|}\int_{-\infty}^{K_{min}}\left(e^{|x|+\beta_1|K_{min}|}-1\right)\nu(\,\mathrm{d}x)+e^{-\beta_2|K_{max}|}\int_{K_{max}}^{+\infty}\left(e^{|x|+\beta_2|K_{max}|}-1\right)\nu(\,\mathrm{d}x)\right)}-1 \\
\leq \;& e^{(T-t)\left(C_5\,e^{-\beta_1|K_{min}|}+C_6\,e^{-\beta_2|K_{max}|}\right)}-1.
\end{aligned}
\tag{3.49}
$$

Finally we compute the boundary for $\mathbb{E}\left[\left|\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}D_s^K\right|\right]$. This process D_t^K can be represented as a sum of two compound Poisson processes: P_t^K with a measure $\nu(\,\mathrm{d}x)\,1_{x>K_{max}}$ and N_t^K with a measure $\nu(\,\mathrm{d}x)\,1_{x<K_{min}}$. More precisely, the process $P_t^K \geq 0$ has only positive jumps not smaller than $K_{max} > 1$ and the process $N_t^K \leq 0$ has only negative jumps not greater than $K_{min} < -1$. Then we have

$$
\begin{aligned}
& \mathbb{E}\left[\left|\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}D_s^K\right|\right] \\
= \;& \mathbb{E}\left[\left|\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}P_s^K+\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}N_s^K\right|\right]
\end{aligned}
$$

$$
\begin{aligned}
&\leq \mathbb{E}\left[\left|\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}P_s^K\right|\right] + \mathbb{E}\left[\left|\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}N_s^K\right|\right] \\
&= \mathbb{E}\left[\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}P_s^K\right] - \mathbb{E}\left[\int_t^T e^{-\alpha(T-s)}\,\mathrm{d}N_s^K\right] \\
&\leq \int_t^T e^{-\alpha(T-s)} \int_{-\infty}^{K_{min}} |x|\;\nu(\,\mathrm{d}x)\;\mathrm{d}s + \int_t^T e^{-\alpha(T-s)} \int_{K_{max}}^{+\infty} |x|\;\nu(\,\mathrm{d}x)\;\mathrm{d}s \\
&\leq \left(1-e^{-\alpha(T-t)}\right)\left(e^{-\beta_1|K_{min}|}\int_{-\infty}^{K_{min}} |x|\,e^{\beta_1|K_{min}|}\,\nu(\,\mathrm{d}x) + e^{-\beta_2|K_{max}|}\int_{K_{max}}^{+\infty} |x|\,e^{\beta_2|K_{max}|}\,\nu(\,\mathrm{d}x)\right) \\
&\leq \left(1-e^{-\alpha(T-t)}\right)\left(C_7\,e^{-\beta_1|K_{min}|} + C_8\,e^{-\beta_2|K_{max}|}\right).
\end{aligned}
\tag{3.50}
$$

□

3.4 Results and discussion

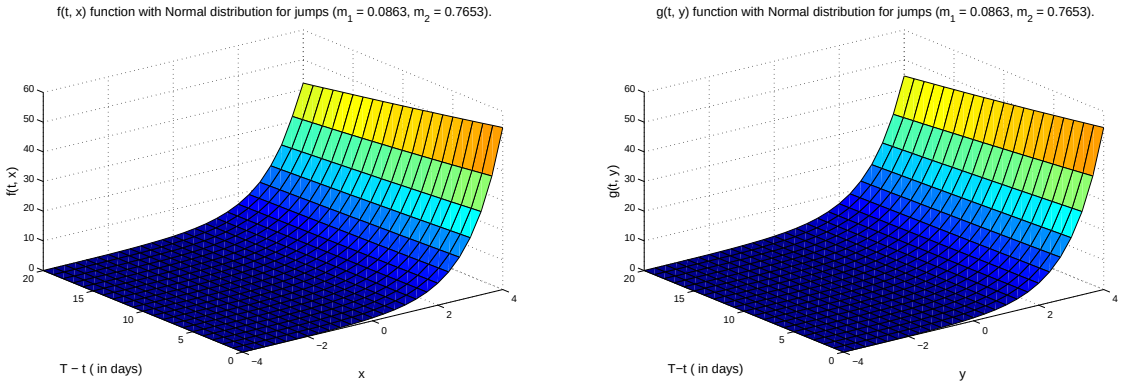
3.4.1 Resulting forwards

This section demonstrates the results of the two models. We plot the forward prices $f(t,x)$ as a solution to Equation (3.24) for the threshold model and the forward prices $g(t,y)$ given in Equation (3.12) for the jump-diffusion model. For both models for the sake of simplicity we assume that the seasonal component $\mu(t)=0$. We use the calibrated parameters obtained from the German spot power market (see Benth *et al.* [2012] and Chapter 5). Table 3.1 contains all the estimated parameters we use.

Table 3.1: An overview over the estimated parameter values for the forward price.

Parameter	Interpretation	Estimated value	Measure unit
α	mean-reversion force	0.6923	approx. 1 day
σ	volatility	2.59	
λ	jump intensity	13.5	spikes per year
$\mathcal{T}$	jump threshold	3	log scale
K_{max}	jump truncation	8	log scale

We start with an assumption of a Gaussian distribution for the jump size. Figure 3.2 shows prices $f(t,x)$ and $g(t,y)$. When we are at the maturity, i.e. $T-t=0$, we observe our boundary condition. We also notice that the prices decrease when time to maturity increases. We detect a rather expected effect that both models produce similar results in terms of the general price level. However, the threshold model produces slightly lower prices compared to the standard jump-diffusion approach. This is perfectly in line with the behaviour of the function h.

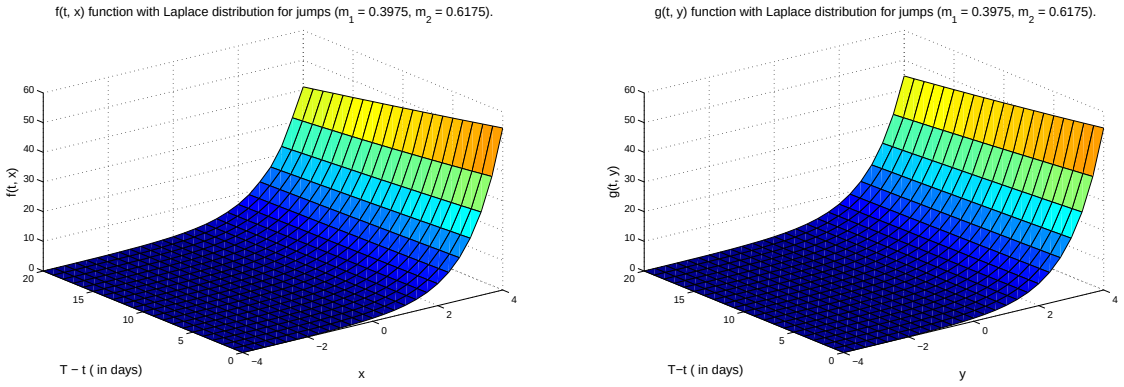


(a) Forward function $f(t,x)$, $\mathcal{T}=3$.

(b) Forward function $g(t,y)$.

Figure 3.2: Normal distribution assumed for jumps with mean $m_1 = 0.0863$ and standard deviation $m_2 = 0.7653$. Time to maturity $T-t = 20$ days.

We continue with an assumption of a Laplace distribution for the jump size. Figure 3.3 shows prices $f(t,x)$ and $g(t,y)$. When we are at the maturity, i.e. $T-t=0$, we can see that our boundary condition is fulfilled. We also note that the prices decrease when time to maturity increases. Again, both models demonstrate similar results in terms of the general price level. Moreover, different jump size distribution does not provide a significant difference to the price level. The threshold model produces slightly lower prices compared to the standard jump-diffusion approach due to the function h.



(a) Forward function $f(t,x)$, $\mathcal{T}=3$.

(b) Forward function $g(t,y)$.

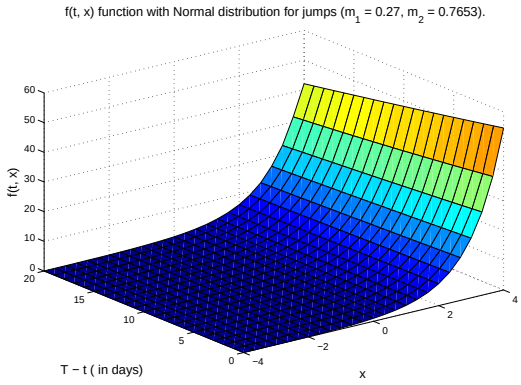
Figure 3.3: Laplace distribution assumed for jumps with location $m_1 = 0.3975$ and scale $m_2 = 0.6175$. Time to maturity $T-t = 20$ days.

Now we triple the average jump size parameter value (both Normal and Laplace) and see how does it effect the forward prices for both models. Figures 3.4 and 3.5 demonstrate that there is no a pronounced change in the level of the prices.

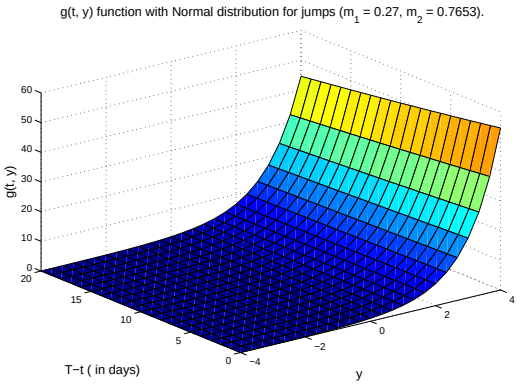
Table 3.2 illustrates this result in the details. We particularly see that in the case of Normal

distribution for the jump size $f(t,x)$ decreases when the average jump size increases. We also note the effect of the function h: when the prices are at the "high" regime, i.e. $x = 4$, then the decrease is more significant. When the prices are at the "low" regime, i.e. $x = 2$, then the decrease is quite small. For the jump-diffusion forward price $g(t,y)$ we notice either a slight increase in the price when the average jump size increase or no change at all.

For the case of Laplace distribution for the jump size $f(t,x)$ decreases dramatically when the average jump size increases compared to the Normal distribution case. This effect is coupled with the impact of the function h when $x = 4$. However, when x is relatively small we even observe an increase in the prices. For the jump-diffusion forward price $g(t,y)$ we see the same effect as for the Normal distribution case.

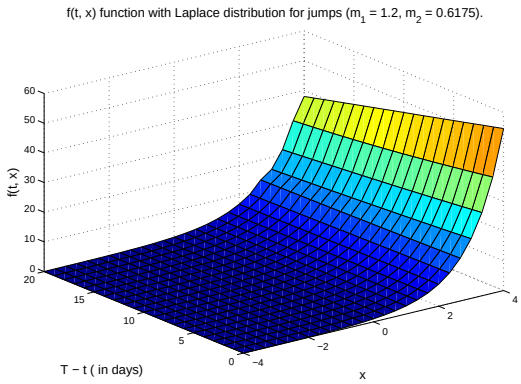


(a) Forward function $f(t,x)$.

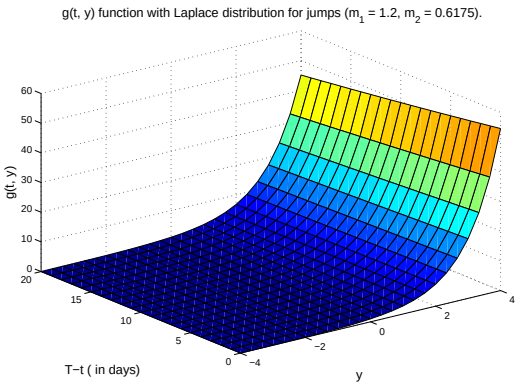


(b) Forward function $g(t,y)$.

Figure 3.4: Normal distribution assumed for jumps with mean $m_1 = 0.27$ and standard deviation $m_2 = 0.7653$. Time to maturity $T - t = 20$ days.



(a) Forward function $f(t,x)$ in the log scale.



(b) Forward function $g(t,y)$ in the log scale.

Figure 3.5: Laplace distribution assumed for jumps with location $m_1 = 1.2$ and scale $m_2 = 0.6175$. Time to maturity $T - t = 20$ days.

Table 3.2: Comparative forward values for the threshold ($\mathcal{T} = 3$) and jump-diffusion models for various starting values x and various average jump size parameter m_1 values. $\tau = 20$ days to maturity.

		x = y = 4		x = y = 2	
		$f(\tau, x)$,	$g(\tau, y)$	$f(\tau, x)$	$g(\tau, y)$
Normal	$m_1 = 0.0863$	42.0727	44.4889	6.5248	6.5248
	$m_1 = 0.27$	42.0345	44.5389	6.5254	6.5248
Laplace	$m_1 = 0.3975$	40.9618	44.6442	6.6176	6.5248
	$m_1 = 1.2$	37.9248	45.1348	6.8678	6.5248

3.5 Conclusion

To summarise our findings we discovered that the Laplace distribution has a remarkable effect on the forward prices compared with the Normal distribution. We demonstrated that the function h shows a noticeable impact on the prices when they are above the specified threshold and when the average jump size is expected to be increasing. This observation is in line with the result of jump size distribution significance obtained in Chapters 2 and 5.

An interesting point, in contrast to modelling spot prices, is that the effect of function h is minor, i.e. it does not demonstrate a prominent difference between forward prices when they are below the threshold.

CHAPTER 4

STORAGE MODELLING

In this chapter we consider the storage problem and present a new approach to the storage value modelling. This approach models the storage level process as a mean-reverting stochastic process within some interval (l, u). To thoroughly describe the problem, we start with the theoretical background and investigate some bounded stochastic processes in Section 4.1. Then, in Section 4.2, we give some application examples. As the main focus of this chapter is on the storage modelling, we further provide an overview of the current modelling approach to the storage problem together with the literature review in Section 4.3 and motivate our modelling approach in Section 4.4. Next, in Section 4.5, we introduce a model for storage process and give a variety of possible payoffs that help a storage owner (and possibly producer) to hedge their market position. Section 4.6 provides some illustrative application examples. Finally, Section 4.7 concludes and gives an outlook for further applications.

4.1 Theoretical development

4.1.1 Preliminary definitions and notations

Karlin & Taylor [1981] (Chapter 15, p. 157) define a diffusion process as
Definition 1 (Diffusion process)**.** *A continuous time parameter stochastic process which possess the (strong) Markov property and for which the sample paths $X(t)$ are (almost always, i.e. with probability 1) continuous functions of t is called a diffusion process.*

The authors also claim (p. 191) that a time homogeneous diffusion process $X(t)_{t\geq 0}$ on a state space $\infty \leq l < u \leq \infty$ satisfies:

- the process $X(t)$ is regular in the interior of the interval (l, u), i.e.

$$\mathbb{P}\{T(y) < \infty \,|\, X(0) = x\} > 0, \quad l < x, y < u,$$

  where $T(y)$ is the first time, if any, the process reaches the value y;

- the process $X(t)$ has infinitesimal parameters $a(x)$ and $\sigma^2(x)$ for (l,u) where $\Delta X = X(h) - X(0)$ and

$$a(x) = \lim_{h\downarrow 0} \frac{1}{h}\mathbb{E}[\Delta X \mid X(0) = x],$$

$$\sigma^2(x) = \lim_{h\downarrow 0} \frac{1}{h}\mathbb{E}[(\Delta X)^2 \mid X(0) = x];$$

- the drift $a(x)$ and the diffusion coefficient $\sigma(x)$ are assumed to be continuous and $\sigma(x) >$ on (l,u).

A scale function $S(x)$ for a diffusion is the unique solution on (l,u) of (up to linear transformation) the following differential equation

$$\frac{1}{2}\sigma^2(x)\,\frac{\mathrm{d}^2 S(x)}{\mathrm{d}x^2} + a(x)\,\frac{\mathrm{d}S(x)}{\mathrm{d}x} = 0. \tag{4.1}$$

A solution to this equation is given as[1]

$$S(x) = \int_{x_1}^{x} \exp\Big\{ -\int_{x_0}^{y} \frac{2a(z)}{\sigma^2(z)}\,\mathrm{d}z\Big\}\,\mathrm{d}y = \int_{x_1}^{x} s(y)\,\mathrm{d}y,$$

for $x \in (l,u)$, x_0, x_1 are points in (l,u) and $s(x)$ is the scale density. A speed measure for a diffusion is denoted by $M(x)$ and determines how fast the process moves through its paths

$$M(x) = \int_{x_0}^{x} \frac{2}{\sigma^2(z)}\,\exp\Big\{\int^{z} \frac{2a(s)}{\sigma^2(s)}\,\mathrm{d}s\Big\}\,\mathrm{d}z = \int_{x_0}^{x} \frac{2}{\sigma^2(z)s(z)}\,\mathrm{d}z = \int_{x_0}^{x} m(z)\,\mathrm{d}z,$$

where $m(x) = \frac{2}{\sigma^2(x)s(x)}$ is the speed measure density.

The following theorem is known as the Feller's test for explosion.

Theorem 1 (Diffusion explosiveness)**.** *The diffusion $X(t)$ is non-explosive if and only if*

$$\lim_{x\to u} \int_{x_0}^{x} (S(x) - S(z))\; m(z)\; \mathrm{d}z = \infty$$

and

$$\lim_{x\to l} \int_{x}^{x_0} (S(z) - S(x))\; m(z)\; \mathrm{d}z = \infty.$$

Proof. The proof can be found in Theorem 5.5.29 in Karatzas & Shreve [1991]. □

If $S(l) = -\infty$ and $S(u) = \infty$, then the boundaries l and u cannot be reached in finite time. If these values are finite then the boundary is called attracting.

[1] For a proof see Proposition 16.78, p. 386, "Probability", Leo Breiman.

There exists the following boundary point b classification:

- A boundary b is **regular**, if $S(b) < \infty$ and $M(b) < \infty$. This means that the process can reach and leave the boundary in finite time.
- If $S(b) < \infty$ and $M[\{b\}] = 0$ the process reflects back into the interior with a infinite speed and this is an **instantaneously** reflecting boundary.
- If $S(b) < \infty$ and $M[\{b\}] = \infty$ the boundary b is **absorbing** or **exit**, i.e. after reaching it the process cannot leave it any more.
- If $S(b) < \infty$ and $M[\{b\}] \in (0, \infty)$ the process reflects back into the interior with an finite speed. The boundary is called the **sticky** boundary in Karlin & Taylor [1981], the **delayed** reflection boundary in Gihman & Skorokhod [1972] and the **slowly** reflecting boundary in Breiman [1968].
- If $S(b) = \infty$ and $M(b) < \infty$, the boundary is called an **entrance** boundary, meaning that the process cannot reach it from within the interval. Although, the process can start from there.
- If $S(b) = \infty$ and $M[\{b\}] = \infty$, the boundary is called **natural**, i.e. it cannot be reached in finite time.

4.1.2 Drift and diffusion coefficients

Consider a time-homogeneous process diffusion $X(t)$ with $t \geq 0$ on the interval (l, u) with $0 < l < u < \infty$ with two boundary points l and u satisfying

$$\mathrm{d}X(t) = a(X(t))\,\mathrm{d}t + \sigma(X(t))\,\mathrm{d}W(t). \tag{4.2}$$

Coefficients $a(x)$ and $\sigma(x)$ uniquely determine the boundary behaviour. We specify the coefficients as follows

$$\begin{aligned} \sigma(x) &= \sqrt{2(u-x)(x-l)}, \\ a(x) &= -a(x-m), \end{aligned} \tag{4.3}$$

where $a > 0$ and $m = \frac{u+l}{2}$.

For our choice of coefficients the scale density is

$$\begin{aligned} s(x) &= \exp\Big\{-\int^x \frac{-a(y-m)}{(u-y)(y-l)}\,\mathrm{d}y\Big\} \\ &= \exp\Big\{\frac{a}{2}\int^x \Big(\frac{1}{u-y}-\frac{1}{y-l}\Big)\,\mathrm{d}y\Big\} \\ &= \exp\{-\frac{a}{2}\ln\left((u-x)(x-l)\right)\} \\ &= ((u-x)(x-l))^{-\frac{a}{2}}, \end{aligned} \tag{4.4}$$

and the speed density is

$$m(x) = ((u-x)(x-l))^{\frac{a}{2}-1} \tag{4.5}$$

for some parameter a which is responsible for the behaviour at the boundaries. If $a = 0$, the boundaries l and u are absorbing points. If $a \in (0,2)$, the process reflects back into the interior instantaneously. If $a \geq 2$ the process never reach the boundaries, although it can start from there.

Recall now definition of the Bessel process $Z(t)$ given in Chapter 11 in Revuz & Yor [1999]
Definition 2 (Bessel process)**.** *For every $\delta \geq 0$ and $x \geq 0$, the unique strong solution of the equation*

$$\mathrm{d}Z_t = \delta\,\mathrm{d}t + 2\sqrt{Z_t}\,\mathrm{d}W_t, \quad Z_0 = x$$

is called the square of δ-dimensional Bessel process started at x and denoted by $BESQ^\delta(x)$. The number δ defines the dimension of $BESQ^\delta(x)$.

In Proposition 1.5 (Chapter 11, p. 442) in Revuz & Yor [1999] the authors prove that when $0 < \delta < 2$ then the process is instantaneously reflecting at the point $\{0\}$; when $\delta \geq 2$ the process never reaches the point $\{0\}$; when $\delta = 0$ the process is absorbed at the point $\{0\}$.

We now can see that our process $X(t)$ is the Bessel process of order a and two boundary points l and u. We recall also that the Bessel process is Markov, for a proof see Theorem 1.9, Chapter 9 in Revuz & Yor [1999].

4.1.3 Infinitesimal operator

For a twice continuously differentiable function $f(x)$ on (l,u) the canonical representation of the differential infinitesimal operator associated with diffusion process is defined as

$$Lf(x) = \frac{\mathrm{d}}{dM}\Big[\frac{\mathrm{d}f}{\mathrm{d}S}\Big] = \frac{1}{m(x)}\frac{\mathrm{d}}{\mathrm{d}x}\Big[\frac{1}{s(x)}\frac{\mathrm{d}f}{\mathrm{d}x}\Big], \tag{4.6}$$

since $dM = m(x)\,\mathrm{d}x$ and $dS = s(x)\,\mathrm{d}x$. This operator can be re-written as

$$Lf(x) = \frac{\sigma^2(x)}{2}\frac{\partial^2 f}{\partial x^2} + a(x)\frac{\partial f}{\partial x}. \tag{4.7}$$

4.1.4 Transformation to ODE

Theorem 2 (Transition probability density)**.** *There exists a continuous map* $q : (0,\infty) \times ((l,u)) \times ((l,u)) \to (0,\infty)$ *such that for all bounded measurable function* ϕ *supported on* (l,u)

$$\begin{aligned} \Phi(t,x) &:= \mathbb{E}_x[\phi(X(t))] \\ &= \int_l^u \phi(y) q_t(x,y)\,\mathrm{d}y, \end{aligned} \tag{4.8}$$

where $q_t(x,y)$ *is the transition probability density function.*

Proof. For a proof see § 4.11 in Ito & McKean [1974]. □

Note also that

$$\begin{aligned} \Phi(0,x) &= \int_l^u \phi(y) q_0(x,y)\,\mathrm{d}y \\ &= \int_l^u \phi(y)\delta(y-x)\,\mathrm{d}y \\ &= \phi(x), \end{aligned} \tag{4.9}$$

by the sifting property of the Dirac delta function $\delta(x)$.Function $\Phi(t,x)$ satisfies the following backward Kolmogorov partial-differential equation with initial condition in (4.10)

$$\frac{\partial}{\partial t}\Phi(t,x) = a(x)\frac{\partial}{\partial x}\Phi(t,x) + \frac{1}{2}\sigma^2(x)\frac{\partial^2}{\partial x^2}\Phi(t,x) \tag{4.10}$$

with two boundary conditions which are the result of our diffusion coefficient $\sigma(x)$ specification

$$\begin{cases} \lim\limits_{x\to l} a(l)\Phi'_x(t,x) = \Phi'_t(t,l), \\ \lim\limits_{x\to u} a(u)\Phi'_x(t,x) = \Phi'_t(t,u). \end{cases} \tag{4.11}$$

To find a solution to this problem we start with applying the Laplace transform to the function $\Phi(t,x)$

$$y_\lambda(x) = \int_0^\infty e^{-\lambda t}\Phi(t,x)\,\mathrm{d}t. \tag{4.12}$$

Then we continue with applying the Laplace transform to Equations (4.10) and (4.11). We obtain a second-order non-homogeneous ordinary differential equation with variable coefficients

$$\frac{\sigma^2(x)}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}y_\lambda(x) + a(x)\frac{\mathrm{d}}{\mathrm{d}x}y_\lambda(x) - \lambda y_\lambda(x) = -\Phi(0,x) \equiv -\phi(x), \tag{4.13}$$

with non-homogeneous boundary conditions

$$\begin{cases} \lim\limits_{x\to l} a(l)\dfrac{\mathrm{d}}{\mathrm{d}x}y_\lambda(x) - \lambda y_\lambda(l) = -\Phi(0,l) \equiv -\phi(l), \\ \lim\limits_{x\to u} a(u)\dfrac{\mathrm{d}}{\mathrm{d}x}y_\lambda(x) - \lambda y_\lambda(u) = -\Phi(0,u) \equiv -\phi(u). \end{cases} \tag{4.14}$$

Due to equivalence of Equations (4.6) to (4.7) we rewrite our equation as

$$\frac{\mathrm{d}}{\mathrm{d}x}\Big[\frac{1}{s(x)}\frac{\mathrm{d}y_\lambda(x)}{\mathrm{d}x}\Big] - \lambda m(x) y_\lambda(x) = -m(x)\phi(x),$$

which we will solve in further section. Before to proceed with this, we discuss some properties of the solution. With this we follow § 24 in Gihman & Skorokhod [1972].

Lemma 1. *Let $\phi(x)$ be a bounded, continuous function and $y_\lambda(x)$ a bounded solution of*

$$\frac{\sigma^2(x)}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}y_\lambda(x) + a(x)\frac{\mathrm{d}}{\mathrm{d}x}y_\lambda(x) - \lambda y_\lambda(x) = -\phi(x) \tag{4.15}$$

for $l < x < u$, satisfying two boundary conditions

$$\begin{cases} a(l)\dfrac{\mathrm{d}}{\mathrm{d}x}y_\lambda(l^+) - \lambda y_\lambda(l) = -\phi(l), \\ a(u)\dfrac{\mathrm{d}}{\mathrm{d}x}y_\lambda(u^-) - \lambda y_\lambda(u) = -\phi(u). \end{cases} \tag{4.16}$$

and let $y'_\lambda(x)\sigma(x)$ be also bounded. Then for all $t > 0$

$$\mathbb{E}[y_\lambda(X(T)) \,|\, \mathcal{F}_t] = e^{-\lambda(T-t)}y_\lambda(X(t)) - \int_t^T e^{-\lambda z}\phi(X(z))\,\mathrm{d}z. \tag{4.17}$$

Proof. Applying Itô's formula, we have

$$\begin{aligned}\mathrm{d}(y_\lambda(X(t))e^{-\lambda t}) &= \Big[-\lambda y_\lambda(X(t)) + \frac{\mathrm{d}}{\mathrm{d}x}y_\lambda(X(t))a(X(t)) \\ &\quad + \frac{1}{2}\sigma^2(X(t))\frac{\mathrm{d}^2}{\mathrm{d}x^2}y_\lambda(X(t))\Big]e^{-\lambda t}\,\mathrm{d}t + y'_\lambda(X(t))\sigma(X(t))e^{-\lambda t}\,\mathrm{d}W(t) \\ &= -e^{-\lambda t}\phi(X(t))\,\mathrm{d}t + y'_\lambda(x)(X(t))\sigma(X(t))e^{-\lambda t}\,\mathrm{d}W(t).\end{aligned}$$

Rewriting the last equation in the integral form and taking the expectation yields the desired result. Note that when $T \to \infty$, then

$$y_\lambda(X(t)) = e^{\lambda t}\mathbb{E}\Big[\int_t^\infty e^{-\lambda z}\phi(X(z))\;\mathrm{d}z \,|\, \mathcal{F}_t\Big]. \tag{4.18}$$

4.1.5 Solution to ODE

Sturm-Liouville problem

The problem we want to solve is given in Equations (4.13) and (4.14). This problem belongs to a class of Sturm-Liouville problems in general given as following

$$\begin{cases} Ly(x) = & -f(x), \\ By(x) = & 0, \end{cases} \tag{4.19}$$

where $x \in [l, u]$ and L denotes the Sturm-Liouville differential operator

$$L = \frac{\mathrm{d}}{\mathrm{d}x}\Big[p(x)\frac{\mathrm{d}}{\mathrm{d}x}\Big] + [q(x) + \lambda r(x)],$$

B is the boundary condition operator

$$B = \begin{cases} \alpha_1 + \alpha_2\,\dfrac{\mathrm{d}}{\mathrm{d}x}, & \text{at } x = l, \\ \beta_1 + \beta_2\,\dfrac{\mathrm{d}}{\mathrm{d}x}, & \text{at } x = u, \end{cases}$$

and λ is a parameter. This problem is *regular* if it satisfies the following properties:

- finite interval $[l, u]$;

- $p(x)$, $p'(x)$, $q(x)$ and $r(x)$ are continuous functions;
- $p(x)$ and $r(x)$ are strictly positive on $[l, u]$.

If any of these conditions is not fulfilled, then the problem is called *singular*. Comparing operator given in 4.6 yields that $q(x) = 0$, $r(x) = m(x)$, $f(x) = m(x)\phi(x)$ and $p(x) = \frac{1}{s(x)}$.

Coefficients transformation

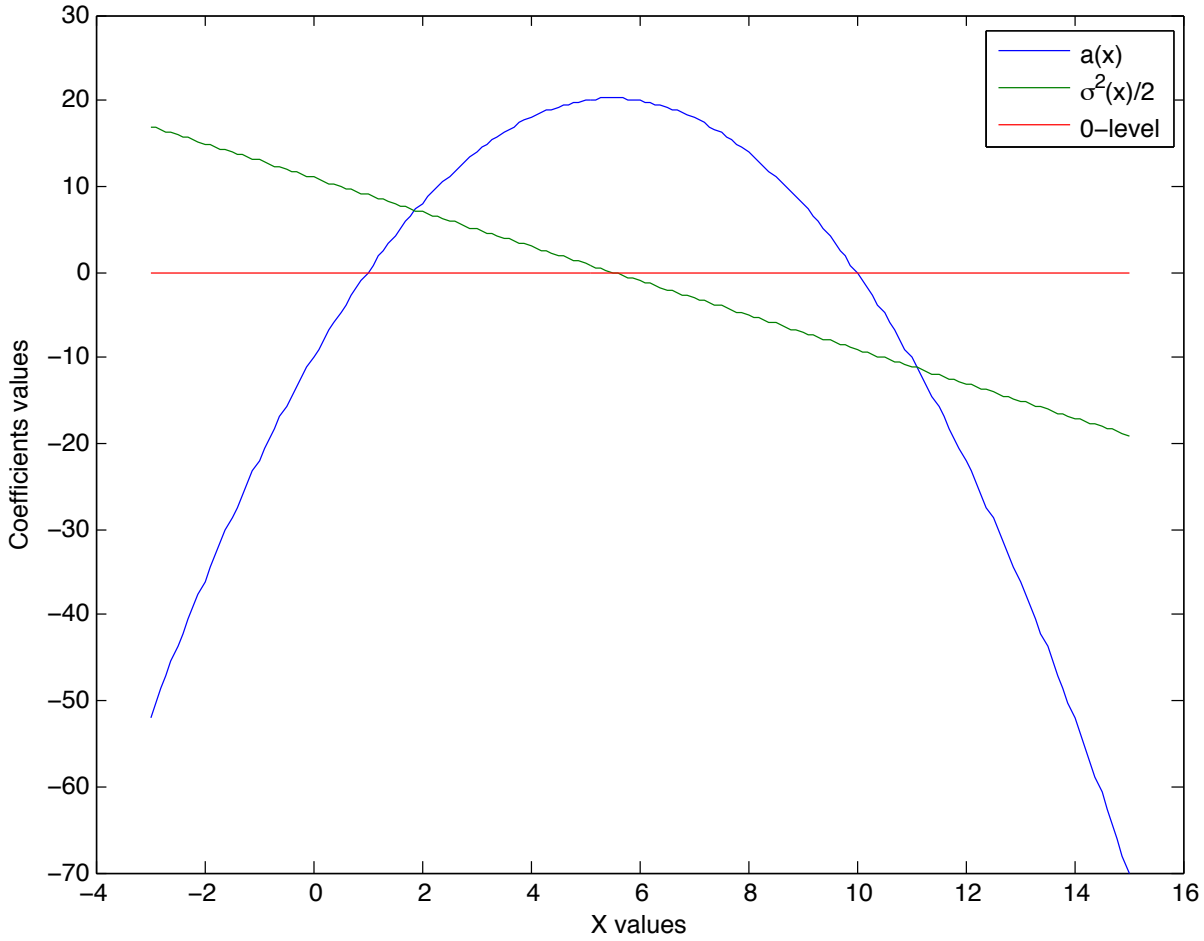


Figure 4.1: Drift $a(x)$ and diffusion $\sigma^2(x)/2$ coefficients.

In Equation (4.3) we specified the form of the drift and diffusion coefficients for the diffusion $X(t)$. Figure 4.1 shows these functions for various values of x.

Let us introduce the following notations we use

- $m := \frac{l+u}{2}$;
- $\Delta := u - l$;

- $z := \frac{2}{\Delta}(x - l) - 1$.

The last substitution gives that $z(x) \in [-1, 1]$, when $x \in [l, u]$. Therefore, $y(x) = y(x(z)) = g(z)$ and our system of homogeneous Equation (4.13) with Conditions 4.14 transform into the following equation

$$\left(1 - z^2\right) g''_{zz} - azg'_z - \lambda g = 0, \tag{4.20}$$

with boundary conditions

$$\begin{cases} ag'(-1) - \lambda g(-1) = & 0, \\ -ag'(1) - \lambda g(1) = & 0. \end{cases} \tag{4.21}$$

The homogeneous version of this equation is equivalent to the Jacobi differential equation

$$(1 - x^2)y'' + (\beta - \alpha - (\alpha + \beta + 2)x)y' + n(n + \alpha + \beta + 1)y = 0, \tag{4.22}$$

when $\alpha = \beta$. The Jacobi polynomials are the solution to this equation. An extensive study of the Jacobi polynomials can be found in the book of Szego [1939]. Let us define Jacobi polynomials and recall some useful and necessary facts. The classical definition is given by Rodrigues' formula

$$(1 - x)^\alpha (1 + x)^\beta P_n^{\alpha,\beta}(x) = \frac{(-1)^n}{2^n n!} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \Big\{ (1 - x)^{n+\alpha} (1 + x)^{n+\beta} \Big\}, \qquad x \in [-1, 1]. \tag{4.23}$$

For $\alpha, \beta > -1$ these polynomials are mutually orthogonal in $\mathcal{L}^2$ (see Theorem 4.2.2 in Szego [1939]), namely

$$\int_{-1}^{1} P_n^{\alpha,\beta}(x)\, P_m^{\alpha,\beta}(x)\, (1 - x)^\alpha (1 - x)^\beta \, \mathrm{d}x = \frac{2^{\alpha+\beta+1} \Gamma(n + \alpha + 1)\Gamma(n + \beta + 1)}{(2n + \alpha + \beta + 1)\Gamma(n + 1)\Gamma(n + \alpha + \beta + 1)} \delta_{n,m}, \tag{4.24}$$

where $\delta_{n,m}$ is the Kronecker delta function. When

- when $\alpha = \beta$, the polynomials are called ultraspherical or the Gegenbauer polynomials;
- when $\alpha = \beta = 0$ the polynomials are called the Legendre polynomials;

These polynomials play a key role in the form of the transition probability density of the diffusion process as it is stated in Cooper *et al.* [1977]. They are needed to construct the Green function, the inverse Laplace transform of which is the transition density function.

Our Equation (4.20) is solved by the Gegenbauer (or ultraspherical) polynomials for $\alpha = \frac{a}{2} - 1$, $\lambda = -n(n + 2\alpha + 1)$ and $\alpha = \gamma - \frac{1}{2}$ which are defined as

$$P_n^{\gamma}(x) = \frac{\Gamma(\gamma + \frac{1}{2})\Gamma(n + 2\gamma)}{\Gamma(2\gamma)\Gamma(n + \gamma + \frac{1}{2})} P_n^{\gamma - \frac{1}{2}, \gamma - \frac{1}{2}}(x), \qquad x \in [-1, 1], \tag{4.25}$$

for $\alpha > -1$ or $\gamma > -\frac{1}{2}$ with a modified Rodrigues' formula

$$P_n^{\gamma}(x) = \frac{(-2)^n \Gamma(n + \gamma)\gamma(n + 2\gamma)}{n!\Gamma(\gamma)\Gamma(2n + 2\gamma)} (1 - x^2)^{-\gamma + \frac{1}{2}} \Big(\frac{\mathrm{d}}{\mathrm{d}x} \Big)^n (1 - x^2)^{n + \gamma - \frac{1}{2}}, \qquad x \in [-1, 1]. \tag{4.26}$$

A mutual orthogonally property is given as

$$\int_{-1}^{1} P_n^{\gamma}(x)\, P_m^{\gamma}(x)\, (1 - x^2)^{\gamma - \frac{1}{2}}\, \mathrm{d}x = \underbrace{\frac{\pi 2^{1-2\gamma}\Gamma(n + 2\gamma)}{n!(n + \gamma)\Gamma^2(\gamma)}}_{:= A(n,\gamma)} \delta_{n,m}. \tag{4.27}$$

So with $\frac{a}{2} - 1 = \alpha = \beta = \gamma - \frac{1}{2}$ with $\alpha > -1$ or $\gamma > -\frac{1}{2}$ (to ensure mutual orthogonality) we define three cases.

Case 1: $a \in (0, 2)$ or $\alpha = \beta \in (-1, 0)$ or $\gamma \in (-\frac{1}{2}, \frac{1}{2})$. With these values of parameters the behaviour at the boundary l or u is characterised by the instantaneous reflection. The eigenfunctions to the differential operator in Equation (4.20) in this case are the Gegenbauer polynomials.

Case 2: $a \geq 2$ or $\alpha \geq 0$ or $\gamma \geq \frac{1}{2}$. With these values of a the process never reaches the boundary l or u in a finite time. Although, the process can start at one of the boundaries. The eigenfunctions to the differential operator in Equation (4.20) in this case are the Legendre polynomials.

Case 3: $a = 0$ or $\alpha = -1$ or $\gamma = -\frac{1}{2}$. With these values of parameters the process will be absorbed at the boundary l or u. The eigenfunctions to Equation (4.20) in this case are the Jacobi polynomials, but in this case the mutual orthogonality condition is lost.

The Green function

Green's function methods is often used to solve the boundary value problems. There is an extensive literature on the Green function application given for example in Stakgold [1979]. In general, the Green function $G(x, \xi)$ is defined as a solution to a homogeneous form of the boundary value problem given as in Equation (4.19)

$$\begin{cases} LG(x,\xi) = 0, \quad a < x,\, \xi < b; \\ BG(x,\xi) = 0; \\ G(x,\xi) \text{ is continuous at } x = \xi; \\ \dfrac{\mathrm{d}G}{\mathrm{d}x}\Big|_{x=\xi^+} - \dfrac{\mathrm{d}G}{\mathrm{d}x}\Big|_{x=\xi^-} = \dfrac{1}{p(x)}. \end{cases} \tag{4.28}$$

Also, the differential operator applied to the Green function should be equal to the Dirac delta function

$$LG(x,\xi) = -\delta(x-\xi), \quad a < x,\, \xi < b; \quad BG(x,\xi) = 0.$$

Another important property of the Green function is that at some conditions it can be represented via the eigenfunctions $v_n(x)$ of the linear differential operator L which is known as bilinear form

$$G_\lambda(x,\xi) = \sum_{n=0}^{\infty} \frac{v_n(x)v_n(\xi)}{(\lambda_n - \lambda)\int_a^b r(x)v_n^2(x)\,\mathrm{d}x},$$

where λ_n are the eigenvalues of the operator. If the eigenfunctions are orthonormal then $c_n \equiv \int_a^b r(x)v_n^2(x)\,\mathrm{d}x = 1$, if not then we can normalise them to have this property. From this representation it can be seen that Green's function has an infinite number of poles at λ_n, these poles are called the point spectrum of the Green function. There is a relation between the Green function and the poles in the complex λ-plane known as the following formula

$$\frac{1}{2\pi i}\oint_C G_\lambda(x,\xi)\ \mathrm{d}\lambda = -\frac{\delta(x-\xi)}{r(x)}. \tag{4.29}$$

The proof is given in Duffy [2001], Chapter 2, p. 49. After finding the Green function, the unique solution to the non-homogeneous equation with homogeneous boundary conditions as in Equation (4.19) is given as

$$y(x) = \int_a^b G(x,\xi)\, f(\xi)\ \mathrm{d}(\xi).$$

This form of the solution is approved by the sifting property of the delta function, meaning that

$$Ly = \int_a^b LG(x,\xi)\, f(\xi)\ \mathrm{d}(\xi) = -\int_a^b \delta(x-\xi)\, f(\xi)\ \mathrm{d}(\xi) = -f(x).$$

Now coming back to our problem which we transformed to the Gegenbauer polynomials with $\alpha = \beta$, we can write down the solution to 4.13 via Green's function with normalised eigenfunctions $u_n(z)$ for Equation (4.20)

$$u_n(z) = \frac{1}{\sqrt{A(n,\gamma)}} P_n^{\gamma}(z) = \frac{1}{\sqrt{A(n,\gamma)}} P_n^{\gamma}(z(x)) = \frac{1}{\sqrt{A(n,\gamma)}} P_n^{\gamma}\Big(\frac{2}{\Delta}(x-l)-1\Big) = v_n(x), \quad (4.30)$$

which are the eigenfunctions for our differential operator $L = \frac{\mathrm{d}}{\mathrm{d}x}\Big[\frac{1}{s(x)}\frac{\mathrm{d}}{\mathrm{d}x}\Big] - \lambda m(x)$ with the eigenvalues $\lambda_n = -n(n+2\gamma+1)$, and

$$y_\lambda(x) = \int_l^u G_\lambda(x,\xi)\,\phi(\xi)\,m(\xi)\;\mathrm{d}\xi = \int_l^u \sum_{n=0}^{\infty} \frac{v_n(x)\,v_n(\xi)}{(\lambda-\lambda_n)}\,\phi(\xi)\,m(\xi)\;\mathrm{d}\xi. \quad (4.31)$$

4.1.6 Transition density formula

This is the last step in getting the transition density formula $q_t(x,y)$. Recall again Equation (4.8)

$$\Phi(t,x) = \int_l^u \phi(y) P(t,x,\,\mathrm{d}y), \quad (4.32)$$

with the transition probability $P(t,x,\,\mathrm{d}y)$. By Fubini's theorem we have

$$\begin{aligned} y_\lambda(x) &= \int_0^\infty e^{-\lambda t}\Phi(t,x)\,\mathrm{d}t \\ &= \int_0^\infty e^{-\lambda t}\int_l^u \phi(y)P(t,x,\,\mathrm{d}y)\,\mathrm{d}t \\ &= \int_l^u \phi(y)\int_0^\infty e^{-\lambda t} q_t(x,y)\,\mathrm{d}t\;\mathrm{d}y \\ &= \int_l^u K_\lambda(x,y)\phi(y)\;\mathrm{d}y \end{aligned} \quad (4.33)$$

where $K_\lambda(x,y)$ is the Laplace transform of the transition density function $q_t(x,y)$. Comparing this expression to the solution given in Equation (4.31) gives

$$K_\lambda(x,y)\Big|_{(l,u)} = G_\lambda(x,y)m(y). \quad (4.34)$$

Now taking the inverse Laplace transform of Equation (4.34) results in the spectral representation of the transition density function of the process in a bounded region

$$\begin{aligned} q_t(x,y)\,|_{(l,u)} &= \mathcal{L}^{-1}\Big(K_\lambda(x,y)\Big) \\ &= \mathcal{L}^{-1}\Big(m(y)\,G_\lambda(x,y)\Big) \\ &= \mathcal{L}^{-1}\Big(m(y)\sum_{n=0}^{\infty}\frac{u_n(x)u_n(y)}{(\lambda-\lambda_n)}\Big) \\ &= m(y)\sum_{n=0}^{\infty} v_n(x)v_n(y)\,e^{-n(n+2\gamma)t} \\ &= m(y)\sum_{n=0}^{\infty}\frac{P_n^\gamma\Big(\frac{2}{\Delta}(x-l)-1\Big)P_n^\gamma\Big(\frac{2}{\Delta}(y-l)-1\Big)}{A(n,\gamma)}\,e^{-n(n+2\gamma)t} \\ &= (1-y^2)^{\gamma-\frac{1}{2}}\sum_{n=0}^{\infty}\frac{P_n^\gamma\Big(\frac{2}{\Delta}(x-l)-1\Big)P_n^\gamma\Big(\frac{2}{\Delta}(y-l)-1\Big)}{A(n,\gamma)}\,e^{-n(n+2\gamma)t} \end{aligned} \tag{4.35}$$

4.1.7 Numerical investigation

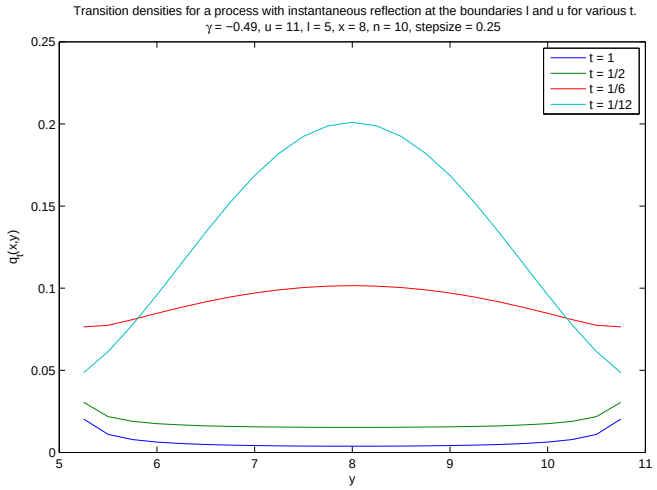
Numerical examples

Here we demonstrate the transition probability density for various parameters values.

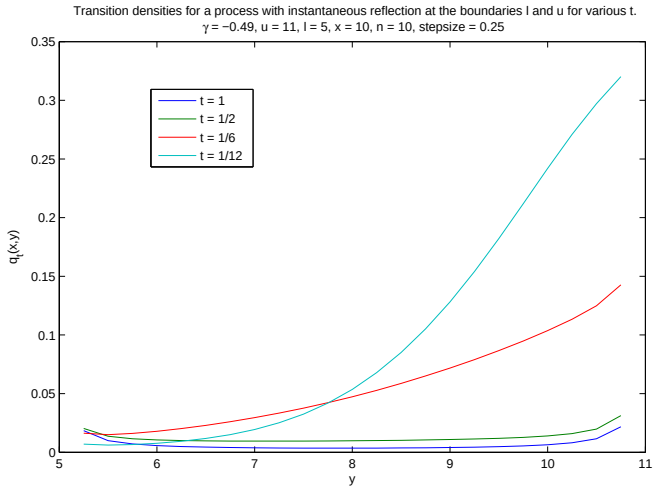
1. Absorption at the boundary: Coefficients $a = 0$ or $\alpha = -1$ or $\gamma = -1/2$ ensure that the boundaries absorb the process, but in this case there is no mutual orthogonality. With this we provide an example of the density with "almost absorption", i.e. we take $\gamma = -0.49$ and see how does it effects the transition probability density near the boundary values. This is shown in Figure 4.2.

2. Instantaneous reflection at the boundary: Coefficients $a \in (0,2)$ or $\alpha \in (-1,0)$ or $\gamma \in (-1/2, 1/2)$ ensure that the boundaries reflect the process with an infinite speed. The behaviour of the transition probability density is shown in Figure 4.3.

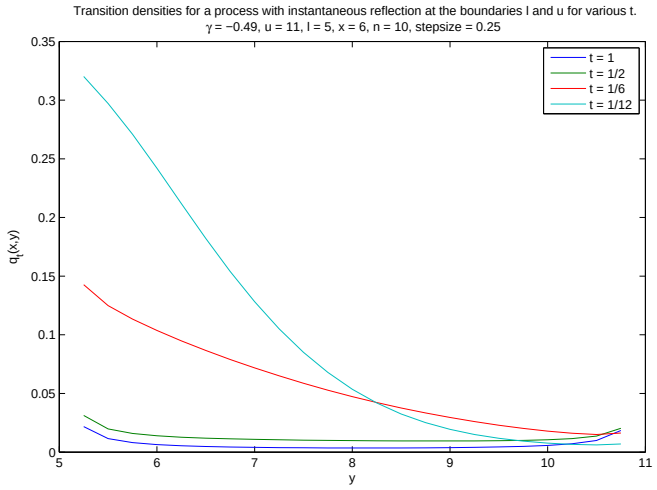
3. Boundary is never attainable: Coefficients $a \geq 2$ or $\alpha \geq 0$ or $\gamma = 1/2$ ensure that the boundary is never reachable. The behaviour of the transition probability density is shown in Figure 4.4.



(a)

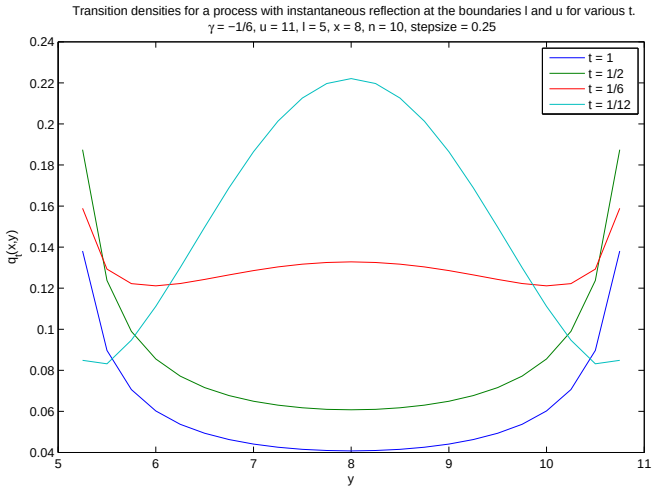


(b)

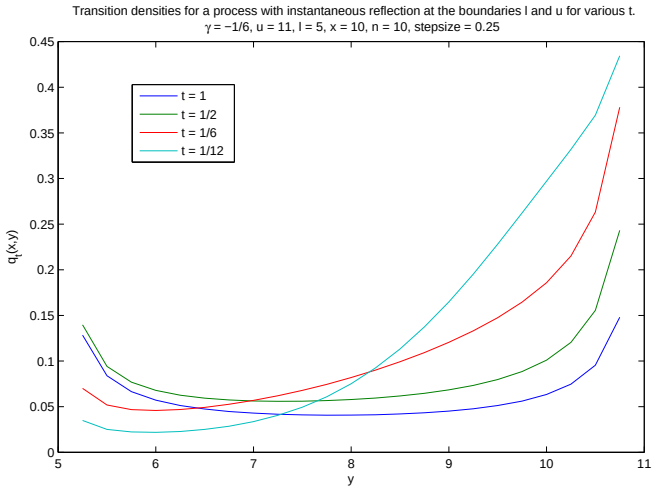


(c)

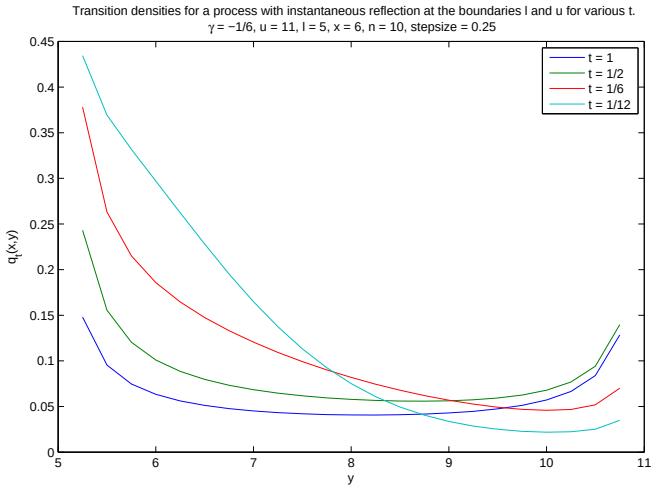
Figure 4.2: Transition probability density function $q_t(x, y)$ for a diffusion that is "almost absorbed" at the boundaries l and u, $n = 10$.



(a)

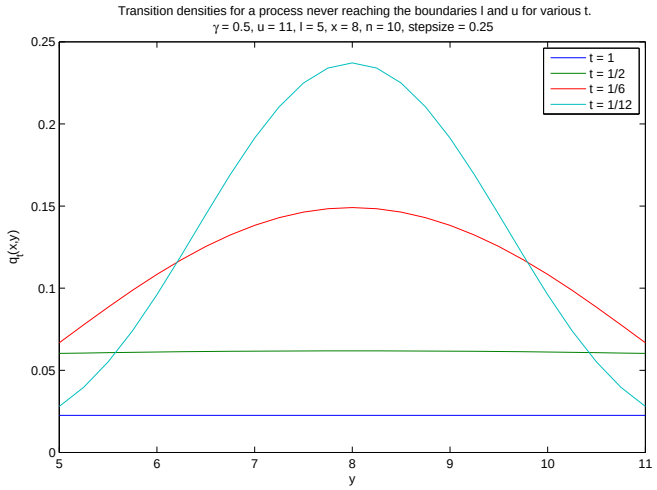


(b)

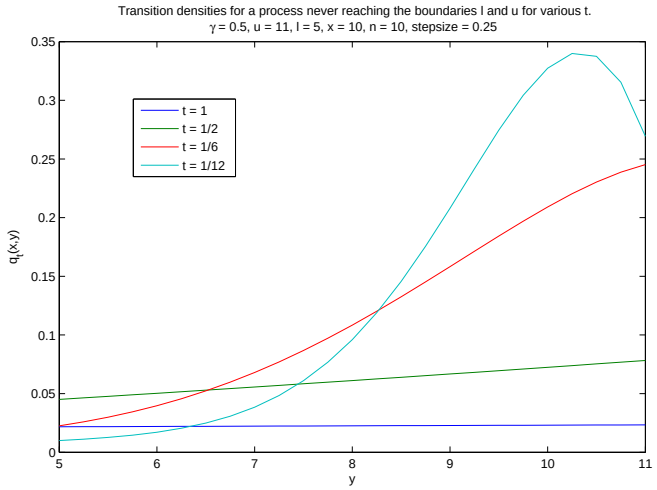


(c)

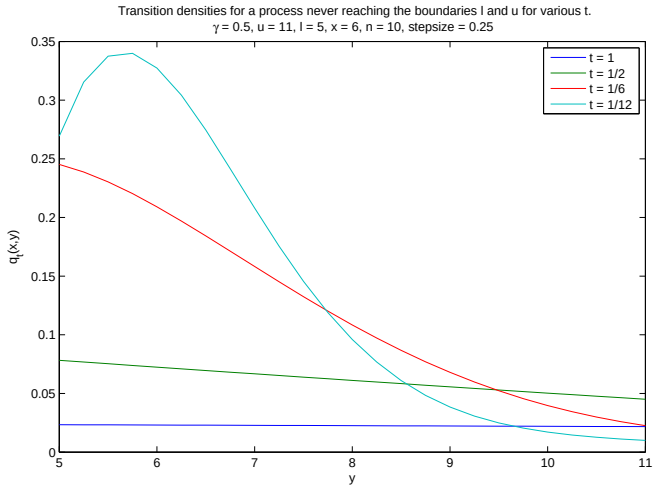
Figure 4.3: Transition probability density function $q_t(x, y)$ for a diffusion that is instantaneously reflected at the boundaries l and u, $n = 10$.



(a)



(b)



(c)

Figure 4.4: Transition probability density function $q_t(x, y)$ for a diffusion that never reaches none of the boundaries l and u, $n = 10$.

Convergence and computational time

The transition density function is given as an infinite sum of the elements. To numerically compute it we need to truncate the sum at some number n which makes the computational time and effort reasonable. In Figures 4.2, 4.3 and 4.4 we use $n = 10$ which is sufficiently fast. We tried several possible values n and found out that after $n = 5$ the difference in the produced values is insignificant.

4.2 Possible applications of a bounded diffusion to the energy-related markets

In this section we briefly overview some possible applications of considered diffusion process $X(t)$ in a bounded region (l, u) to the energy-related markets.

In a contrast to other financial markets, energy markets exhibit very specific behaviour which is a result of physical and regulatory constraints associated with traded energy commodities. To model a price process as a bounded diffusion can be very beneficial at the attempts to investigate and understand market performance.

One of the potential area of a application is a market of European emissions certificates. These financial instruments are a result of the emissions trading scheme introduced in European countries in 2005 with a main goal to regulate climate policy and to reduce emissions of CO_2 and greenhouse gas emissions. Every certificate allows a holder to emit 1 tonne of CO_2, so the companies have to buy these allowances at the market to fulfil the commitments. Otherwise, they have to pay a penalty and on top of that they have to buy missing certificates. To increase investments in the renewable energy production sector some countries (i.e. United Kingdom) introduce the so-called floor price for this emission certificate, which mainly can be regarded as an additional tax that a company has to pay. With these peculiarities, a carbon price as a combination of emissions price together with the tax and the penalty can be viewed as a bounded stochastic process living between these tax- and penalty-boundaries. The most curious question is what kind of a boundary behaviour could we model to keep our assumptions as much as realistic as possible? This investigation is beyond the scope of this thesis, but definitely will be a focus of our future research activity.

Other example of the markets where a price process can exhibit a boundary behaviour is a storage valuation problem. We study this case in details in a subsequent section.

4.3 Current setting and literature

The problem of modelling storage is not new, but remains very demanding and challenging due to its direct purpose of matching the supply and demand in energy markets. The key purposes of controlling a storage process include keeping the balance in the reservoir, meeting the changing demand, hedging market positions, insuring against various sudden events, performing market speculations and others.

To meet the seasonal changes in demand both renewable- and fuel-driven production industries have storage reservoirs. For example, the hydro-dominated Nordic power market (i.e. Norway and Sweden) represents private and public parties operating hydro reservoirs. As stated in Kauppi & Liski [2008], this hydro system has some specific features such as a weather-dependency in spring and fall, many different inflow and outflow technical constraints in hydro turbines and others. Since inflow is highly seasonal and exhibits some instability, there is a strong interconnection between the markets in Scandinavian countries. For instance, depending on the conditions, the necessary amount of hydro power can safely be transported from one region to another. Moreover, there exists a cumulative hydro storage index, available at the Nord Pool exchange, that shows current hydro reservoir level across the countries and total maximum capacity.

Other storage alternatives to have a quick access to are facilities to easily store fuels, mainly gas. These facilities have some specific properties and characteristics one should keep in mind. Among them are the reservoir capacity constraint and injection and withdrawal rate constraints. The latter rate constraint regulates the speed of injection or withdrawal depending on the current reservoir level. Other important operating characteristics are the base gas (cushion) level that ensures the critical pressure in the pipeline and the working gas level which allows one to operate in the market. Also, there is a cyclability constraint representing a number of cycles of injection or withdrawal per year. Furthermore gas storage entails various operational and managerial costs. Additional to these costs there are possible pipeline seepage rates which describe the amount of gas that is lost during injection or withdrawal. On top of that there might be some regulatory constraints.

Technically, there exist three types of underground gas storage facilities: salt caverns, aquifers and depleted oil or gas reservoirs [Federal Energy Regulatory Commission and others, 2004]. The first type of facility has relatively high deliverability and injection rates and is often used for short-term purposes. The second type of facility has high cushion level requirements and a high deliverability rate. The last one is the most common gas storage provision and is used for seasonal system supply or for peak-day demands.

We fairly note that hydro storage and gas storage problems have some issues in common. Particularly, the hydro storage problem addresses the questions of when and how much water to

release or to save and how much power to produce respectively. The gas storage problem addresses the questions of when to withdraw and sell and when to buy in the market and inject. While the former problem has not extensively been discussed in the literature to the best of our knowledge, the latter problem was under a quite focus for the last decade in the literature. Papers of Ahn *et al.* [2002], Chen & Forsyth [2007], Kjaer & Ronn [2008], Thompson *et al.* [2009], Carmona & Ludkovski [2010] investigate the working gas storage value problem as a stochastic optimal control problem.

They consider a control policy which defines the periods of injection, withdrawal or "doing nothing" in such a way that the total profit of a storage holder is maximised with respect to some constraints. Let us describe some details. There is a physical storage at level S_t which is limited up to the maximal storage capacity S_{max}. The market price of gas P_t can either be considered as a futures price $F(t,t)$ with some respective assumptions on $F(t,T)$ or can alternatively be modelled as a stochastic mean-reverting process possibly with jumps. There are two considered rates: injection $a_{in}(S_t) > 0$ and withdrawal $a_{out}(S_t) < 0$, not necessarily equal to each other by their absolute value. There are possibly some costs of injection or withdrawal together with some other operational and managerial costs of storage. Furthermore, there is a finite (or infinite) horizon with either continuous (or discrete) time setting. All this sets up the following optimisation problem of finding an optimal switching policy between injection, withdrawal or "doing nothing" regimes. This is a stochastic optimal control problem, since one seeks for an optimal strategy c from the class $\mathcal{C}_t$ of all admissible strategies. Given starting values at time t one has the following formulation

$$V(t,P_t,S_t) = \sup_{c\in\mathcal{C}_t} \mathbb{E}\Big[\int_t^T h(c_s,P_s,S_s)\,\mathrm{d}s\Big], \tag{4.36}$$

where h is a specified payoff that a storage owner receives at time t implementing the strategy c. Depending on the assumptions and modelling properties, this Hamilton-Jacobi-Bellman type of problem can be solved with several techniques extensively available in the literature. The resulting optimal stochastic control policy reveals three regimes: if the current working gas level in storage is low, then with the gas price increasing one is moving from a strategy of pumping up to "doing nothing" with gas. Conversely, if the current working gas level in the storage facility is high, then with the gas price increasing one has an opportunity of releasing gas from storage to sell. This control strategy corresponds to the following policy: one sells the gas and it results in the highest value when the prices are high and the reservoir is full. Respectively, if the prices are high and the reservoir is empty, then one neither sells or buys. In this paper we will not focus on solving the stochastic optimal control problem, instead we assume that the optimal policy of injection or withdrawal is given and investigate a number of important financial products that an owner can use in order to hedge the market position.

4.4 Motivation

In a contrast to the approach described above, we look at the problem differently and develop a method that allows us to study the storage value problem from the stochastic modelling and statistical points of view. We consider a working storage (gas or hydro) as a mean-reverting bounded stochastic process, assuming that the control policy to inject or withdraw is already given.

The motivation for modelling storage level as a random process is the following. Consider a producer who owns a storage reservoir: she has to regularly decide on injection or withdrawal or "doing nothing" policy depending on various external factors. When she deals with the hydro storage problem, one of the key factors is the power price P_t. A producer tends to release the water to produce power if the current power price level is relatively high and if the current water level in the reservoir allows her to do so. Alternatively, if the current price level is low, one can only opt for a small rate of production. When a producer deals with the gas storage problem, then one of these factors has been considered in the literature as a gas spot price, also P_t. [1] However, since the spot price contains information up to time t, we assume that this producer looks at the futures market and takes a decision respectively. If the market is currently in contango, meaning that the value $D_t = F(t,T) - \mathbb{E}^{\mathbb{Q}}[P_T \,|\, \mathcal{F}_t] > 0$, then a producer can expect that the market is willing to pay more in the future. The opposite case is the backwardation, meaning that the value $D_t = F(t,T) - \mathbb{E}^{\mathbb{Q}}[P_T \,|\, \mathcal{F}_t] < 0$, then a producer can expect that the market is willing to pay less in the future. So this would help to either inject during contango or withdraw during backwardation. Since one usually observes contango in summer and backwardation in winter, we can think of the value D_t as a process which is reverting around zero. This would imply that we follow the strategy to inject when the market is in contango the storage level is below some mean level m. And we follow the strategy to withdraw when the market is in backwardation the storage level is above the mean level m.

Since demand is highly seasonal, managing inventories plays a big role in various risk hedging methods. A stochastic model for storage which does not include the stochastic control component would shed some light on the storage value dynamics and gives an intuition to hedging against a price collapse or other unexpected events. Another benefit of such a setup is that it gives a quick and simple way to estimate the value of owning a storage facility knowing the current market price.

In the next section the model with all the necessary components for storage dynamics, spot market price dynamics and the value process. For the sake of comparison, we also consider several payoffs for hydro and gas storage problems respectively. Further we give several illustrative examples and provide a discussion on the results.

[1] Further in the text we refer P_t as a fuel price which can either be power or gas price respectively.

4.5 Storage process modelling

4.5.1 Modelling setup

Let $(\Omega, \mathbb{P}, \mathcal{F}, \{\mathcal{F}_t\})$ be a complete filtered probability space. We specify the model assumptions and parameters:

- a storage owner can also be an electricity producer;
- continuous time setting;
- finite time horizon $t, T \in [T_1, T_2]$;
- S_t is the current level of working storable commodity in the reservoir at moment t measured in MWh;
- reservoir capacity is restricted naturally by $0 < l < u < \infty$ with l is the minimum reservoir level, u is the maximum reservoir level, $m = \frac{u+l}{2}$ is the average level and $\Delta = u - l$ is the total reservoir capacity;
- $a(S_t)$ is the rate at which we inject or withdraw;
- P_t is the spot price (gas or power);
- $F(t, T)$ is the futures price (gas or power) with maturity T;
- $V_t(S_t, P_t, C_t)$ is the storage value at time t;
- $r(t, T)$ is the discount factor over the period of (t, T);
- C_t represents some cumulative (operational, managerial or switching) costs.

We model the storage level dynamics S_t as a stochastic mean-reverting process which stays between (l, u) as follows

$$\mathrm{d}S_t = -2(S_t - m)\ \mathrm{d}t + \sqrt{2(S_t - l)(u - S_t)}\ \mathrm{d}W_t^S. \tag{4.37}$$

This is an example of the diffusion process in a bounded domain with a "never-reaching boundary" behaviour considered in the previous section. An illustration of such a process is given in Figure 4.5. This formulation suggests that the injection and withdrawal rates are defined by $\mathrm{d}S_t$. The drift term becomes positive when the reservoir is relatively empty and needs to be re-filled and the drift term becomes negative when the reservoir is relatively full and needs to be emptied. The diffusion term ensures the fact that the process S_t always stays inside the interval (l, u) and never reaches the boundaries l and u, which is exactly the case for the real storage

level process due to regulatory constraints on the minimum and maximum reservoir levels l and u.

Here we focus on the hydro- and gas-driven storage reservoirs. We consider the hydro-driven reservoir which is naturally filled with melted snow or rain. For some European countries like Switzerland, Austria, Norway and Sweden such a hydro reservoir is playing a significant in the electricity production. So the amount of precipitation can be regarded as a random process. Fleten [2013] presented some data from Norwegian producers operating hydro storage reservoirs. His data show the random nature of the inflow process. Another important issue discussed in his presentation was the so-called target level set by the producer. This target level is given by a time-dependent component that can be explained by seasonal behavior of the inflow to the storage facility: due to high power demand in winter and low power demand in summer. This effect can be captured by incorporating some circular (e.g. trigonometric) function. In our model for the sake of simplicity we refer to this seasonal component as a constant level m, namely we assume that $m := \mathbb{E}[m(t)]$.

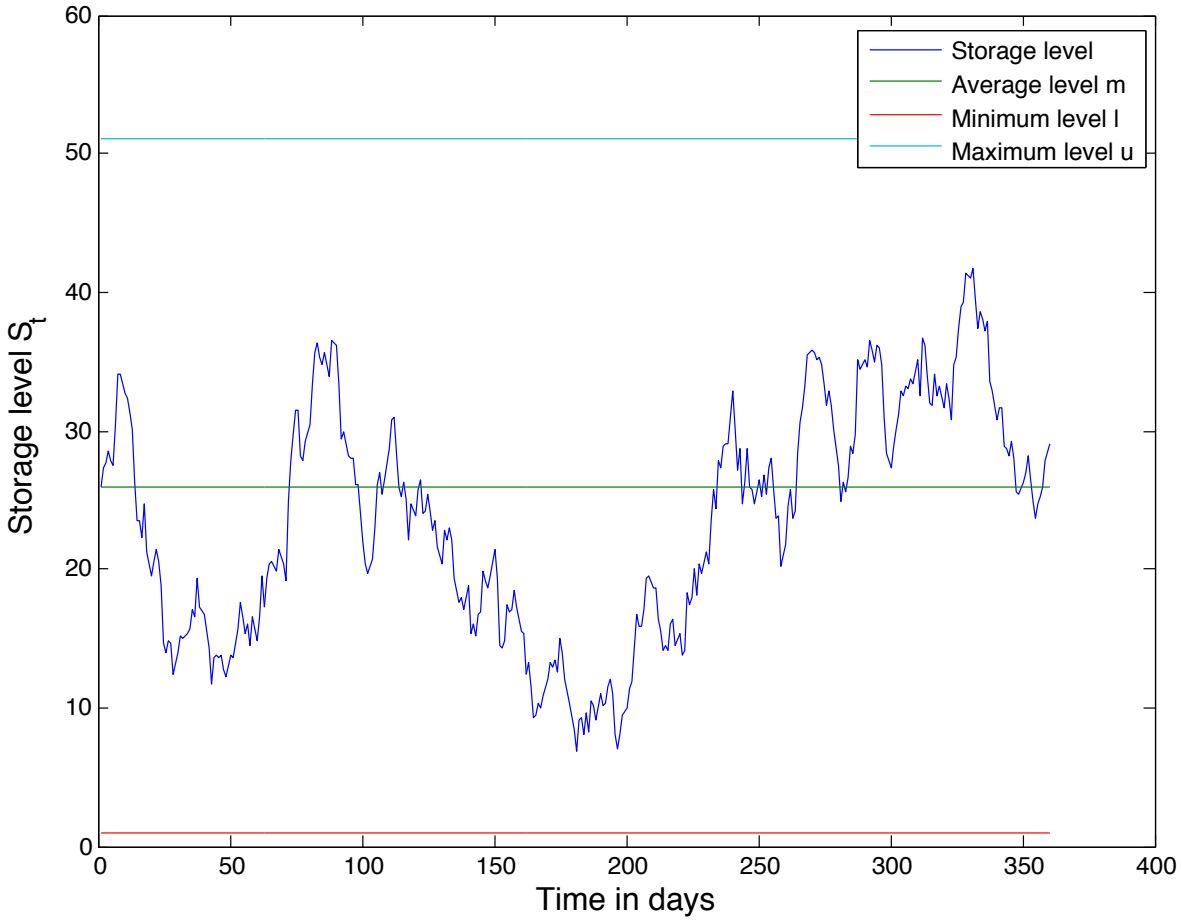


Figure 4.5: An example of a storage level process with $l = 1$ and $u = 51$.

The fuel (gas or power) spot price dynamics is described by an exponential Ornstein-Uhlenbeck process without jumps which ensures the price positivity (here we ignore the fact that sometimes power can exhibit negative prices), namely

$$\begin{aligned} P_t &= e^{f(t)+X_t}, \\ \mathrm{d}X_t &= -\alpha X_t\,\mathrm{d}t + \sigma\,\mathrm{d}W_t^X, \\ \mathrm{d}P_t &= \alpha\Big(\mu(t) - \ln P_t\Big)P_t\,\mathrm{d}t + \sigma P_t\,\mathrm{d}W_t^X, \end{aligned} \tag{4.38}$$

where α is the speed of mean-reversion to the mean level $f(t)$ (possibly capturing the seasonal component), σ is a constant volatility and $\mu(t) := \frac{1}{\alpha}(\frac{\sigma^2}{2} + f_t') + f(t)$. The logarithm of the price P_t follows a Gaussian process, i.e. for $w > t$ we have that $\ln P_w = (f(w) + X_w) \sim \mathcal{N}(\underbrace{f(w) + X_t e^{-\alpha(w-t)}}_{=:m_1}, \underbrace{\sigma^2(1 - e^{-2\alpha(w-t)})/(2\alpha)}_{=:m_2})$. By $\phi_t(x, y)$ we denote a probability density function of this process.

We assume the following correlation structure with

$$\mathrm{d}W_t^X\,\mathrm{d}W_t^S = \rho\,\mathrm{d}t. \tag{4.39}$$

We also assume a linear consistency on the correlation structure, particularly
if $\rho = \mathrm{corr}(X_t, S_t) < 0$, then $\hat{\rho} = \mathrm{corr}\big((X_t - x)\mathbb{1}_{X_t>x}, (S_t - s)\mathbb{1}_{S_t>s}\big) < 0$. To motivate this assumption, one can think of the following: if the storage facilities are relatively full or increasing (e.g., extra precipitation) and market is aware of the lack of a storable asset shortage, then the market power price would be relatively low or decreasing.

We introduce now the value process $V_t(P_t, S_t, C_t)$ as

$$V_t(P_t, S_t, C_t) = \mathbb{E}\Big[\int_t^T e^{-r(s,t)}\,H_s(P_s, S_s, C_s)\,\mathrm{d}s|\mathcal{F}_t\Big], \tag{4.40}$$

where $H_t(P_t, S_t, C_t)$ is a payoff including the various costs C_t. Simply speaking, we consider a value process as a discounted payoff which is a combination of two stochastic processes (fuel price and storage level). Since we know the statistical properties of these two processes, we aim to investigate their product process to have some approximation of the storage value.

Before we proceed with the investigation of various payoffs, we need to recall some technical properties of the process S_t. We can characterise it by the transition density function $p_{t-t_0}(x, y)$ derived in the previous section. We will use the following notations: $c_n = \frac{\Delta}{2}$,

$v_n(x) := \sqrt{\frac{2n+1}{2}} P_n\Big(\frac{2}{\Delta}(x-l)-1\Big)$ with $P_n(x)$ being Legendre's[1] series of order n and $C(n,x,t) := \frac{2n+1}{\Delta} P_n\Big(\frac{2}{\Delta}(x-l)-1\Big)e^{-n(n+1)(t-t_0)}$. Then the transition probability density function $p_{t-t_0}(x,y)$ is given as

$$\begin{aligned} p_{t-t_0}(x,y)\,|_{(l,u)} &= \sum_{n=0}^{\infty} \frac{v_n(x)v_n(y)}{c_n}\, e^{-n(n+1)(t-t_0)} \\ &= \sum_{n=0}^{\infty} \frac{2n+1}{\Delta}\, P_n\Big(\frac{2}{\Delta}(x-l)-1\Big) P_n\Big(\frac{2}{\Delta}(y-l)-1\Big)\, e^{-n(n+1)(t-t_0)} \\ &= \sum_{n=0}^{\infty} C(n,x,t) P_n\Big(\frac{2}{\Delta}(y-l)-1\Big). \end{aligned} \tag{4.41}$$

We will use the following properties of the Legendre series (for the details see Abramowitz & Stegun [1970], p. 786 and Bell [2004], pp. 56-58):

- if $f(z)$ is a polynomial with a degree less than $P_n(z)$ then

$$\int_{-1}^{1} f(z)P_n(z)\ \mathrm{d}z = 0; \tag{4.42}$$

- for $n \geq 1$

$$\int_{x}^{1} P_n(z)\ \mathrm{d}z = \frac{P_{n-1}(x) - P_{n+1}(x)}{2n+1}, \tag{4.43}$$

  we denote this quantity as $P^*_{n-1,n+1}(x)$ for future calculations;

- for $n \geq 2$

$$\int_{x}^{1} zP_n(z)\ \mathrm{d}z = \frac{n(2n+3)P_{n-2}(x) - (2n+1)P_n(x) - (n+1)(2n-1)P_{n+2}(x)}{(4n^2-1)(2n+3)}, \tag{4.44}$$

  we denote this quantity as $P^*_{n-2,n,n+2}(x)$ for future calculations;

- for $n \geq 3$

[1] There is a variety of literature on the Legendre polynomials available, for instance, Whittaker & Watson [1996] and Bell [2004]. The first few polynomials are: $P_0(x) = 1$, $P_1(x) = x$, $P_2(x) = \frac{3}{2}x^2 - \frac{1}{2}$, $P_3(x) = \frac{5}{2}x^3 - \frac{3}{2}x$, $P_4(x) = \frac{35}{8}x^4 - \frac{30}{8}x^2 + \frac{3}{8}$.

$$\begin{aligned}\int_x^1 z^2 P_n(z)\,\mathrm{d}z &= \frac{n(n-1)}{(4n^2-1)(2n-3)}P_{n-3}(x) - \frac{(n+1)(n+2)}{(2n+1)(2n+3)(2n+5)}P_{n+3}(x)\\ &- \frac{n^2+3n-1}{(4n^2-1)(2n+5)}P_{n+1}(x) + \frac{n^2-n-3}{(4n^2-9)(2n+1)}P_{n-1}(x), \qquad (4.45)\end{aligned}$$

we denote this quantity as $P^*_{n-3,n+3}(x)$ for future calculations;

- for even n

$$\int_0^1 z^2 P_n(z)\,\mathrm{d}z = \frac{(-1)^n(n-1)(3/2)}{2(-1)(n+5/2)}, \qquad (4.46)$$

we denote this quantity as P^*_{2n} for future calculations;

- for odd n

$$\int_0^1 z^2 P_n(z)\,\mathrm{d}z = \frac{(-1)^n(n-1/2)(2)}{2(n+3)(-1/2)}, \qquad (4.47)$$

we denote this quantity as P^*_{2n+1} for future calculations;

We will use the following expression and notation for the expected value of the process $(S_t - m)$

$$\begin{aligned}E_0(t,T) &:= \mathbb{E}\big[S_T - m|\mathcal{F}_t\big]\\ &= \mathbb{E}\big[S_T - m|S_t\big]\\ &= \int_l^u (S_T - m)\, p_{T-t}(S_t, S_T)\,\mathrm{d}S_T\\ &= \int_l^u (y - m)\, p_{T-t}(x, y)\,\mathrm{d}y\\ &= \int_l^u (y-m)\left(\sum_{n=0}^{\infty}\frac{2n+1}{\Delta}P_n\left(\frac{2}{\Delta}(y-l)-1\right)P_n\left(\frac{2}{\Delta}(x-l)-1\right)e^{-n(n+1)(T-t)}\right)\mathrm{d}y\\ &= \sum_{n=0}^{\infty}\frac{2n+1}{\Delta}P_n\left(\frac{2}{\Delta}(x-l)-1\right)e^{-n(n+1)(T-t)}\int_l^u (y-m)P_n\left(\frac{2}{\Delta}(y-l)-1\right)\,\mathrm{d}y\\ &= \sum_{n=0}^{\infty}\frac{2n+1}{\Delta}P_n\left(\frac{2}{\Delta}(x-l)-1\right)e^{-n(n+1)(T-t)}\frac{\Delta^2}{4}\int_{-1}^1 zP_n\big(z\big)\,\mathrm{d}z\\ &= \frac{\Delta}{4}\int_{-1}^1 z\,\mathrm{d}z + \frac{3\Delta}{4}P_1\left(\frac{2}{\Delta}(x-l)-1\right)e^{-2(T-t)}\int_{-1}^1 z^2\,\mathrm{d}z\end{aligned}$$

$$
\begin{aligned}
&+ \underbrace{\sum_{n=2}^{\infty} \frac{2n+1}{\Delta} P_n\left(\frac{2}{\Delta}(x-l)-1\right) e^{-n(n+1)(T-t)} \frac{\Delta^2}{4} \int_{-1}^{1} z\, P_n(z)\,\mathrm{d}z}_{=0, \text{ due to Equation (4.42)}} \\
&= (S_t - m) e^{-2(T-t)}, \qquad (4.48)
\end{aligned}
$$

where we use a substitution $z := \frac{2}{\Delta}(y-l)-1$.

We will use the following expression and notation for the variance of the process $(S_t - m)$

$$
\begin{aligned}
V_0(t,T) &:= \mathrm{Var}\big[S_T - m | \mathcal{F}_t\big] \\
&= \mathrm{Var}\big[S_T - m | S_t\big] \\
&= \mathbb{E}\big[(S_T-m)^2 | S_t\big] - \Big(\mathbb{E}\big[S_T - m | S_t\big]\Big)^2 \\
&= \int_l^u (S_T-m)^2\, p_{T-t}(S_t, S_T)\,\mathrm{d}S_T - (S_t-m)^2 e^{-4(T-t)} \\
&= \int_l^u (y-m)^2\, p_{T-t}(x,y)\,\mathrm{d}y - (S_t-m)^2 e^{-4(T-t)} \\
&= \int_l^u (y-m)^2 \Big(\sum_{n=0}^{\infty} \frac{2n+1}{\Delta} P_n\left(\frac{2}{\Delta}(y-l)-1\right) P_n\left(\frac{2}{\Delta}(x-l)-1\right) e^{-n(n+1)(T-t)}\Big)\,\mathrm{d}y \\
&- (S_t-m)^2 e^{-4(T-t)} \\
&= \sum_{n=0}^{\infty} \frac{2n+1}{\Delta} P_n\left(\frac{2}{\Delta}(x-l)-1\right) e^{-n(n+1)(T-t)} \frac{\Delta^3}{8} \int_{-1}^{1} z^2 P_n\big(z\big)\,\mathrm{d}z \\
&= \frac{\Delta^2}{8}\int_{-1}^{1} z^2\,\mathrm{d}z + \frac{3\Delta^2}{8} P_1\left(\frac{2}{\Delta}(x-l)-1\right) e^{-2(T-t)} \int_{-1}^{1} z^3\,\mathrm{d}z \\
&+ \frac{5\Delta^2}{32} P_2\left(\frac{2}{\Delta}(x-l)-1\right) e^{-6(T-t)} \int_{-1}^{1} z^2 P_2(z)\,\mathrm{d}z \\
&+ \underbrace{\sum_{n=3}^{\infty} \frac{2n+1}{\Delta} P_n\left(\frac{2}{\Delta}(x-l)-1\right) e^{-n(n+1)(T-t)} \frac{\Delta^3}{8} \int_{-1}^{1} z^2\, P_n(z)\,\mathrm{d}z}_{=0, \text{ due to Equation (4.42)}} \\
&- (S_t-m)^2 e^{-4(T-t)} \\
&= \frac{\Delta^2}{12} + e^{-6(T-t)}\Big((S_t-m)^2 - \frac{\Delta^2}{12}\Big) - (S_t-m)^2 e^{-4(T-t)} \\
&= (S_t-m)^2 e^{-4(T-t)}\Big(e^{-2(T-t)} - 1\Big) + \frac{\Delta^2}{12}\Big(1 - e^{-6(T-t)}\Big), \qquad (4.49)
\end{aligned}
$$

where we also substitute $z := \frac{2}{\Delta}(y-l)-1$.

We will use the following expression and notation for the expected value of P_t

$$E_1(t,T) := \mathbb{E}\big[P_T|\mathcal{F}_t\big] = \mathbb{E}\big[P_T|P_t\big] = \exp\Big(f(T) + X_t e^{-\alpha(T-t)} + \frac{\sigma^2}{4\alpha}(1-e^{-2\alpha(T-t)})\Big). \quad (4.50)$$

We will use the following expression and notation for the variance of the process P_t

$$\begin{aligned} V_1(t,T) &:= \text{Var}\big[P_T|\mathcal{F}_t\big] = \text{Var}\big[P_T|P_t\big] \\ &= \big(\mathbb{E}\big[P_T|P_t\big]\big)^2\Big(\exp\Big(\frac{\sigma^2}{2\alpha}\big(1-e^{-2\alpha(T-t)}\big)\Big) - 1\Big) \\ &= E_1^2(t,T)\Big(e^{\frac{\sigma^2}{2\alpha}\left(1-e^{-2\alpha(T-t)}\right)} - 1\Big). \end{aligned} \quad (4.51)$$

Now we are ready to consider various financial instruments associated with the storage level and the fuel price.

4.5.2 Probability measure

Before we proceed with pricing, we need to clarify some points on a pricing measure. From mathematical finance theory we know that in a complete market a contingent claim's price is the discounted expected value of the future payoff under the **equivalent martingale measure** $\mathbb{Q}$ different from a real-world pricing measure $\mathbb{P}$. However, the energy-related markets are incomplete, since due to specific market characteristics many payoffs cannot be replicated by other trading financial instruments. In our case the "spot price" is the storage level process which can be, for instance, considered as an index of current state reservoir level (hydro). Hence, we cannot think of $\mathbb{Q}$ being the martingale measure since the process S_t does not need to be a martingale under $\mathbb{Q}$. Instead, we can take **any** measure $\mathbb{Q}$ equivalent to the real-world measure $\mathbb{P}$, i.e. $\mathbb{Q} = \mathbb{P}$, and price derivatives respectively. So then this measure can be called as a **pricing** measure which is the probability measure that takes into account all the risk associated with maintaing the storage. In other words, we assume that the process S_t is already under pricing measure $\mathbb{Q} = \mathbb{P}$.

4.5.3 Simple financial products

Equipped with the Legendre series properties together with the expression for the transition probability density $p_{t-t_0}(x,y)$ in Equation (4.41), we now study some fundamental financial products: futures and options on a reservoir level. Since the storage level S_t at time t is a random process, these financial instruments gamble that the current reservoir level S_t rise or

fall above or below some level. They can also be used by the producer, retailer or market maker to hedge their risk when maintaining the storage or, for example, to speculate (since the power price can be quite volatile). There are various risks here associated with the storage maintenance. Consider, for instance, the case when the power price is high and the water level is low. Then our producer and/or storage owner does not have an opportunity to produce and has a potential loss. Another case is when the power price is low and the water level is high, a producer still bears the costs on maintaining the storage, but it is not profitable to produce power due to low power price level. Since the high power price volatility is a constant source of uncertainty and risk, the producer is willing to hedge against it, especially if she has the fixed price contracts.

Futures on the reservoir level

Under some pricing measure $\mathbb{Q} = \mathbb{P}$ we can due to Equation (4.48) write the futures prices on the water (or gas) level with maturity T as

$$\begin{aligned} F(t,T,m) &= \mathbb{E}[S_T - m|\mathcal{F}_t] = \mathbb{E}[S_T - m|S_t] = E_0(t,T), \\ F(t,T) &= \mathbb{E}[S_T|\mathcal{F}_t] = \mathbb{E}[S_T|S_t] = S_t e^{-2(T-t)} + m(1 - e^{-2(T-t)}). \end{aligned} \tag{4.52}$$

We can also price futures on the average water level over some period of time by considering

$$\begin{aligned} F(t,T_1,T_2) &= \frac{1}{T_2 - T_1}\,\mathbb{E}\Big[\int_{T_1}^{T_2}(S_u - m)\,\mathrm{d}u|\mathcal{F}_t\Big] \\ &= \frac{1}{T_2 - T_1}\int_{T_1}^{T_2}\mathbb{E}\Big[S_u - m|S_t\Big]\,\mathrm{d}u \\ &= \frac{1}{T_2 - T_1}\,\frac{1}{2}\,(S_t - m)\big(e^{-2(T_1-t)} - e^{-2(T_2-t)}\big). \end{aligned} \tag{4.53}$$

European options on the reservoir level

Let us continue with a European Call option and some strike K which can be interpreted as marginal cost for maintaining the reservoir

$$
\begin{aligned}
C(t,T) &= \mathbb{E}^Q[\max\{S_T - K, 0\}|\mathcal{F}_t] \\
&= \mathbb{E}^Q[\max\{S_T - K, 0\}|S_t] \\
&= \int_K^u (y-K) p_{T-t}(x,y)\ \mathrm{d}y \\
&= \int_K^u (y-K) p_{T-t}(x,y)\ \mathrm{d}y \\
&= \int_K^u (y-K)\Big(\sum_{n=0}^{\infty} \frac{2n+1}{\Delta} P_n\left(\frac{2}{\Delta}(y-l)-1\right) P_n\left(\frac{2}{\Delta}(x-l)-1\right) e^{-n(n+1)(T-t)}\Big)\,\mathrm{d}y \\
&= \sum_{n=0}^{\infty} \frac{2n+1}{\Delta} P_n\left(\frac{2}{\Delta}(x-l)-1\right) e^{-n(n+1)(T-t)} \int_K^u (y-K) P_n\left(\frac{2}{\Delta}(y-l)-1\right)\ \mathrm{d}y \\
&= \sum_{n=0}^{\infty} \frac{2n+1}{\Delta} P_n\left(\frac{2}{\Delta}(x-l)-1\right) e^{-n(n+1)(T-t)} \int_{\tilde{K}}^1 \frac{\Delta}{2}\Big(\frac{\Delta}{2}z + (m-K)\Big) P_n(z)\ \mathrm{d}z \\
&= \frac{1}{\Delta}\Big(\frac{\Delta^2}{4}\frac{1-\tilde{K}^2}{2} + \frac{\Delta}{2}(m-K)(1-\tilde{K})\Big) \\
&+ \frac{3}{\Delta} P_1\left(\frac{2}{\Delta}(S_t-l)-1\right) e^{-2(T-t)}\Big(\frac{\Delta^2}{4}\frac{1-\tilde{K}^3}{3} + \frac{\Delta}{2}(m-K)\frac{1-\tilde{K}^2}{2}\Big) \\
&+ \sum_{n=2}^{\infty} \frac{2n+1}{\Delta} P_n\left(\frac{2}{\Delta}(S_t-l)-1\right) e^{-n(n+1)(T-t)}\Big(\frac{\Delta^2}{4} P^*_{n-2,n,n+2}(\tilde{K}) + \frac{\Delta}{2}(m-K) P^*_{n-1,n+1}(\tilde{K})\Big),
\end{aligned}
\tag{4.54}
$$

where we made the replacements $z := \frac{2}{\Delta}(y-l)-1$ and $\tilde{K} := \frac{2}{\Delta}(K-l)-1$. Expressions for $P^*_{n-1,n+1}(x)$ and $P^*_{n-2,n,n+2}(x)$ are given above in Equations (4.43) and (4.44) respectively.

Analogously, a European Put option can be computed.

4.5.4 Hydro-driven power plant

In this section we consider payoffs which can be used to find the value of a hydro-driven power plant and study its properties in a similar manner as in Chapter 5. In general, we consider a producer who wants to hedge against some unfavourable situations such as too low a water level in the reservoir, and too low or high prices. Therefore, such a producer could be interested in an option with which she can hedge against both water levels and price, as low water does not necessarily lead to high prices, only if demand is very high at the same time. Our producer might have contracts that she needs to fulfill with fixed prices, and thereby is concerned with too high a price or too low a water level. But too low a water level and too low a price may be connected with above average temperatures, and then the producer does not risk that much since she does not need to retail much power anyway. Since simple financial products only take into account

the current level in the storage facility, it is not enough to hedge against various complex cases, for these one needs to have some advanced financial products with more sophisticated payoffs. These payoffs are similar to the quanto options considered in Benth *et al.* [2013].

Payoff 1

Consider a hydro-driven power plant and a payoff that includes an average power price level M and average storage level m, namely

$$H_t(P_t, S_t, C_t) = \max\{P_t - M, 0\} \times \max\{S_t - m, 0\} - C_t. \tag{4.55}$$

Since the hydro reservoir depends on the natural inflow and we cannot "inject" any water additionally, this payoff has the following interpretation:

- **Case 1:** $P_t > M$ **(power prices are relatively high) and** $S_t > m$ **(reservoir is relatively full).** This is the most favourable situation which results in a positive value that an owner can have by releasing some water, producing power and selling it at the market.
- **Case 2:** $P_t > M$ **(power prices are relatively high) and** $S_t < m$ **(reservoir is relatively empty).** Here an owner does not have much water in the reservoir to produce power, though she has to still keep the storage maintenance.

At time t we find that

$$\begin{aligned}
V_t(P_t, S_t, C_t) &= \mathbb{E}\Big[\int_t^T e^{-r(w,t)}\, H_w(P_w, S_w, C_w)\,\mathrm{d}w|\mathcal{F}_t\Big] \\
&= \int_t^T e^{-r(w,t)} \Big(\mathbb{E}\big[H_w(P_w, S_w)|\mathcal{F}_t\big]\Big)\,\mathrm{d}w \\
&= \int_t^T e^{-r(w,t)} \Big(\mathbb{E}\big[\max\{P_w - M, 0\} \times \max\{S_w - m, 0\}|\mathcal{F}_t\big] - C_w\Big)\,\mathrm{d}w \\
&= \int_t^T e^{-r(w,t)} \Big(\underbrace{\mathbb{E}\big[\max\{P_w - M, 0\}|\mathcal{F}_t\big]}_{=:E_3(t,w)} \underbrace{\mathbb{E}\big[\max\{S_w - m, 0\}|\mathcal{F}_t\big]}_{=:E_2(t,w)} \\
&+ \rho_1 \sqrt{\underbrace{\mathrm{Var}(\max\{P_w - M, 0\}|\mathcal{F}_t)}_{=:V_3(t,w)} \underbrace{\mathrm{Var}(\max\{S_w - m, 0\}|\mathcal{F}_t)}_{=:V_2(t,w)}} - C_w\Big)\,\mathrm{d}w \\
&= \int_t^T e^{-r(w,t)} \Big(E_3(t,w) \cdot E_2(t,w) + \rho_1 \cdot \sqrt{V_3(t,w) \cdot V_2(t,w)} - C_w\Big)\,\mathrm{d}w,
\end{aligned} \tag{4.56}$$

where $\rho_1 = \mathrm{corr}(\max\{P_w - M, 0\}, \max\{S_w - m, 0\})$.

Let us now compute the values of E_2, E_3, V_2 and V_3. So, we obtain

$$\begin{aligned} E_3(t,w) &:= \mathbb{E}\big[\max\{P_w - M, 0\}|\mathcal{F}_t\big] \\ &= \mathbb{E}\big[\max\{e^{f(w)+X_w} - M, 0\}|P_t\big] \\ &= \int_{\ln M}^{\infty} (e^y - M)\phi_t(x,y)\ \mathrm{d}y \\ &= e^{m_1+\frac{m_2}{2}}\Phi\Big(\frac{m_1+m_2-\ln M}{\sqrt{m_2}}\Big) - M\Phi\Big(\frac{m_1-\ln M}{\sqrt{m_2}}\Big) \\ &= e^{m_1+\frac{m_2}{2}}\Phi(d_2) - M\Phi(d_1), \end{aligned}$$

where $d_1 := \frac{f(w)+X_t e^{-\alpha(w-t)}-\ln M}{\sqrt{\sigma^2/(2\alpha)(1-e^{-2\alpha(w-t)})}}$ and $d_2 := d_1 + \sqrt{\sigma^2/(2\alpha)(1-e^{-2\alpha(w-t)})}$.

Further, knowing the transition probability density function $p_{t-t_0}(x,y)$ for the process $S_t \in (l,u)$, we obtain for $w > t$

$$\begin{aligned} E_2(t,w) &:= \mathbb{E}\big[\max\{S_w - m, 0\}|\mathcal{F}_t\big] \\ &= \mathbb{E}\big[\max\{S_w - m, 0\}|S_t\big] \\ &= \int_m^u (y-m)\, p_{w-t}(x,y)\ \mathrm{d}y \\ &= \int_m^u (y-m) \sum_{n=0}^{\infty} \frac{2n+1}{\Delta} P_n\left(\frac{2}{\Delta}(x-l)-1\right) P_n\left(\frac{2}{\Delta}(y-l)-1\right) e^{-n(n+1)(w-t)}\ \mathrm{d}y \\ &= \sum_{n=0}^{\infty} \frac{2n+1}{\Delta} P_n\left(\frac{2}{\Delta}(x-l)-1\right) e^{-n(n+1)(w-t)} \int_m^u (y-m) P_n\left(\frac{2}{\Delta}(y-l)-1\right)\ \mathrm{d}y \\ &= \sum_{n=0}^{\infty} \frac{2n+1}{\Delta} P_n\left(\frac{2}{\Delta}(x-l)-1\right) e^{-n(n+1)(w-t)}\, \frac{\Delta^2}{4} \int_0^1 z P_n(z)\ \mathrm{d}z \\ &= \frac{\Delta}{8} + \frac{\Delta}{4} P_1\left(\frac{2}{\Delta}(x-l)-1\right) e^{-2(w-t)} \\ &+ \sum_{n=2}^{\infty} \frac{(2n+1)\Delta}{4} P_n\left(\frac{2}{\Delta}(x-l)-1\right) e^{-n(n+1)(w-t)}\, P^*_{n-2,n,n+2}(0) \\ &= \frac{\Delta}{8} + \frac{\Delta}{4} P_1\left(\frac{2}{\Delta}(S_t-l)-1\right) e^{-2(w-t)} \\ &+ \sum_{n=2}^{\infty} \frac{(2n+1)\Delta}{4} P_n\left(\frac{2}{\Delta}(S_t-l)-1\right) e^{-n(n+1)(w-t)}\, P^*_{n-2,n,n+2}(0), \end{aligned} \tag{4.57}$$

where the expression for $P^*_{n-2,n,n+2}(0)$ is given in Equation (4.44). Now we compute the value of V_3

$$
\begin{aligned}
V_3(t,w) &:= \mathrm{Var}(\max\{P_w - M, 0\}|\mathcal{F}_t) \\
&= \mathrm{Var}(\max\{e^{f(w)+X_w} - M, 0\}|P_t) \\
&= \int_{\ln M}^{\infty} (e^y - M)^2 \phi_t(x,y)\ \mathrm{d}y - \Big(\int_{\ln M}^{\infty} (e^y - M)\phi_t(x,y)\ \mathrm{d}y \Big)^2 \\
&= \int_{\ln M}^{\infty} e^{2y}\phi_t(x,y)\ \mathrm{d}y - 2M \int_{\ln M}^{\infty} e^y \phi_t(x,y)\ \mathrm{d}y \\
&+ M^2 \int_{\ln M}^{\infty} \phi_t(x,y)\ \mathrm{d}y - E_3^2(t,w) \\
&= e^{2(m_1+m_2)}\Phi(d_3) - 2Me^{m_1+\frac{m_2}{2}}\Phi(d_2) + M^2\Phi(d_1) - E_3^2(t,w),
\end{aligned}
\tag{4.58}
$$

where d_1 and d_2 are given above and $d_3 := d_1 + 2\sqrt{\sigma^2/(2\alpha)(1-e^{-2\alpha(w-t)})}$.

And finally we compute the value of V_2

$$
\begin{aligned}
V_2(t,w) &:= \mathrm{Var}(\max\{S_w - m, 0\}|\mathcal{F}_t) \\
&= \int_m^u (y-m)^2 p_{t-w}(x,y)\ \mathrm{d}y - E_2^2(t,w) \\
&= \sum_{n=0}^{\infty} \frac{2n+1}{\Delta} P_n\left(\frac{2}{\Delta}(x-l)-1\right) e^{-n(n+1)(w-t)} \int_m^u (y-m)^2 P_n\left(\frac{2}{\Delta}(y-l)-1\right)\ \mathrm{d}y - E_2^2(t,w) \\
&= \sum_{n=0}^{\infty} \frac{2n+1}{\Delta} P_n\left(\frac{2}{\Delta}(x-l)-1\right) e^{-n(n+1)(w-t)}\ \frac{\Delta^3}{8} \int_0^1 z^2 P_n(z)\ \mathrm{d}z - E_2^2(t,w), \\
&= \sum_{n=0}^{\infty} \frac{(2n+1)\Delta^2}{8} P_n\left(\frac{2}{\Delta}(x-l)-1\right) e^{-n(n+1)(w-t)}\ \{P_{2n}^*, P_{2n+1}^*\} - E_2^2(t,w),
\end{aligned}
\tag{4.59}
$$

where P_{2n}^* and P_{2n+1}^* are given in Equations (4.46) and (4.47) respectively.

Payoff 2

Let us the modify the payoff of the previous section and introduce an extra term responsible for the power production rate

$$\kappa(S_t) = \frac{1}{2} + \frac{S_t - m}{u - l} = \frac{S_t - l}{u - l},$$

$$H_t(P_t, S_t, C_t) = \max\{P_t - M, 0\} \times \kappa(S_t) \times S_t - C_t. \tag{4.60}$$

The difference to the previous payoff is that in this case we can produce at the rate κ which is greater than 50% if the $S_t > m$. There is the following interpretation for this payoff allowing for more flexibility in the production rate compared to the previous one:

- **Case 1: $P_t > M$ (power prices are relatively high) and $S_t > m$ (reservoir is relatively full).** This is the most favourable situation which results in a positive value that a storage owner can realise by releasing some water, producing power at the rate $\kappa(S_t)$ and selling it at the market.

- **Case 2: $P_t > M$ (power prices are relatively high) and $S_t < m$ (reservoir is relatively empty).** There is not much water available in the reservoir to produce intensively, but since the prices are high a producer would not like to completely stop operating, instead she has an option to produce at some rate at least.

At time t we find that

$$
\begin{aligned}
V_t(P_t, S_t, C_t) &= \mathbb{E}\Big[\int_t^T e^{-r(w,t)}\, H_w(P_w, S_w, C_t)\,\mathrm{d}w|\mathcal{F}_t\Big] \\
&= \int_t^T e^{-r(w,t)} \Big(\mathbb{E}\big[H_w(P_w, S_w)|\mathcal{F}_t\big]\Big)\,\mathrm{d}w \\
&= \int_t^T e^{-r(w,t)} \Big(\mathbb{E}\big[\max\{P_w - M, 0\} \times \kappa(S_w) \times S_w|\mathcal{F}_t\big] - C_w\Big)\,\mathrm{d}w \\
&= \int_t^T e^{-r(w,t)} \Big(\mathbb{E}\Big[\max\{P_w - M, 0\} \times \frac{S_w(S_w - l)}{u - l}|\mathcal{F}_t\Big] - C_w\Big)\,\mathrm{d}w \\
&= \int_t^T e^{-r(w,t)} \Big(\underbrace{\mathbb{E}\big[\max\{P_w - M, 0\}|\mathcal{F}_t\big]}_{=:E_3(t,w)}\underbrace{\mathbb{E}\Big[\frac{S_w(S_w - l)}{u - l}|\mathcal{F}_t\Big]}_{=:E_4(t,w)} \\
&+ \rho_2 \sqrt{\underbrace{\mathrm{Var}(\max\{P_w - M, 0\}|\mathcal{F}_t)}_{=:V_3(t,w)}\underbrace{\mathrm{Var}\Big(\frac{S_w(S_w - l)}{u - l}|\mathcal{F}_t\Big)}_{=:V_4(t,w)}} - C_w\Big)\,\mathrm{d}w \\
&= \int_t^T e^{-r(w,t)} \Big(E_3(t,w) \cdot E_4(t,w) + \rho_2 \cdot \sqrt{V_3(t,w) \cdot V_4(t,w)} - C_w\Big)\,\mathrm{d}w,
\end{aligned}
\tag{4.61}
$$

where $\rho_2 = \mathrm{corr}(\max\{P_w - M, 0\}, \frac{S_w(S_w - l)}{u - l})$.

From above we know the values of E_3 and V_3. Let us compute the values of E_4 and V_4. We start with

$$\begin{aligned}
E_4(t,w) &:= \mathbb{E}\Big[\frac{S_w(S_w-l)}{u-l}\big|\mathcal{F}_t\Big] \\
&= \frac{1}{\Delta}\mathbb{E}\big[S_w^2|S_t\big] - \frac{l}{\Delta}\mathbb{E}\big[S_w|S_t\big] \\
&= \frac{1}{\Delta}\mathbb{E}\big[(S_w-m)^2|S_t\big] + \frac{u}{\Delta}\mathbb{E}\big[S_w-m|S_t\big] + \frac{m}{2} \\
&= \frac{1}{\Delta}\int_l^u (S_w-m)^2 p_{w-t}(S_t,S_w)\,\mathrm{d}S_w + \frac{u}{\Delta}E_0(t,w) + \frac{m}{2} \\
&= \frac{1}{\Delta}\int_l^u (y-m)^2 p_{w-t}(x,y)\,\mathrm{d}y + \frac{u}{\Delta}E_0(t,w) + \frac{m}{2} \\
&= \frac{1}{\Delta}\Big(\frac{\Delta^2}{12} + \big((x-m)^2 - \frac{\Delta^2}{12}\big)e^{-6(w-t)}\Big) + \frac{4}{\Delta}(x-m)e^{-2(w-t)} + \frac{m}{2} \\
&= \Big(\frac{(x-m)^2}{\Delta} - \frac{\Delta}{12}\Big)e^{-6(w-t)} + \frac{4}{\Delta}(x-m)e^{-2(w-t)} + \frac{2u-l}{6} \\
&= \Big(\frac{(S_t-m)^2}{\Delta} - \frac{\Delta}{12}\Big)e^{-6(w-t)} + \frac{4}{\Delta}(S_t-m)e^{-2(w-t)} + \frac{2u-l}{6}. \qquad (4.62)
\end{aligned}$$

Now we continue with the value of V_4

$$\begin{aligned}
V_4(t,w) &:= \mathrm{Var}\Big(\frac{S_w(S_w-l)}{u-l}\big|\mathcal{F}_t\Big) \\
&= \frac{1}{\Delta^2}Var(S_w^2 - lS_w\big|S_t) \\
&= \frac{1}{\Delta^2}\int_l^u (S_w^2 - lS_w)^2 p_{w-t}(S_t,S_w)\,\mathrm{d}S_w - E_4^2(t,w) \\
&= \frac{1}{\Delta^2}\int_l^u y^4\, p_{w-t}(x,y)\,\mathrm{d}y - \frac{2l}{\Delta^2}\int_l^u y^3\, p_{w-t}(x,y)\,\mathrm{d}y + \frac{l^2}{\Delta^2}\int_l^u y^2\, p_{w-t}(x,y)\,\mathrm{d}y - E_4^2(t,w).
\end{aligned} \qquad (4.63)$$

The first three terms mainly involve the following expression for some a

$$y(a,n) := \int_l^u y^a P_n\Big(\frac{2}{\Delta}(y-l)-1\Big)\,\mathrm{d}y = \frac{\Delta}{2}\int_{-1}^1 \Big(\frac{\Delta}{2}(z+1)+l\Big)^a P_n(z)\,\mathrm{d}z = 0, \qquad (4.64)$$

when $a < n$. So we need to compute the following integrals to finish with Equation (4.63)

$$y(2,0) := \int_l^u y^2 P_0\Big(\frac{2}{\Delta}(y-l)-1\Big)\,\mathrm{d}y = \frac{u^3-l^3}{3}, \qquad (4.65)$$

$$y(2,1) := \int_l^u y^2 P_1\Big(\frac{2}{\Delta}(y-l)-1\Big)\,\mathrm{d}y = \frac{(u-l)^2}{6}(u+l), \qquad (4.66)$$

$$y(2,2) := \int_l^u y^2 P_2\Big(\frac{2}{\Delta}(y-l)-1\Big)\,\mathrm{d}y = \frac{(u-l)^3}{30}, \tag{4.67}$$

$$y(3,0) := \int_l^u y^3 P_0\Big(\frac{2}{\Delta}(y-l)-1\Big)\,\mathrm{d}y = \frac{u^4-l^4}{4}, \tag{4.68}$$

$$y(3,1) := \int_l^u y^3 P_1\Big(\frac{2}{\Delta}(y-l)-1\Big)\,\mathrm{d}y = \frac{(u-l)^2}{20}(3u^2+4lu+3l^2), \tag{4.69}$$

$$y(3,2) := \int_l^u y^3 P_2\Big(\frac{2}{\Delta}(y-l)-1\Big)\,\mathrm{d}y = \frac{(u-l)^3}{20}(u+l), \tag{4.70}$$

$$y(3,3) := \int_l^u y^3 P_3\Big(\frac{2}{\Delta}(y-l)-1\Big)\,\mathrm{d}y = \frac{(u-l)^4}{140}, \tag{4.71}$$

$$y(4,0) := \int_l^u y^4 P_0\Big(\frac{2}{\Delta}(y-l)-1\Big)\,\mathrm{d}y = \frac{u^5-l^5}{5}, \tag{4.72}$$

$$y(4,1) := \int_l^u y^4 P_1\Big(\frac{2}{\Delta}(y-l)-1\Big)\,\mathrm{d}y = \frac{(u-l)^2}{15}(2u^3+3lu^2+3ul^2+2l^3), \tag{4.73}$$

$$y(4,2) := \int_l^u y^4 P_2\Big(\frac{2}{\Delta}(y-l)-1\Big)\,\mathrm{d}y = \frac{(u-l)^3}{35}(2u^2+3lu+2l^2), \tag{4.74}$$

$$y(4,3) := \int_l^u y^4 P_3\Big(\frac{2}{\Delta}(y-l)-1\Big)\,\mathrm{d}y = \frac{(u-l)^4}{70}(u+l), \tag{4.75}$$

$$y(4,4) := \int_l^u y^4 P_4\Big(\frac{2}{\Delta}(y-l)-1\Big)\,\mathrm{d}y = \frac{(u-l)^5}{630}. \tag{4.76}$$

Then coming back to Equation (4.63) we obtain

$$\begin{aligned}
\ldots &= \frac{1}{\Delta^2}\int_l^u y^4 p_{w-t}(x,y)\,\mathrm{d}y - \frac{2l}{\Delta^2}\int_l^u y^3 p_{w-t}(x,y)\,\mathrm{d}y + \frac{l^2}{\Delta^2}\int_l^u y^2 p_{w-t}(x,y)\,\mathrm{d}y - E_4^2(t,w) \\
&= C(4,w-t,x)\frac{1}{\Delta^2}y(4,4) \\
&+ C(3,w-t,x)\Big(\frac{1}{\Delta^2}y(4,3) - \frac{2l}{\Delta^2}y(3,3)\Big)
\end{aligned}$$

$$
\begin{aligned}
&+ \quad C(2, w-t, x)\Big(\frac{1}{\Delta^2}y(4,2) - \frac{2l}{\Delta^2}y(3,2) + \frac{l^2}{\Delta^2}y(2,2)\Big) \\
&+ \quad C(1, w-t, x)\Big(\frac{1}{\Delta^2}y(4,1) - \frac{2l}{\Delta^2}y(3,1) + \frac{l^2}{\Delta^2}y(2,1)\Big) \\
&+ \quad C(0, w-t, x)\Big(\frac{1}{\Delta^2}y(4,0) - \frac{2l}{\Delta^2}y(3,0) + \frac{l^2}{\Delta^2}y(2,0)\Big) - E_4^2(t,w) \\
&= \quad \frac{1}{\Delta^2} \times \begin{pmatrix} C(4, w-t, x) \\ C(3, w-t, x) \\ C(2, w-t, x) \\ C(1, w-t, x) \\ C(0, w-t, x) \end{pmatrix}^T \times \begin{pmatrix} y(4,4) & 0 & 0 \\ y(4,3) & y(3,3) & 0 \\ y(4,2) & y(3,2) & y(2,2) \\ y(4,1) & y(3,1) & y(2,1) \\ y(4,0) & y(3,0) & y(2,0) \end{pmatrix} \times \begin{pmatrix} 1 \\ -2l \\ l^2 \end{pmatrix} - E_4^2(t,w),
\end{aligned}
\tag{4.77}
$$

where $C(n, x, w-t) := \frac{2n+1}{\Delta} P_n\Big(\frac{2}{\Delta}(x-l) - 1\Big)e^{-n(n+1)(w-t)}$.

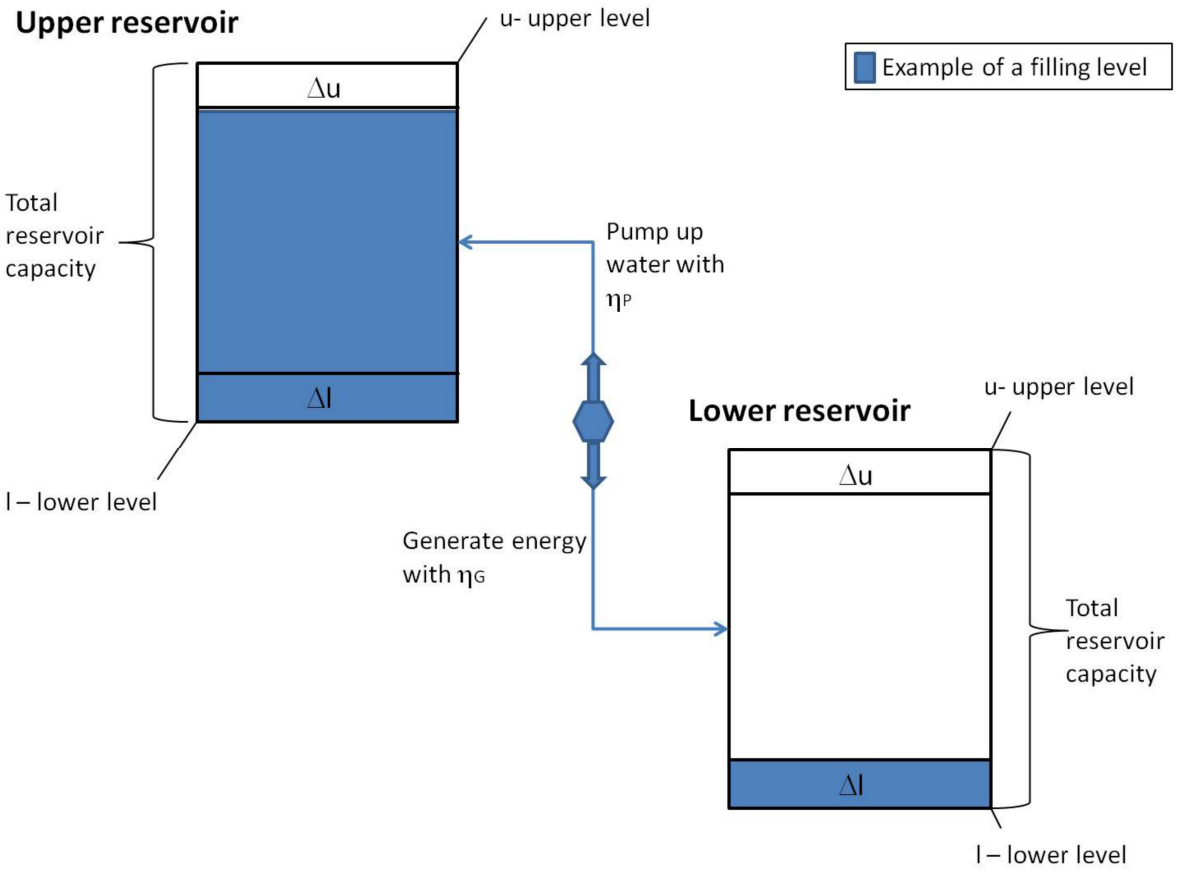


Figure 4.6: Pumped storage reservoir.

Payoff 3

Let us now consider the hydro power station with two reservoirs R_1 and R_2. The scheme of this pumped storage example is given in Figure 4.6[1]. We make the following assumptions:

- the inflow to R_1 is random, since it depends on precipitation and thaw;
- no other water inflow into the reservoirs is possible;
- by injection water from R_1 to R_2 we produce power and by pumping water up from R_2 to R_1 we refill R_1 for our future production purposes when needed and possible;
- it takes more energy to pump water up than to produce energy;
- both reservoirs have the same capacity of (l, u);
- our hydro-driven power plant contains water of one full reservoir total capacity;

Our assumptions yield that $u - S_t^1 = S_t^2 - l$. Then for some levels $l < K_1 \leq m \leq K_2 < u$ in R_1 and R_2 we construct the payoff

$$\begin{aligned} H_t(P_t, S_t^1, S_t^2, C_t) &= \max\{P_t - M^+, 0\} \max\{S_t^1 - K_1, 0\} - \max\{M^- - P_t, 0\} \max\{S_t^2 - K_2, 0\} - C_t \\ &= \max\{P_t - M^+, 0\} \max\{S_t^1 - K_1, 0\} - \max\{M^- - P_t, 0\} \max\{K_1 - S_t^1, 0\} - C_t \\ &= H_t(P_t, S_t^1, C_t) \end{aligned} \tag{4.78}$$

Due to physical reasons we assume that pumping water up needs more energy than producing electricity. There exists a quantity ΔP such that with $M^- = M - \Delta P$ and $M^+ = M + \Delta P$ we define an interval (M^-, M^+) such that for $P_t \in (M^-, M^+)$ it is not efficient to generate or to buy electricity. This ΔP can be computed via an average price M and efficiency rates of pumping and generating. The details on calculation of the ΔP are given in Connolly *et al.* [2011].

This payoff has the following interpretation

- **Case 1:** $P_t > M^+$ **(power prices are relatively high) and** $S_t^1 > K_1$ **(reservoir** R_1 **is relatively full which implies that reservoir** R_2 **is relatively empty).** This is the most favourable situation which results in a positive value that an owner realises by releasing some water into reservoir R_2, producing power and selling it at the market.
- **Case 2:** $P_t > M^+$ **(power prices are relatively high) and** $S_t^1 < K_1$ **(reservoir** R_1 **is relatively empty which implies that reservoir** R_2 **is relatively full).** This situation is

[1]This picture has been generated by the master student of our chair Elisabeth Tropp.

quite unfavourable, since to produce power we first need to pump the water up from R_2 to R_1. To do so, we need to buy power, but since the prices are high we may want to suspend the activity for a while.

- **Case 3: $P_t < M^-$ (power prices are relatively low) and $S_t^2 > K_2$ (reservoir R_1 is relatively empty which implies that reservoir R_2 is relatively full).** This is also a quite favourable situation for an owner, since she can buy power at a relatively low price and pump water immediately up to get the reservoir R_1 full.

- **Case 4: $P_t < M^-$ (power prices are relatively high) and $S_t^2 < K_1$ (reservoir R_1 is relatively full which implies that reservoir R_2 is relatively empty).** Here our producer is not interested in buying power since despite the price level the reservoir R_1 is full.

At time t we find that

$$
\begin{aligned}
V_t(P_t, S_t^1, S_t^2, C_t) &= \mathbb{E}\Big[\int_t^T e^{-r(w,t)}\, H_w(P_w, S_w^1, S_t^2, C_w)\,\mathrm{d}w|\mathcal{F}_t\Big] \\
&= \int_t^T e^{-r(w,t)} \Big(\mathbb{E}\big[H_w(P_w, S_w^1, S_w^2, C_w)|\mathcal{F}_t\big]\Big)\,\mathrm{d}w \\
&= \int_t^T e^{-r(w,t)} \Big(\mathbb{E}\big[\max\{P_w - M^+, 0\} \times \max\{S_w^1 - K_1, 0\} \\
&- \max\{M^- - P_w, 0\} \times \max\{S_w^2 - K_2, 0\}|\mathcal{F}_t\big] - C_w\Big)\,\mathrm{d}w \\
&= \int_t^T e^{-r(w,t)} \Big(\mathbb{E}\big[\max\{P_w - M^+, 0\} \times \max\{S_w^1 - K_1, 0\} \\
&- \max\{M^- - P_w, 0\} \times \max\{K_1 - S_w^1, 0\}|\mathcal{F}_t\big] - C_w\Big)\,\mathrm{d}w \\
&= \int_t^T e^{-r(w,t)} \Big(\underbrace{\mathbb{E}\big[\max\{P_w - M^+, 0\}|\mathcal{F}_t\big]}_{=:\tilde{E}_3(t,w)}\, \underbrace{\mathbb{E}\big[\max\{S_w^1 - K_1, 0\}|\mathcal{F}_t\big]}_{=:E_8(t,w)} \\
&- \underbrace{\mathbb{E}\big[\max\{M^- - P_w, 0\}|\mathcal{F}_t\big]}_{=:\tilde{E}_5(t,w)}\, \underbrace{\mathbb{E}\big[\max\{K_1 - S_w^1, 0\}|\mathcal{F}_t\big]}_{=:E_{10}(t,w)} \\
&+ \rho_3\Big(\sqrt{\underbrace{\mathsf{Var}(\max\{P_w - M^+, 0\}|\mathcal{F}_t)}_{=:\tilde{V}_3(t,w)}\, \underbrace{\mathsf{Var}(\max\{S_w^1 - K_1, 0\}|\mathcal{F}_t)}_{=:V_8(t,w)}} \\
&- \sqrt{\underbrace{\mathsf{Var}(\max\{M^- - P_w, 0\}|\mathcal{F}_t)}_{=:\tilde{V}_5(t,w)}\, \underbrace{\mathsf{Var}(\max\{K_1 - S_w^1, 0\}|\mathcal{F}_t)\Big)}_{=:V_{10}(t,w)}} - C_w\Big)\,\mathrm{d}w \\
&= \int_t^T e^{-r(w,t)} \Big(\tilde{E}_3(t,w) \cdot E_8(t,w) - \tilde{E}_5(t,w) \cdot E_{10}(t,w) \\
&+ \rho_3\big(\sqrt{\tilde{V}_3(t,w) \cdot V_8(t,w)} - \sqrt{\tilde{V}_5(t,w) \cdot V_{10}(t,w)}\big) - C_w\Big),\,\mathrm{d}w. \qquad (4.79)
\end{aligned}
$$

where $\rho_3 = \text{corr}(\max\{P_w - M^+, 0\}, \max\{S_w^1 - K_1, 0\}) = \text{corr}(\max\{M^- - P_w, 0\}, \max\{S_w^2 - K_2, 0\})$.

The expressions for $\tilde{E}_3$ and $\tilde{V}_3$ can be obtained from E_3 and V_3 given above by replacing M by M^+. The same way the expressions for $\tilde{E}_5$ and $\tilde{V}_5$ can be obtained from E_5 and V_5 given below by replacing M by M^-. Now let us compute the values of E_8, E_{10}, V_8 and V_{10}. We start with the value of E_8 which is analogous to the value in Equation (4.54)

$$\begin{aligned}
E_8(t,w) &:= \mathbb{E}\big[\max\{S_w^1 - K_1, 0\}|\mathcal{F}_t\big] \\
&= \int_{K_1}^{u} (y - K_1) p_{t-w}(y, x)\, \mathrm{d}y \\
&= \frac{1}{\Delta}\Big(\frac{\Delta^2}{4}\frac{1-\tilde{K}_1^2}{2} + \frac{\Delta}{2}(m - K_1)(1 - \tilde{K})\Big) \\
&+ \frac{3}{\Delta}P_1\big(\frac{2}{\Delta}(S_t^1 - l) - 1\big)e^{-2(T-t)}\Big(\frac{\Delta^2}{4}\frac{1-\tilde{K}_1^3}{3} + \frac{\Delta}{2}(m - K_1)\frac{1-\tilde{K}_1^2}{2}\Big) \\
&+ \sum_{n=2}^{\infty}\frac{2n+1}{\Delta}P_n\left(\frac{2}{\Delta}(S_t^1 - l) - 1\right)e^{-n(n+1)(T-t)}\Big(\frac{\Delta^2}{4}P^*_{n-2,n,n+2}(\tilde{K}_1) \\
&\qquad + \frac{\Delta}{2}(m - K_1)P^*_{n-1,n+1}(\tilde{K})\Big),
\end{aligned} \tag{4.80}$$

where $\tilde{K}_1 := \frac{2}{\Delta}(K_1 - l) - 1$ and expressions for $P^*_{n-2,n,n+2}(x)$ and $P^*_{n-1,n+1}(x)$ are given above in Equations (4.44) and (4.43) respectively. Then

$$\begin{aligned}
E_{10}(t,w) &:= \mathbb{E}\big[\max\{S_w^2 - K_2, 0\}|\mathcal{F}_t\big] \\
&= \mathbb{E}\big[\max\{K_1 - S_w^1, 0\}|S_t\big] \\
&= K_1 + m - E_0(t,w) - E_8(t,w).
\end{aligned} \tag{4.81}$$

We continue with the values of V_8 and V_{10}. When $K_1 = m$ the case is identical to V_2, but here we assume that K_1 is different from m, then

$$\begin{aligned}
V_8(t,w) &:= \text{Var}\big(\max\{S_w^1 - K_1, 0\}|\mathcal{F}_t\big) \\
&= \int_{K_1}^{u}(y - K_1)^2 p_{t-w}(x, y)\, \mathrm{d}y - E_8^2(t,w) \\
&= \sum_{n=0}^{\infty}\frac{2n+1}{\Delta}P_n\left(\frac{2}{\Delta}(x - l) - 1\right)e^{-n(n+1)(w-t)}\int_{K_1}^{u}(y - K_1)^2 P_n\left(\frac{2}{\Delta}(y - l) - 1\right)\mathrm{d}y
\end{aligned}$$

$$- \quad E_8^2(t,w),$$

(4.82)

now continue with computing an integral term and substituting $z := \frac{2}{\Delta}(y-l)-1$ and $\tilde{K_1} := \frac{2}{\Delta}(K_1-l)-1$ gives

$$\begin{aligned}
\int_{K_1}^{u}(y-K_1)^2 P_n\left(\frac{2}{\Delta}(y-l)-1\right)\,\mathrm{d}y &= \frac{\Delta}{2}\int_{\tilde{K_1}}^{1}\Big(\frac{\Delta}{2}z+(m-K_1)\Big)^2 P_n(z)\,\mathrm{d}z \\
&= \frac{\Delta^3}{8}\int_{\tilde{K_1}}^{1} z^2 P_n(z)\,\mathrm{d}z + \frac{\Delta^2}{2}(m-K_1)\int_{\tilde{K_1}}^{1} z P_n(z)\,\mathrm{d}z \\
&+ \frac{\Delta}{2}(m-K_1)^2\int_{\tilde{K_1}}^{1} P_n(z)\,\mathrm{d}z \\
&= \frac{\Delta^3}{8}P^*_{n-3,n+3}(\tilde{K_1}) + \frac{\Delta^2}{2}(m-K_1)P^*_{n-2,n,n+2}(\tilde{K_1}) \\
&+ \frac{\Delta}{2}(m-K_1)^2 P^*_{n-1,n+1}(\tilde{K_1}) \qquad (4.83)
\end{aligned}$$

where we use Equations (4.43), (4.44) and (4.45). Now we come back to solving Equation (4.82)

$$\begin{aligned}
\ldots &= \frac{1}{\Delta}\Big(\frac{\Delta^3}{8}\frac{1-\tilde{K_1}^3}{3} + \frac{\Delta^2}{2}(m-K_1)\frac{1-\tilde{K_1}^2}{2} + \frac{\Delta}{2}(m-K_1)^2(1-\tilde{K_1})\Big) \\
&+ \frac{3}{\Delta}P_1\left(\frac{2}{\Delta}(x-l)-1\right)e^{-2(w-t)}\Big(\frac{\Delta^3}{8}\frac{1-\tilde{K_1}^4}{4} + \frac{\Delta^2}{2}(m-K_1)\frac{1-\tilde{K_1}^3}{3} + \frac{\Delta}{2}(m-K_1)^2\frac{1-\tilde{K_1}^2}{2}\Big) \\
&+ \frac{5}{\Delta}P_2\left(\frac{2}{\Delta}(x-l)-1\right)e^{-6(w-t)}\Big(\frac{\Delta^3}{8}\frac{-9\tilde{K_1}^5+5\tilde{K_1}^3+4}{30} + \frac{\Delta^2}{2}(m-K_1)\frac{-3\tilde{K_1}^4+2\tilde{K_1}^2+1}{8} \\
&+ \frac{\Delta}{2}(m-K_1)^2\frac{\tilde{K_1}-\tilde{K_1}^3}{2}\Big) \\
&+ \sum_{n=3}^{\infty}\frac{2n+1}{\Delta}P_n\left(\frac{2}{\Delta}(x-l)-1\right)e^{-n(n+1)(w-t)}\Big(\frac{\Delta^3}{8}P^*_{n-3,n+3}(\tilde{K_1}) \\
&+ \frac{\Delta^2}{2}(m-K_1)P^*_{n-2,n,n+2}(\tilde{K_1}) + \frac{\Delta}{2}(m-K_1)^2 P^*_{n-1,n+1}(\tilde{K_1})\Big) - E_8^2(t,w).
\end{aligned}$$

(4.84)

The last element is the value of V_{10}

$$\begin{aligned} V_{10}(t,w) &:= \mathrm{Var}\big(\max\{K_1 - S^1_w, 0\}|\mathcal{F}_t\big) \\ &= \int_l^{K_1} (K_1 - y)^2 p_{t-w}(x,y)\ \mathrm{d}y - E^2_{10}(t,w) \\ &= \int_l^{u} (K_1 - y)^2 p_{t-w}(x,y)\ \mathrm{d}y - \int_{K_1}^{u} (K_1 - y)^2 p_{t-w}(x,y)\ \mathrm{d}y - E^2_{10}(t,w) \\ &= V_0(k,t) + \big(K_1 + m - E_0(k,t)\big)^2 - (V_8(t,w) + E^2_8(t,w)) - E^2_{10}(t,w). \end{aligned} \quad (4.85)$$

4.5.5 Gas-driven storage

In this section we consider one gas-storage-driven payoff which is similar to the hydro-driven payoffs studied above. Consider a storage owner who regularly sells or buys gas and respectively fills or empties the storage facility. Assume, that this is her stochastic optimal control policy and the decision to inject or withdraw is a result of the optimisation problem under some constraints. The power price and costs on the managing storage facility are the key drivers to find the optimal policy. If taking the costs as a deterministic function of time, one can think of this policy as solely dependent of the stochastic gas price. So we can further assume that the resulting storage level S_t is indirectly a function of the optimal stochastic control. In this sense we can regard a payoff that would be hedging the position of this storage owner in case of a low reservoir level and low prices.

Payoff 4

We consider a gas storage facility and an storage owner who injects and withdraws the necessary amount of gas into the reservoir. Then with a cost function C_t the payoff is defined as

$$H_t(P_t, S_t, C_t) = \max\{P_t - M, 0\} \times \max\{S_t - m, 0\} - \max\{M - P_t, 0\} \times \max\{m - S_t, 0\} - C_t, \quad (4.86)$$

where M is the average gas price level and all the rest notations are as in the previous section. This payoff has the following interpretation:

- **Case 1:** $P_t > M$ **(gas prices are relatively high) and** $S_t > m$ **(reservoir is relatively full).** This is the most favourable situation which results in a positive value that an owner can have by withdrawing and selling the storable asset at the market.
- **Case 2:** $P_t > M$ **(gas prices are relatively high) and** $S_t < m$ **(reservoir is relatively empty).** This situation corresponds to the "doing nothing" regime, since the prices are

quite high to buy. So, the value a producer receives is negative due to the costs she has to pay to maintain the storage facility.

- **Case 3: $P_t < M$ (gas prices are relatively low) and $S_t > m$ (reservoir is relatively full).** This situation also corresponds to "doing nothing" regime, since the prices are quite low to withdraw and sell despite the fact that the reservoir is full.

- **Case 4: $P_t < M$ (gas prices are relatively low) and $S_t < m$ (reservoir is relatively empty).** This is an auspicious situation for a producer to buy and inject the storable asset, although the value is negative.

At time t we find that

$$
\begin{aligned}
V_t(P_t,S_t,C_t) &= \mathbb{E}\Big[\int_t^T e^{-r(w,t)}\, H_w(P_w,S_w,C_w)\,\mathrm{d}w|\mathcal{F}_t\Big] \\
&= \int_t^T e^{-r(w,t)}\Big(\mathbb{E}\big[H_w(P_w,S_w,C_w)|\mathcal{F}_t\big]\Big)\,\mathrm{d}w \\
&= \int_t^T e^{-r(w,t)}\Big(\mathbb{E}\big[\max\{P_w-M,0\}\times\max\{S_w-m,0\} \\
&- \max\{M-P_w,0\}\times\max\{m-S_w,0\}-C_w|\mathcal{F}_t\big]\Big)\,\mathrm{d}w \\
&= \int_t^T e^{-r(w,t)}\Big(\underbrace{\mathbb{E}\big[\max\{P_w-M,0\}|\mathcal{F}_t\big]}_{=:E_3(t,w)}\,\underbrace{\mathbb{E}\big[\max\{S_w-m,0\}|\mathcal{F}_t\big]}_{=:E_2(t,w)} \\
&- \underbrace{\mathbb{E}\big[\max\{M-P_w,0\}|\mathcal{F}_t\big]}_{=:E_5(t,w)}\,\underbrace{\mathbb{E}\big[\max\{m-S_w,0\}|\mathcal{F}_t\big]}_{=:E_6(t,w)} \\
&+ \rho_4\Big(\sqrt{\underbrace{\mathsf{Var}(\max\{P_w-M,0\}|\mathcal{F}_t)}_{=:V_3(t,w)}\,\underbrace{\mathsf{Var}(\max\{S_w-m,0\}|\mathcal{F}_t)}_{=:V_2(t,w)}} \\
&- \sqrt{\underbrace{\mathsf{Var}(\max\{M-P_w,0\}|\mathcal{F}_t)}_{=:V_5(t,w)}\,\underbrace{\mathsf{Var}(\max\{m-S_w,0\}|\mathcal{F}_t)\Big)}_{=:V_6(t,w)}} - C_w\Big)\,\mathrm{d}w \\
&= \int_t^T e^{-r(w,t)}\Big(E_3(t,w)\cdot E_2(t,w) - E_5(t,w)\cdot E_6(t,w) \\
&+ \rho_4\big(\sqrt{V_3(t,w)\cdot V_2(t,w)} - \sqrt{V_5(t,w)\cdot V_6(t,w)}\big) - C_w\Big)\,\mathrm{d}w, \qquad (4.87)
\end{aligned}
$$

where $\rho_4 = \mathsf{corr}(\max\{P_w-M,0\},\max\{S_w-m,0\}) = \mathsf{corr}(\max\{M-P_w,0\},\max\{S_w-m,0\})$.

We need to compute the values of E_5, E_6, V_5 and V_6. But this can easily be done since we know the values of E_2, E_3, V_2, and V_3 from the previous sections. So we start with the value of E_5

$$\begin{aligned} E_5(t,w) &:= \mathbb{E}\big[\max\{M - P_w, 0\}|\mathcal{F}_t\big] \\ &= \mathbb{E}\big[\max\{M - e^{f(w)+X_w}, 0\}|P_t\big] \\ &= \int_{-\infty}^{\ln M} (M - e^y)\phi_t(x,y)\ \mathrm{d}y \\ &= \int_{-\infty}^{\infty} (M - e^y)\phi_t(x,y)\ \mathrm{d}y + E_3(t,w) \\ &= M - e^{m_1+\frac{m_2}{2}} + e^{m_1+\frac{m_2}{2}}\Phi(d_2) - M\Phi(d_1) \\ &= M\Phi(-d_1) - e^{m_1+\frac{m_2}{2}}\Phi(-d_2), \end{aligned} \tag{4.88}$$

where d_1 and d_2 are given above and continue with the value of V_5

$$\begin{aligned} V_5(t,w) &:= \mathsf{Var}(\max\{M - P_w, 0\}|\mathcal{F}_t) \\ &= \mathsf{Var}(\max\{M - e^{f(w)+X_w}, 0\}|P_t) \\ &= \int_{-\infty}^{\ln M} (M - e^y)^2\phi_t(x,y)\ \mathrm{d}y - \Big(\int_{-\infty}^{\ln M} (M - e^y)\phi_t(x,y)\ \mathrm{d}y\Big)^2 \\ &= \int_{-\infty}^{\infty} (M - e^y)^2\phi_t(x,y)\ \mathrm{d}y - \int_{\ln M}^{\infty} (M - e^y)^2\phi_t(x,y)\ \mathrm{d}y - E_5^2(t,w) \\ &= \mathsf{Var}(M - e^y) + \mathbb{E}^2[M - e^y] - \int_{\ln M}^{\infty} (e^y - M)^2\phi_t(x,y)\ \mathrm{d}y - E_5^2(t,w) \\ &= \mathsf{Var}(e^y) + (M - \mathbb{E}[e^y])^2 - \int_{\ln M}^{\infty} (e^y - M)^2\phi_t(x,y)\ \mathrm{d}y - E_5^2(t,w) \\ &= (e^{m_2} - 1)e^{2m_1+m_2} + \big(M - e^{m_1+\frac{m_2}{2}}\big)^2 - (V_3(t,w) + E_3^2(t,w)) - E_5^2(t,w). \end{aligned} \tag{4.89}$$

Then we compute the value of E_6

$$\begin{aligned} E_6(t,w) &:= \mathbb{E}\big[\max\{m - S_w, 0\}|\mathcal{F}_t\big] \\ &= \int_{l}^{m} (m - y)p_{w-t}(x,y)\ \mathrm{d}y \\ &= E_2(t,w) - E_0(t,w). \end{aligned} \tag{4.90}$$

and finally proceed with the value of V_6

$$
\begin{aligned}
V_6(t,w) &:= \mathrm{Var}(\max\{m - S_w, 0\}|\mathcal{F}_t) \\
&= \int_l^m (m - S_w)^2 p(S_w, k, S_t, t)\,\mathrm{d}S_w - E_6^2(t,w) \\
&= \int_l^m (m - y)^2 p_{w-t}(x,y)\,\mathrm{d}y - E_6^2(t,w) \\
&= \int_l^u (m - y)^2 p_{w-t}(x,y)\,\mathrm{d}y - \int_m^u (m - y)^2 p_{w-t}(x,y)\,\mathrm{d}y - E_6^2(t,w) \\
&= V_0(t,w) - V_2(t,w) + E_0^2(t,w) - E_2^2(t,w) - E_6^2(t,w). \qquad (4.91)
\end{aligned}
$$

4.6 Numerical examples

4.6.1 Hydro storage

Simple products

In section 4.5.3 we discussed fundamental financial products that can be used for hedging purposes in the storage industry. Those formulas for the futures and options demand negligible computational efforts.

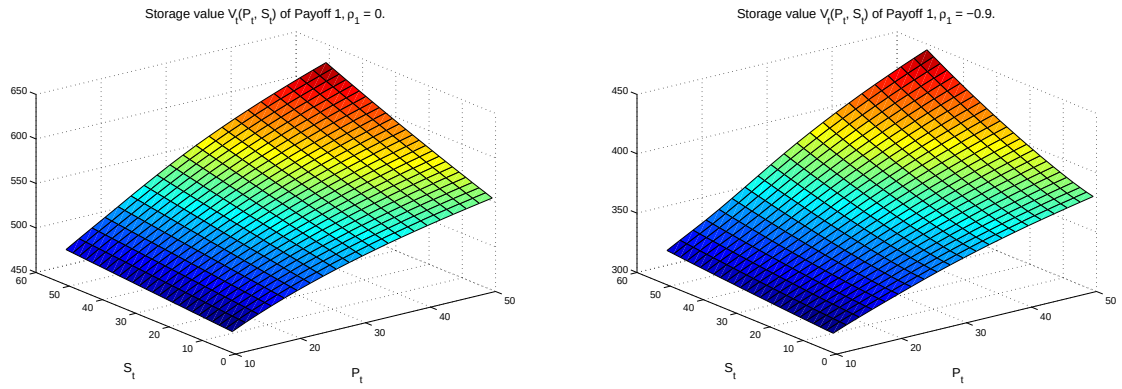
Hydro-Driven Storage: Payoffs 1, 2, 3

Here we illustrate the hydro storage value problem described above by the payoffs 1, 2, and 3. For some fixed parameters values we plot the payoff for a range of S_t and P_t. For all the examples we consider $t = 0.5$ and $T = 5$ in years. For the sake of simplicity we also fix the discount factor $r(t,T)$ and the costs of storage maintenance C_t. We take the following parameters values: $l = 1$, $u = 51$, $m = 26$, $\alpha = 1.5$, $\sigma = 0.2$, $M = 30$, $C = 0$ and $r = 0.03$.

To investigate the role of correlation parameters ρ_1, ρ_2 and ρ_3, we study two cases: zero and negative correlation. Zero correlation implies that there is no relationship between the current power price and the storage level. In other words, when the inflow increases, the power price stays ineffectual. In the markets where various fossil fuels dominate over a hydro-driven electricity production, zero or negligible correlation can exactly be the case since there are many other power price drivers apart from the current reservoir level. However, in the markets with a significant or even dominating share of hydro facilities we can fairly expect an effect of negative correlation. When the inflow increases and the cumulative reservoir is getting full of water, the supply uncertainty decreases and all the market participants are aware of this. So since there is no lack of water in the reservoir, the power price decreases. We investigate how large is the

effect of the correlation on the price for our financial instruments. This can be considered as a correlation sensitivity analysis.

Figure 4.7 depicts the value driven by the payoff 1 given in Equation (4.56) for two values of the correlation parameter ρ_1. We notice that a relatively high power price together with a full water reservoir yield increase possible profit. We also observe that a relatively low power price together with an empty water reservoir decrease the profit. We further note that with negative correlation $\rho_1 = -0.9$ producer's profit lessens in comparison with $\rho_1 = 0$. We interpret this gap as a premium that a producer has to pay for the market information about current reservoir level.

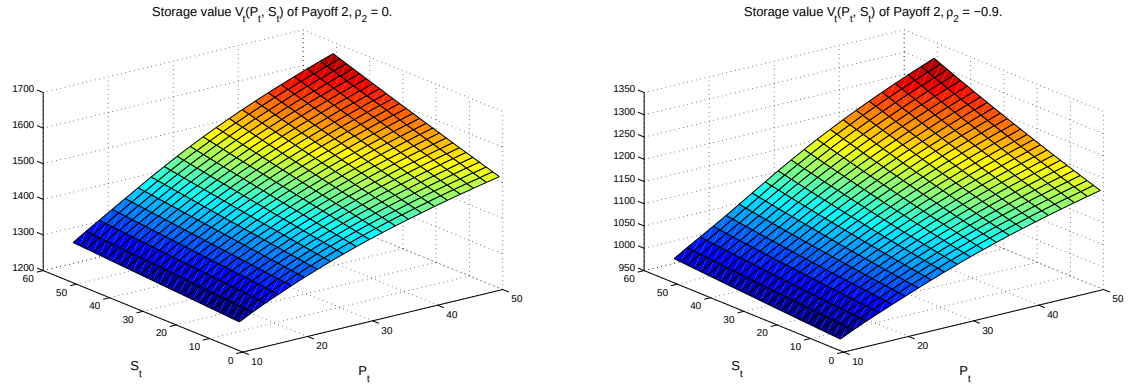


(a) (b)

Figure 4.7: Storage value with payoff 1 with power price P_t and storage level S_t. Parameters: l = 1, u = 51, m = 26, $\alpha = 1.5$, σ = 0.2, M = 30, C = 0, r = 0.03.

Figure 4.8 depicts the value driven by the payoff 2 given in Equation (4.61) for two values of the correlation parameter ρ_2. We observe that a relatively high power price together with a full water reservoir increase possible profit. And a relatively low power price together with an empty water reservoir decrease the profit. We also see the same negative correlation effect. The main difference here is that the profit of the payoff 2 is almost three times higher compared to the profit of the payoff 1. We reason this with the flexibility of the payoff 2 to produce power even even if the current reservoir level is less than m at some rate $\kappa(S_t)$.

Figure 4.9 depicts the value driven by the payoff 3 given in Equation (4.79) for two values of the correlation parameter ρ_3 and the parameter K_1. We mainly state two dependencies: the profit decrease when the correlation coefficient ρ_3 together with the coefficient K_1 increase. The reasoning for the first case can be regarded as an information premium for a producer. The explanation for the second case is intuitively clear: a lower critical production level K_1 at which we are allowed to produce leads to a larger capacity to produce and benefit. A higher critical production level K_1 results in much smaller capacity for electricity production. The value can even be negative and our producer has losses.



(a) (b)

Figure 4.8: Storage value with payoff 2 with power price P_t and storage level S_t. Parameters: l = 1, u = 51, m = 26, $\alpha = 1.5$, σ = 0.2, M = 30, C = 0, r = 0.03.

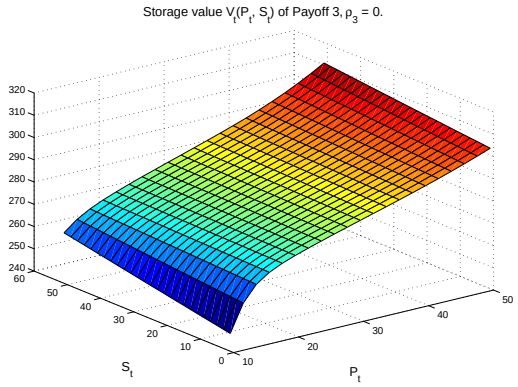
4.6.2 Gas storage

Gas-Driven Storage: Payoff 4

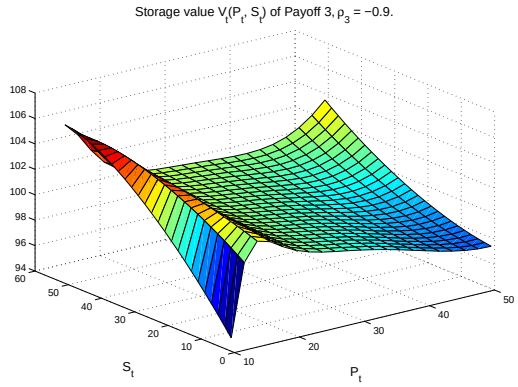
In this section we illustrate the gas storage value driven by the payoff 4 given in Equation (4.87). For some fixed parameters values we plot the payoff for a range of S_t and P_t. For all the examples we consider $t = 0.5$ and $T = 5$ in years. For the sake of simplicity we also fix the discount factor $r(t, T)$ and the costs of storage maintenance C_t. We assume the following parameters values: $l = 1$, $u = 51$, m = 26, $\alpha = 1.5$, σ = 0.2, M = 30, C = 0 and r = 0.03.

Here we also study two cases: zero and negative correlation coefficient ρ_4. However, here we assume that in the markets where different fossil fuels (coal, gas) dominate over a hydro-driven power production, negative correlation can exactly be the case, since the market is aware of the current gas supply level. In the markets with dominating share of hydro facilities we can fairly expect zero correlation, since the current gas storage level will not be critical for power production.

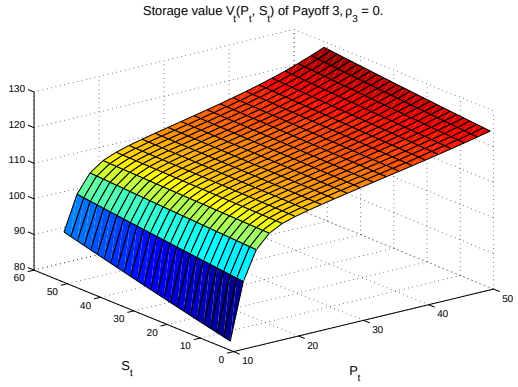
Figure 4.10 depicts the value obtained with the payoff 4 for two values of the correlation parameter ρ_4. This result is consistent with the hydro-storage case considered in the previous section. Particularly, when a higher gas price together with a relatively full storage facility gives a positive value to a storage owner, as she withdraws, sells the gas in the market and obtains profit. A relatively low gas price together with a relatively empty storage facility gives small profit. We also remark the same effect with the correlation coefficient ρ_4: a negative correlation leads to the lower profit value.



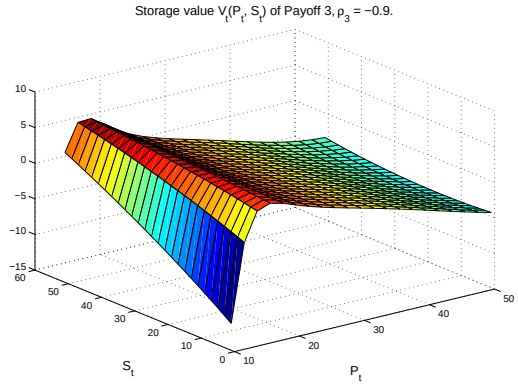
(a) $K_1 = 12$



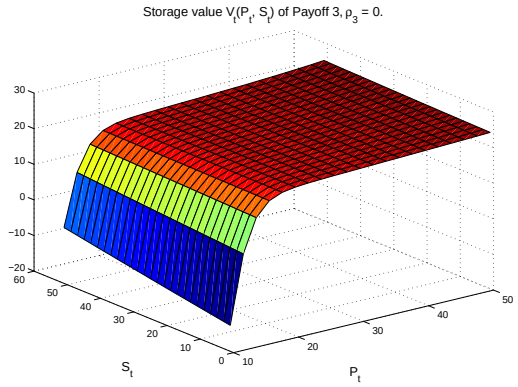
(b) $K_1 = 12$



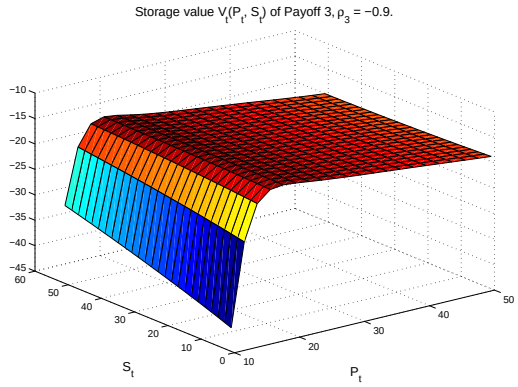
(c) $K_1 = 26$



(d) $K_1 = 26$

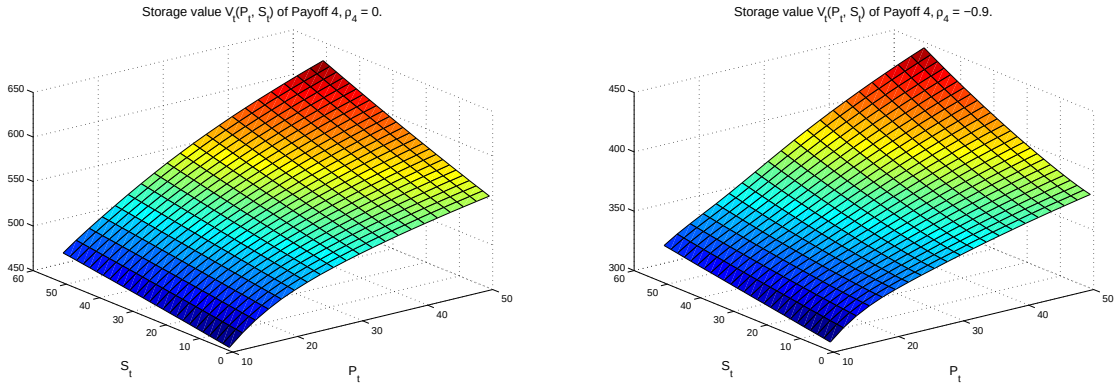


(e) $K_1 = 40$



(f) $K_1 = 40$

Figure 4.9: Storage value with payoff 3 with power price P_t and storage level S_t. Parameters: $l = 1$, $u = 51$, $m = 26$, $\alpha = 1.5$, $\sigma = 0.2$, $M^- = 25$, $M^+ = 50$, $C = 0$, $r = 0.03$.



(a) (b)

Figure 4.10: Storage value with payoff 4 with gas price P_t and storage level S_t. Parameters: $l = 1$, $u = 51$, $m = 26$, $\alpha = 1.5$, $\sigma = 0.2$, $M = 30$, $C = 0$, $r = 0.03$.

4.7 Discussion and conclusion

In this chapter we have studied storage value problem. We proposed a new stochastic approach to storage value modelling to complement widely used optimal stochastic control methods. The main advantage of our approach is its straightforwardness and easy implementation, which is a result of a closed-form transition probability density formula for a bounded diffusion to model storage level process.

We scrutinised a number of financial instruments useful for hedging market position of a storage owner and/or power producer. We reduced our attention mainly to hydro and gas storage facilities and found out that despite they have some similarities, the associated payoffs should be treated differently due to physical storage differences.

We may suggest several extensions to continue with this topic. One of the assumptions made above on the constant mean-reverting level m is not quite realistic due to seasonal inflow. Also it would be more relevant to include a jump component for a better fit to power price modelling. Namely,

$$\mathrm{d}S_t = -2(S_t - m(t))\ \mathrm{d}t + \sqrt{2(S_t - l)(u - S_t)}\ \mathrm{d}W_t^S, \tag{4.92}$$

where $m(t)$ could be some trigonometric function capturing seasonal behaviour of the storage level. And

$$\begin{aligned} P_t &= e^{f(t)+X_t}, \\ \mathrm{d}X_t &= -\alpha X_t \,\mathrm{d}t + \sigma \,\mathrm{d}W_t^X, \\ \mathrm{d}P_t &= \alpha\Big(\mu(t) - \ln P_t\Big) P_t \,\mathrm{d}t + \sigma P_t \,\mathrm{d}W_t^X + \,\mathrm{d}Q_t, \end{aligned} \tag{4.93}$$

where Q_t is a compound Poisson process with some finite intensity λ. We can also consider various jump size distributions: Gaussian, exponential, Pareto, Laplace studied in Chapters 2 and 5. Depending on this choice we either can directly compute the payoff value or we need to simulate the processes P_t and S_t to get the value of the virtual hydro-driven power plant in the same manner as we did in Chapter 5.

It would also be interesting to investigate the role of rate κ dependent of the current reservoir level and responsible for the production rate. In this study we assumed a fixed value for it, which is a quite simplification, though we still managed to detect its great impact on the storage value.

One more beneficial thing would be to benchmark our model against a classical stochastic optimal control approach and to investigate the comparative results. This will be the scope of our future research.

CHAPTER 5

MODEL RISK

5.1 Introduction

Ever since the financial crisis struck the importance of models has been in the centre of attention. In particular, it has been realised that risk management is subject to model risk and that model risk has to be adequately measured. In some cases, one might be able to assign probabilities to the different models (resp. parameters within a specific model), where one ends up according to the terminology of Knight [1921] with model or parameter risk. For standard financial markets the issue has been addressed extensively in recent years. For instance, Avellaneda *et al.* [1995] and Cont [2006] consider worst-case scenarios and obtain a range of possible prices for derivatives. Rebonato [2010] addresses model risk issues concerning stress testing, while Glasserman & Xu [2012] and Ruehlicke [2013] discuss robust approaches to risk management including model risk.

In contrast, model risk has not been discussed in the context of energy markets. In view of the recent changes in European energy market, especially the German "Energiewende", with the increasing impact of volatile renewable energies, it is clear that model risk is of particular interest. One important aspect is the need for reinvestment (replacement investments and building more capacity) in the power plant park on Company and European level. The financial streams of such an investment can be generated on the market for energy derivatives in terms of spread options. For instance, a gas-fired power plants can be represented as a clean crack spread option, where the owner of such an option is long electricity and short gas and emission certificates. A positive investment decision is made in case such a contract is in the money, meaning that we observe a positive spread on the time interval under consideration.

We will consider the model risk inherent in conventional gas-fired power plants since these are in particularly affected by the increasing share of renewable. Flexible gas-fired power plants have been build to address the need in peak hours during the day. So their use is based on short term demand and this peak demand is highly affected by the uncertain in-feed of renewable generated by solar or wind power plants.

This chapter is organised as follows. In Section 5.2 we review the methodology introduced by Bannör & Scherer [2013] which we use to access model risk. In Section 5.3 we explain power plant valuation in terms of spread options and introduce the stochastic models used to fit the price processes. In Section 5.4 we undertake our empirical investigation. In Section 5.5 we calculate the relevant risk measures according to techniques introduced in Section 5.2, discuss our results and put them into context. Section 5.6 finally concludes.

5.2 Incorporating parameter risk

Modelling electricity prices is a considerable task since the electricity market is still developing and subject to changes in regulation and market design. Nevertheless, there are numerous attempts trying to model the dynamics of electricity prices, for recent discussions see Culot [2013], De Jong [2006], Meyer-Brandis & Tankov [2008] or Benth *et al.* [2012] and for textbook accounts Eydeland & Wolyniec [2003] or Burger *et al.* [2008]. Having chosen a specific model, one still has to determine the model's correct parameters. In electricity markets, one typically relies on time series analysis to obtain a model's parameters due to the lack of liquid derivatives prices to calibrate to. Thus, the standard procedure is to estimate the parameters from time series of electricity prices and to plug the point estimate into the desired calculations afterwards, e.g. the calculation of electricity derivatives prices. But, when simply plugging in the obtained parameter for price determination, one disregards the whole information which is contained in the distribution of the estimator. If a parameter may be difficult to estimate (like, e.g., in presence of a small sample size), one faces tremendous risk that one does not obtain the right parameter due to the estimator's bias and/or variance. This risk is not neglectable: when calculating derivatives prices, taking a slightly different parameter than the right one may result in considerable different prices (as demonstrated in Schoutens *et al.* [2004]).

Following the terminology of Knight [1921], the above problem is described as *parameter risk*: via the estimator's distribution, one has an idea about the likelihood of the different parameters, but one does not know for sure whether the point estimate parameter is the right one. To account for this, Bannör & Scherer [2013] introduce the framework of parameter risk-captured pricing. In this chapter several ideas on treating parameter risk or uncertainty suggested in Cont [2006]; Gupta *et al.* [2010]; Lindström [2010] are generalised and a concise framework to incorporate parameter and estimation risks into financial prices is provided. We also briefly sum up the guidelines to parameter risk-captured pricing as described in Bannör & Scherer [2013].

5.2.1 Measuring parameter risk and risk-captured prices

Our methodology to measure parameter risk is based on *convex risk measures*. The notion of convex and coherent risk measures (see the seminal paper of Artzner *et al.* [1999]) have emerged

from the shortcomings of the Value-at-Risk. The Value-at-Risk, simply being some upper quantile, is popular among practitioners and easy to interpret, but there are settings where the diversification of financial instruments is penalized, i.e. a diversified portfolio of financial positions is regarded more risky than the single positions. Since this property is not desirable, alternative measures, most notably convex risk measures, have been developed. Convex risk measures have been treated and extended in many papers like Kusuoka [2001]; Föllmer & Schied [2002]; Frittelli & Scandolo [2006] and there are numerous tractable examples for convex risk measures available like, e.g., the Average-Value-at-Risk (cf. Acerbi & Tasche [2002]).

When considering parameter risk, i.e. there is a distribution R on the parameter space Θ available (which may be induced, e.g., from an estimator statistics $\hat{\theta} = \hat{\theta}(X_1, \dots, X_N)$ via using the pushforward measure), we can define the risk-captured price as a convex risk measure, evaluated on the price regarded as a function of the parameter θ. A formal definition is as follows:

Definition 3 (Risk-captured price)**.** Given a parameter space Θ with a distribution R on the parameters, a parameterised family of valuation measures $(Q_\theta)_{\theta\in\Theta}$ and a law-invariant, normalised convex risk measure ρ as a generator (defined on a proper domain of functions on Θ) we calculate the *risk-captured price* of a contingent claim X by

$$\Gamma(X) := \rho(\theta \mapsto \mathbb{E}_\theta[X]).$$

The risk-captured price of X may be interpreted as an ask price of the contingent claim X. On the other hand, one can introduce the dual analogue by $\bar{\Gamma}(X) := -\Gamma(-X)$ which can be interpreted as a bid price. A detailed discussion as well as technical details can be found in Bannör & Scherer [2013]. The idea behind this definition is quite intuitive. If we have different (potentially correct) parameters available and we know the probability that some parameter is the correct one (given by the measure R on Θ), then we immediately get a distribution of prices for X, since each parameter θ can be plugged into the pricing formula $\mathbb{E}_\theta[X]$. Then, we apply the risk measure ρ to weight the different prices according to the probability measure R and incorporate the parameter risk. This is illustrated in Figure 5.1.

To shed some light on the concept, we consider the parameter risk-captured price generated by the Average-Value-at-Risk (AVaR)[1]. We follow Bannör & Scherer [2013] and define the AVaR w.r.t. the significance level $\alpha \in (0, 1)$ of some random variable X as the integrated upper tail of X, i.e.

$$\mathrm{AVaR}_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha q_X(1-\beta)\,\mathrm{d}\beta,$$

denoting by $q_X(\gamma)$ the (lower) γ-quantile of the random variable X.

When applying the regular AVaR to a contingent claim X (instead of taking the expectation

[1] A detailed discussion on the properties of the Average-Value-at-Risk is provided in Acerbi & Tasche [2002].

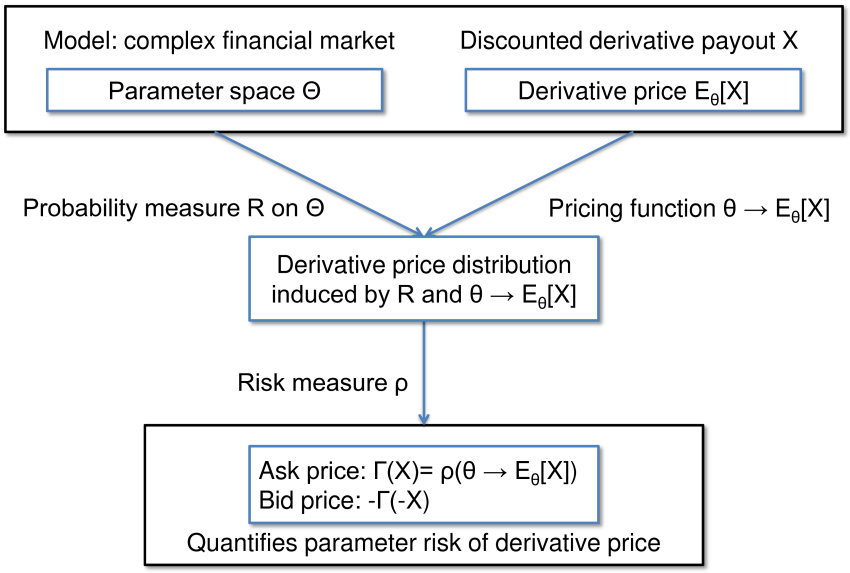


Figure 5.1: Visualisation of the steps of parameter risk-capturing valuation.

w.r.t. an obtained model Q_θ), the AVaR measures the risk which may occur according to the previously specified model Q_θ. In contrast, when calculating the parameter risk-captured price of X being induced by the AVaR, risk-neutral prices $(\mathbb{E}_\theta[X])_{\theta\in\Theta}$ w.r.t. different models $(Q_\theta)_{\theta\in\Theta}$ are compared and subsumed by the AVaR risk measure. Hence, the AVaR is used to quantify the parameter risk we are exposed to when pricing X.

5.2.2 Using asymptotic distributions for determining parameter risk-captured prices

In [Bannör & Scherer, 2013, Proposition 4.2], it has been shown that the AVaR-induced risk-captured price is continuous w.r.t. the topology of weak convergence on the parameter distributions when the price evaluation function for the contingent claim $\theta \mapsto \mathbb{E}_\theta[X]$ is continuous and bounded. In particular, if the parameter distribution R is complicated to calculate or even unknown, one might conveniently replace R by the asymptotic distribution $\tilde{R}$ in case of large sample size. Due to the continuity, the approximation of risk-captured prices by substituting the original parameter distribution R with the (more tractable) asymptotic distribution $\tilde{R}$ is feasible. The advantage of this procedure is that for very wide classes of estimators (e.g. Maximum Likelihood estimators under mild conditions), the asymptotic distribution is known and follows a Normal distribution. In these cases, the calculation of the AVaR-induced risk-captured prices is particularly comfortable, since it can be done in a closed-form. Therefore, the application of the theory of risk-captured prices for asymptotically normal estimators can be done in a straightforward manner.

5.3 Spread options

5.3.1 Spread options and power plant valuation

From a financial point of view we can consider the owner of a (electricity generating) power plant as long electricity and short the fuels needed for production. With the introduction of carbon emission certificates traded at the European Emission Trading Scheme (EU ETS) the price of these certificates, of which the owner is short, has to be considered as well. This leads naturally to viewing a power plant as a spread option consisting of the difference of these prices. Typical fuel spread options are the dark spread, i.e. the difference between power and coal and the spark spread, the difference between power and gas (see Burger *et al.* [2008] for further discussion). The variants of these options taking the price of carbon into account are labelled clean dark resp. spark spread options.

In our investigation we will focus on the clean spark spread to model the value of a gas power plant. Gas-fired power plants are particular important as they were supposed to replace coal-fired power plants in the short- to medium term triggered by the introduction of emission certificates as they are cleaner and more efficient. However, the low carbon price and the huge inflow of renewables during peak hours made most gas-fired power plants a very costly investment for utilities. For our analysis we consider the evaluation of power plant dispatch (i.e. the dispatch of power plant stages) in terms of its financial position.

We will use spot price processes in order to assess the day-by-day risk position of such a position. Thus, we will model the daily profit (or loss) of the virtual power plant position as

$$V_t = \max\{P_t - h\, G_t - \eta E_t, 0\}, \tag{5.1}$$

where P_t is the power price, G_t is the gas price, E_t is the carbon certificate price, h is the heat rate of the power plant, and η is the CO_2 emission rate of the power plant.

5.3.2 Energy price models

In this section we introduce the energy price models. We use models widely used in the literature, see Benth *et al.* [2012], where different models are compared and Benth & Koekebakker [2008], Benth & Kufakunesu [2009], where the models used here are put to work. Let us point out that the methodology to capture parameter risk also applies to alternative models as, e.g., recently proposed in Culot [2013] or summarised in textbooks such as Burger *et al.* [2008].

We model the emission price as a geometric Brownian motion

$$\mathrm{d}E_t = \alpha^E\, E_t\, \mathrm{d}t + \sigma^E\, E_t\, \mathrm{d}W_t^E, \tag{5.2}$$

the gas price as a mean-reverting process[1]

$$\begin{aligned} G_t &= e^{g(t)+Z_t}, \\ \mathrm{d}Z_t &= -\alpha^G\, Z_t\, \mathrm{d}t + \sigma^G\, \mathrm{d}W_t^G, \end{aligned} \tag{5.3}$$

and the power price as a sum of two mean-reverting processes[2]

$$\begin{aligned} P_t &= e^{f(t)+X_t+Y_t}, \\ \mathrm{d}X_t &= -\alpha^P\, X_t\, \mathrm{d}t + \sigma^P\, \mathrm{d}W_t^P, \\ \mathrm{d}Y_t &= -\beta\, Y_t\, \mathrm{d}t + J_t\, \mathrm{d}N_t, \end{aligned} \tag{5.4}$$

where α^G, α^P, and β are mean-reversion forces for gas and power prices respectively; N is a Poisson process with intensity λ and J_t are independent identically distributed (i.i.d.) random variables representing the jump size. For the jump size, we consider two different scenarios: first, we suggest to use a non-central Laplace distribution to capture the heavy-tail nature of spikes. Second, for comparison, we employ the Gaussian distribution as has already been done in Cartea & Figueroa [2005]. Functions $g(t)$ and $f(t)$ are seasonal trend components for gas and power respectively defined as

$$\begin{aligned} f(t) &= a_1 + a_2\, t + a_3 \cos(a_5 + 2\pi t) + a_4 \cos(a_6 + 4\pi t), \\ g(t) &= b_1 + b_2\, t + b_3 \cos(b_5 + 2\pi t) + b_4 \cos(b_6 + 4\pi t), \end{aligned} \tag{5.5}$$

where a_1 and b_1 are the production expenses, a_2 and b_2 are the slopes of increase in these costs. The rest parameters are responsible for two seasonal changes in summer and winter respectively. In the current setting we also assume that W^E, W^G, and N are mutually independent processes, but W^P and W^G are allowed to be correlated, i.e.,

$$\mathrm{d}W_t^P\, \mathrm{d}W_t^G = \rho\, \mathrm{d}t. \tag{5.6}$$

[1] See Lucia & Schwartz [2002].
[2] See Hambly *et al.* [2009].

Here we introduce two processes to capture the power price movements. The first one X_t is a zero mean-reverting process, responsible for the so-called base signal, i.e. daily price fluctuations. The second process Y_t is a mean-reverting jump process, responsible for price shocks (which may ocur due to sudden inflow of power from renewable energy, or an outage of significant capacity). The number of processes used in the modelling approach is one of the critical choices one has to take when analysing the fine power price structure. This has been discussed in Chapter 2.4.

One can obtain the following expressions for conditional mean and variance of the logarithmic prices:

$$\begin{cases} \mathbb{E}[\ln E_t] = \ln E_0 + \left(\alpha^E - \dfrac{(\sigma^E)^2}{2}\right) t, \\ Var(\ln E_t) = (\sigma^E)^2\, t, \\ \mathbb{E}[\ln G_t] = g(t) + \left(\ln G_0 - g(0)\right) e^{-\alpha^G t}, \\ Var(\ln G_t) = \dfrac{(\sigma^G)^2}{2\alpha^G}\left(1 - e^{-2\alpha^G t}\right), \\ \mathbb{E}[\ln P_t] = f(0) + X_0\, e^{\alpha^P t} + Y_0\, e^{\beta t} + \dfrac{\lambda}{\beta}\left(1 - e^{-\beta t}\right)\mathbb{E}[J], \\ Var(\ln P_t) = \dfrac{(\sigma^P)^2}{2\alpha^P}\left(1 - e^{-2\alpha^P t}\right) + \dfrac{\lambda}{\beta}\left(1 - e^{-2\beta t}\right)\mathbb{E}[J^2]. \end{cases} \tag{5.7}$$

The total set of parameters includes $\{\alpha^E, \sigma^E, g(t), \alpha^G, \sigma^G, f(t), \alpha^P, \beta, \sigma^P, \lambda, \mathbb{E}[J], \mathbb{E}[J^2], \rho\}$.[1] Hence, the hybrid model we have chosen for modelling the clean spark spread is not parsimonious and allows for several degrees of freedom. Consequently, the risk of determining parameters in a wrong way is considerable and it will turn out that even the determination of single parameters may lead to tremendous results for prices obtained in the model.

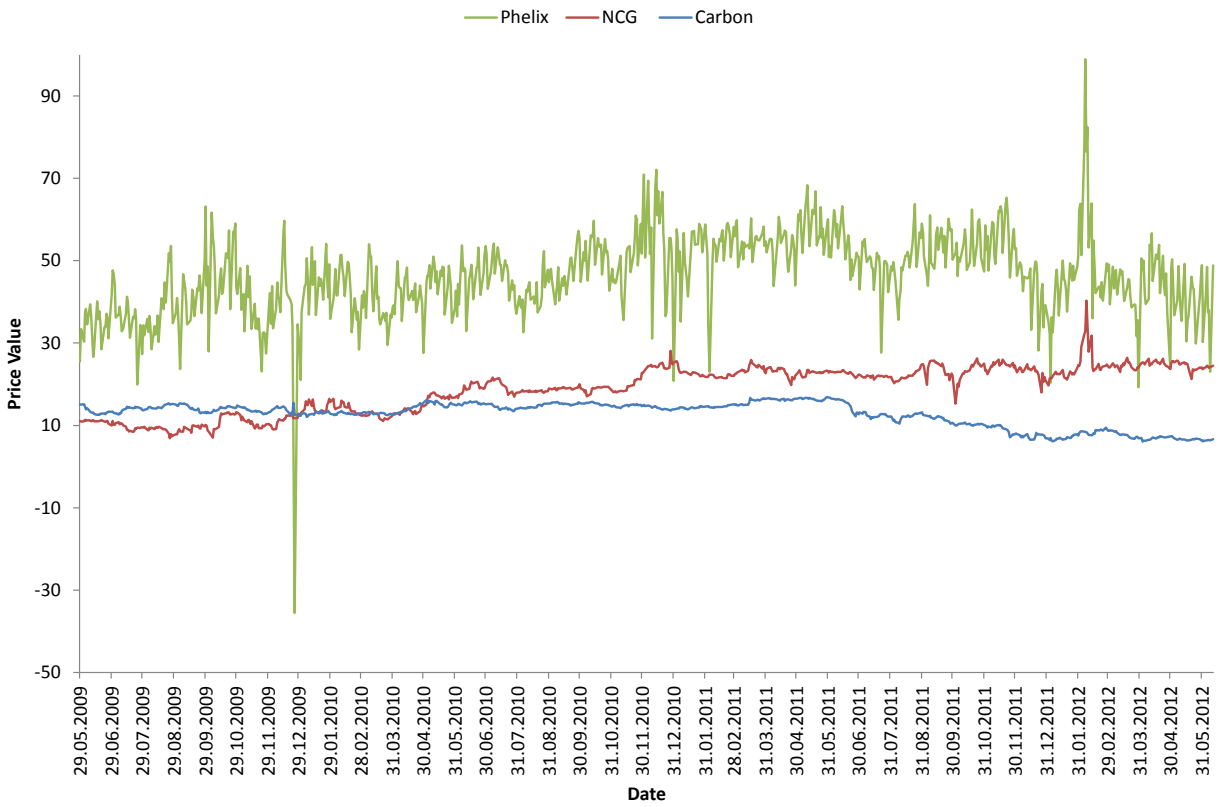


Figure 5.2: Evolution of the power (base load), gas, and carbon prices between 25.09.2009 and 08.06.2012.

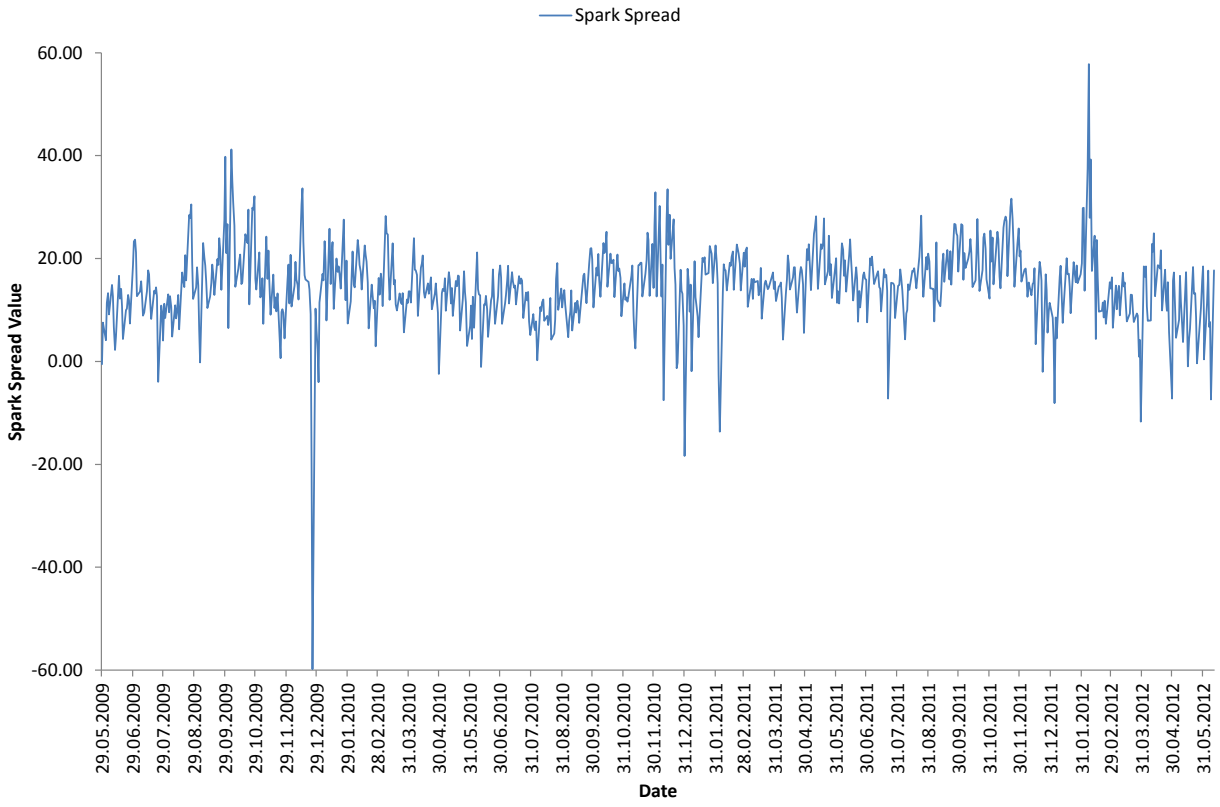


Figure 5.3: Evolution of the clean spark spread between 25.09.2009 and 08.06.2012.

5.4 Empirical investigation

5.4.1 Data and estimating procedure

We use the following data sets[1]: Phelix Day Base[2] (EUR/MWh), NCG[3] daily price (EUR/MWh), and emissions daily price[4] (EUR/EUA). Figures 5.2 and 5.3 depict the paths of the prices together with the spark spread path. The period of observations covers three years: 25.09.2009 - 08.06.2012.

We rely on Maximum Likelihood estimators (ML estimators) as far as possible for our estimation procedure. ML estimators exhibit asymptotic normality and their asymptotic variance is given

[1]In the parametric form of the Laplace and Gaussian distributions, we do not directly use the second moment as a parameter, but the standard scaling parameters for Laplace and Gaussian distributions (variance for Gaussian, mean absolute deviation from median for Laplace). Obviously, using the second moment is only an equivalent re-parametrisation.

[1]All the data sets are taken from the European Energy Exchange, `www.eex.com`.

[2]It is the average price of the hours 1 to 24 for electricity traded on the spot market. It is calculated for all calendar days of the year as the simple average of the auction prices for the hours 1 to 24 in the market area Germany/Austria disregarding power transmission bottlenecks.

[3]Delivery is possible at the virtual trading hub in the market areas of NetConnect Germany GmbH & Co KG.

[4]One EU emission allowance confers the right to emit one tonne of carbon dioxide or one tonne of carbon dioxide equivalent.

by the inverse Fisher information. These properties are particularly suited for our calculation procedure to obtain risk-capturing functionals (see section 2). We estimate the parameters for the emission prices, the joint base signal of power and gas prices, and the jump size distribution by ML estimators by well-established methods (which can be found together with detailed Matlab algorithms in Brigo *et al.* [2007]). The estimation procedure for the parameters of the power price process includes several steps:

1. *Estimation of the seasonal trend and deseasonalisation.* This aims to remove the seasonal pattern from the data. Power prices usually display weekly, monthly, and yearly seasonalities. One of the ways of doing so is to fit the market data to the specified trend function. There are, of course, many candidates for it, one of which is a class of circular functions. For a recent discussion on seasonal functions and their fit see Nowotarski [2013] and Janczura [2013].

2. *Separation of the processes.* To identify which price variations belong to the jump process and which ones are driven by the continuous part of the process is a difficult task. There exists an extensive collection of methods to implement this filtering. One of the simplest ways is to iteratively fix a reasonable threshold and to filter out the jumps as data points above the threshold until no such points occur anymore. Cartea & Figueroa [2005] apply such iterative procedure when filtrating the UK power price. A problem of the method is that it delivers just jump values, but not a complete jump process path together with a mean-reverting component.

3. *Estimation of the mean-reversion rates.* For every component of the power price we have to estimate the forces that push the process back to its mean levels. If there is only one such force, like it is in Geman & Roncoroni [2006], then it will do the job for both base and spike signals and bring some additional noise into the price path as it is justified in Benth *et al.* [2012]. Therefore, it is preferable to individually model the mean-reversion parameters for the base and the spike process. Once we have filtered the X_t process, we can identify it as a first order autoregressive model in continuous time. Discretising the process (to an AR(1) time series) and applying the Maximum Likelihood method yields the estimates. For the details see Knittel & Roberts [2005]. To estimate the mean-reversion rate for the jump process one can take advantage of the approach based on the autocorrelation function (ACF) as surgested by Barndorff-Nielsen & Shephard [2001] and implemented in Meyer-Brandis & Tankov [2008] and Benth *et al.* [2012].

4. *Estimation of the base signal X_t.* As the stationary process in our model is Gaussian, we can estimate the parameters using a ML estimator.[1].

[1]It is more difficult when this process X_t is a Lévy process. Then there are just a few situations in which one can estimate the parameters. Barndorff-Nielsen & Shephard [2001] discuss some distributions like Gamma and refer to possible estimation ways. An application to energy price modelling was suggested in Benth *et al.* [2007] and further implemented in Chapter 2 with an aid of the prediction-based estimating functions method, which is introduced in Sørensen [2000].

5. *Estimation of the spike signal Y_t.* Here we estimate the intensity λ and the spike size parameters. We take the intensity as a constant (a time-dependent intensity is discussed in Meyer-Brandis & Tankov [2008]. Then it can be estimated as spikes frequency, i.e. a number of detected spikes per period. For the spike size distribution we use a parametric assumptions and use ML estimators for the parameter values. One should note that different filtering procedures yield quite different jump process values, mostly large jumps are detected. Therefore, when estimating the jump size distribution from this data via Maximum Likelihood estimation, one obtains considerably different location parameters for the Normal and the Laplace distribution, since the ML estimator for the location parameter of the Laplace distribution is the median, while the ML estimator for the location parameter of the Normal distribution is the mean. This methodology might be criticised, but it is still standard in energy finance, see, e.g., Cartea & Figueroa [2005].

6. *Estimation of the correlation.* We estimate it jointly with the estimation of multivariate normal regression estimation for power and gas prices.

Following the above steps, we estimate the set of parameters

$$\{\alpha^E, \sigma^E, g(t), \alpha^G, \sigma^G, f(t), \alpha^P, \beta, \sigma^P, \lambda, \mu_s, \sigma_s, \rho\}.$$

The result is given in Table 5.1 (estimates for the correlation matrices are in Appendix A) .

5.4.2 Measuring parameter risk

Since the whole parameter distribution is very complex and difficult to obtain (Bunn *et al.* [2013]), we reduce the problem here by considering the distributions of the single parameters separately (e.g. the correlation coefficient, the jump size distribution parameters). Hence, we scrutinise the parameter risk w.r.t. selected parameters separately, disregarding the remaining parameter risk. This procedure is described by the following steps.

Each parameter θ_j is to be estimated by an estimator $\hat{\theta}_j(X_1, \ldots, X_N)$ under the real-world measure and we assume the other parameters $\theta_1, \ldots, \theta_{j-1}, \theta_{j+1}, \theta_N$ to be known. Afterwards, we assume the plug-in estimator as the true value and calculate the asymptotic distribution of the estimator.

We calculate the parameter risk-captured price generated by the Average-Value-at-Risk (AVaR), which is a widely used convex risk measure thoroughly discussed in Acerbi & Tasche [2002]. As mentioned in section 5.2.2, the AVaR-induced risk-captured prices are continuous w.r.t. the weak topology on the parameter distributions, if the price function $\theta \mapsto \mathbb{E}_\theta[X]$ is continuous and bounded in θ.

Estimation Step	Product	Estimates	Method
GBM	Emissions	$\alpha^E = -0.2843$, $\sigma^E = 0.4079$	MLE
Seasonal trend	Power	$a_1 = 3.6716$, $a_2 = 0.0980$, $a_3 = -0.0274$ $a_4 = 0.0368$, $a_5 = 0.6524$, $a_6 = 0.9530$	OLS
Seasonal trend	Gas	$b_1 = 2.3420$, $b_2 = 0.3503$, $b_3 = 0.0218$ $b_4 = -0.0445$, $b_5 = 0.7829$, $b_6 = 1.6126$	OLS
Filtering	Power		3×Std.Dev rule
Base process	Gas	$\alpha^G = 13.5827$, $\sigma^G = 0.7768$	Multivariate
Base process	Power	$\alpha^P = 121.8684$, $\sigma^P = 2.5943$, $\rho = 0.1247$	normal regression
Spike mean-reversion	Power	$\beta = 243.7240$	$2\times\alpha^P$
Spike intensity	Power	$\lambda = 13.4936$	Annual frequency
Spike size (Laplace)	Power	$\mu_s(median) = 0.3975$, $\sigma_s(scale) = 0.6175$	MLE
Spike size (normal)	Power	$\mu_s(mean) = 0.0863$, $\sigma_s^2(variance) = 0.5857$	MLE
Heat rate	Gas	h = 2.5	Technical constant
Emission rate	Gas	η = 0.4	Technical constant
Interest rate		r = 3%	Market-quoted

Table 5.1: An overview over the parameters of the hybrid model for emissions, gas, and power prices and their respective estimation methods.

Since we used ML estimators we know that our estimators $(\theta_n)_{n\in\mathbb{N}}$ form an asymptotically normal sequence of estimators for the true parameter $\theta_0 \in \Theta \subset \mathbb{R}_m$ with positive definite covariance matrix Σ. So

$$\sqrt{N}\left(\theta_n - \theta_0\right) \to \mathcal{N}_m\left(0, \Sigma\right).$$

Since $\theta \mapsto \mathbb{E}_\theta(X)$ is continuously differentiable and $\nabla\mathbb{E}_{\theta_0} \neq 0$, we can approximate its distribution by the so-called Delta-method (see Shao [1999], page 45). So

$$\sqrt{N}\left(\mathbb{E}_{\theta_n}(X) - \mathbb{E}_{\theta_0}(X)\right) \to \mathcal{N}\left(0, \left(\nabla\mathbb{E}_{\theta_0}\right)' \Sigma \nabla\mathbb{E}_{\theta_0}\right)$$

For the risk-capturing functional $\theta_n \star AVaR_\alpha(X)$ we can thus calculate the AVaR as for a normally distributed variable

$$\theta_N \star AVaR_\alpha(X) \approx \mathbb{E}_{\theta_0}(X) + \frac{\varphi\left(\Phi^{-1}(\alpha)\right)}{\alpha\sqrt{N}} \sqrt{\left(\nabla\mathbb{E}_{\theta_0}\right)' \Sigma \nabla\mathbb{E}_{\theta_0}},$$

with φ (resp. Φ) density (resp. distribution) function of a standard Normal.

In our application, the evaluation expectation is given by the value of a strip of spread options. So we consider a time period for which we want to analyse the value of the power plant and our pricing functional is given as

$$VPP(t,T) = \int_t^T e^{-r(s-t)}\, V(s)\, \mathrm{d}s. \tag{5.8}$$

Using the above, we can employ the asymptotic distribution of our estimators to quantify parameter risk and employ the closed-form formula for the normal AVaR to compute risk-captured prices efficiently. We use the risk-captured prices induced by the AVaR w.r.t.. different significance levels $\alpha \in (0,1)$.

The general procedure reads as follows:

1. After we estimated all parameters of our price processes, we simulate the processes for the future time period we consider and compute the spark spread value $V(t)$ given in Equation (5.1) for every day t in the period. For our illustration we will consider a three year period starting immediately after the observation period.

2. Then, by fixing all parameters except the one of interest (which we generically denote be θ) and setting the shift value ξ (e.g. $\xi = 1\%$), we compute shifted up and down spark spread values, i.e. $V_t^{up}(\theta + \xi)$ and $V_t^{down}(\theta - \xi)$.

3. Further, we compute the value of the power plant (VPP) by means of Monte Carlo simula-

tions. For a fixed large number N and a fixed period $T = 3$ years we have

$$VPP(t,T) = \frac{1}{N}\sum_{i=1}^{N} VPP_i(t,T), \tag{5.9}$$

where from (5.8)

$$VPP_i(t,T) = \int_t^T e^{-r(s-t)}\, V_i(s)\, \mathrm{d}s,$$

with i referring to the simulation run. At this step we also compute power plant values where the parameter θ is shifted by some some value of $\xi > 0$, i.e. $VPP^{up}(t,T;\theta) := VPP(t,T;\theta+\xi)$ and $VPP^{down}(t,T;\theta) := VPP(t,T;\theta-\xi)$ (e.g. w.r.t. shifted spark spread values at each time point as in the previous step) and estimate the sensitivity (the derivative) of the VPP with respect to the parameter θ with the central finite difference

$$\nabla_\theta VPP := \frac{\partial VPP(\theta)}{\partial\theta} \approx \frac{VPP^{up}(t,T;\theta) - VPP^{down}(t,T;\theta)}{2\cdot\xi}. \tag{5.10}$$

4. Finally, we compute the bid and ask prices, i.e. we use a closed-form approximation formula for the AVaR to get the risk-captured prices by subtracting and adding risk-adjustment value to $VPP(t,T)$ respectively. For a specified significance level $\alpha \in (0,1)$ this risk-adjustment value is computed as follows

$$\frac{\varphi(\Phi^{-1}(1-\alpha))}{\alpha}\sqrt{\frac{(\nabla_\theta VPP)'\cdot\Sigma_\theta\cdot\nabla_\theta VPP}{N}},$$

denoting by Σ_θ the asymptotic covariance matrix of the estimator for the parameter θ.[1]

5.5 Illustrative example

We consider the financial equivalent of a gas fired power plant. Our objective is to analyse the impact of model risk for a three year valuation period starting immediately after the estimation period. For this we can simulate the price processes throughout the future period and calculate the price of the clean spark spread for every day within the period. The value is then given by Equation (5.9).

[1] Note that this (normally approximated) risk-adjustment value does only depend on the significance level α w.r.t. the factor $\varphi(\Phi^{-1}(1-\alpha))/\alpha$. In particular, the ratio of AVaR-induced risk-captured bid-ask spreads w.r.t. different significance levels is constant, regardless of the parameter in doubt. This is a crucial property of the AVaR w.r.t. a Normal distribution.

5.5.1 Risk values results

In Equations (5.2)-(5.6) we discussed the estimation of the relevant parameters which all are sources of parameter risk. We study these risks separately since a joint assessment would require knowledge of the joint estimator's distributions, which is not available in closed form. Below the discussion of every source of risk is given together with illustrative figures. In Table 5.2 all final values of the relative bid-ask spreads for every risk source are given. One can clearly see that the major and critical source of risk lies in the jump size distribution. Even when we measure the model risk in all parameters simultaneously, except of the jump size distribution, we see that still the values are very small compared to those which are from the jump size distribution alone.

- *parameter risk in spike size – Laplace distribution assumed for jumps* - The spikes, in many cases with an upward movement, have a large impact on the option evaluation. In case of an upward spike, the option usually jumps into the money and the option value rises. Furthermore, the distribution of the spike size is more difficult to determine due to the small sample size; the threshold filtering produces only a small sample size of jumps to estimate the jump size distribution from. Hence, particularly the scale of the spikes is crucial for valuation purposes. When incorporating the parameter risk of the spike distribution, regarding both the location parameter μ as well as the scale parameter σ, one obtains a relative difference of 49% for a safety level of $\alpha = 50\%$. If one is more risk averse towards parameter risk in the jump size distribution (e.g. a safety level of $\alpha = 10\%$), one even obtains a relative width of the bid-ask spread of 107%.

- *parameter risk in spike size – Normal distribution assumed for jumps* - Again, also in case of the Normal distribution, the spikes have a large impact, although the impact is a bit smaller than in the case of Laplacian jumps. For a safety level of 50%, one obtains a relative width of the parameter risk-implied bid-ask spread of 33%, while this spread widens considerably with increasing risk aversion (e.g. 73% relative width of the bid-ask spread for a safety level of 10%).

- *parameter risk in correlation* - The correlation is a major driver steering the width of the spread between the gas and the power price. Using the AVaR w.r.t. a security level of 10% employing the Fisher transform of the correlation estimator, the relative difference between the risk-captured bid and ask prices is 2.17% in case of Laplace distributed jumps and 4.68% in case of normally distributed jumps.

- *parameter risk in gas signal* - The gas price process is a minor driver of the bid-ask price width. Using the AVaR w.r.t. a security level of 10%, the relative difference between the risk-captured bid and ask prices is approx. 4% in case of normally distributed jumps and 1.92% in case of Laplace distributed jumps.

- *joint parameter risk in gas and base power signal* - The joint continuous signal of gas and electricity prices is also a minor driver of the bid-ask price width. Using the AVaR w.r.t. a security level of 10%, the relative difference between the risk-captured bid and ask prices is approx. 4.2% in case of normal distributed jumps and approx. 2% in case of Laplace distributed jumps.
- *joint parameter risk in gas, power, and emissions (all processes, except of jump size parameter)* - As was shown before, the spike size distribution is a major risk factor when modelling the bid-ask price width. Therefore, it would be interesting to re-check this from the opposite side by considering the joint risk in every driver except of spike size and intensity. Using the AVaR w.r.t. a security level of 10%, the relative difference between the risk-captured bid and ask prices is approx. 5.5% in case of normal distributed jumps and approx. 2.6% in case of Laplace distributed jumps.

Table 5.2 shows all sources of investigated model risk for various confidence levels α and shift sizes ξ. Due to the large sample size of $M = 790$ observations, the estimation of the base signals in gas and power has high accuracy, provided that the parametric form of the model is correctly chosen. Hence, parameter risk from the estimation of the parameters in base gas and power signals is very moderate and the parameter risk-captured bid-ask spreads are relatively narrow. A completely different picture is shown for the estimation of the spike size. Due to the threshold filtering technique, the number of spikes is relatively small ($\tilde{M} = 41$), which naturally enlarges the variance. Furthermore, the jump size distribution is crucial for ensuring that the real option representing the gas power plant gets deep into the money: When there is a large upward spike in the power price process, the payoff of the real option immediately jumps into the money. Hence, the probability of producing large upward spikes in the power price plays a major role in determining the future value of a gas power plant.

Besides the considered risk sources, one should also care for other model risk factors. In this chapter for the sake of simplicity we use models with non-stochastic volatilities, constant mean-reversion forces, parsimonious filtering procedure, and fixed spike intensity. Therefore, we ignore these risk drivers to identify some sort of a low boundary value for the potential model and parameters risk in terms of chosen modelling frames.

5.5.2 Absolute and relative bid-ask prices

In this section we illustrate our results. We start by investigating how the confidence level α influences the bid and ask prices. By taking a large number of simulations $N = 5000$ and confidence level $\alpha = [0.01, 0.1, 0.5]$ we can see how the risk-captured bid and ask prices behave. Together with bid and ask prices, we also investigate the relative width of the bid-ask spread which is computed as

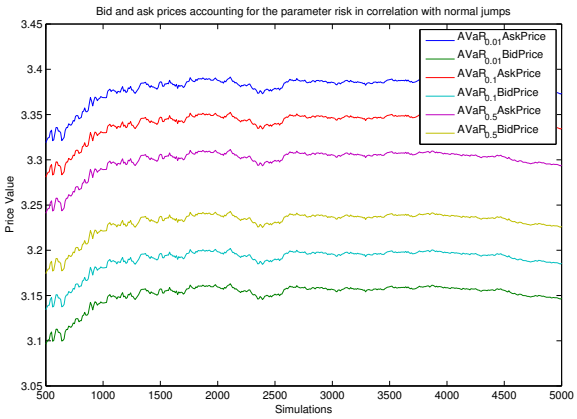
$$\Delta = \frac{bidPrice - askPrice}{midPrice}.$$

Model Risk	Jump size distribution	Gaussian				Laplace				
		α_1	α_2	α_3	ξ_2	ξ_3	α_1	α_2	α_3	ξ_2	ξ_3
Jumps	111.9%	73.7%	33.5%	73.7%	73.8%	163.5%	107.7%	48.9%	109.3%	124.5%
Correlation	6.9%	4.6%	2.1%	4.6%	4.6%	3.3%	2.2%	0.9%	2.2%	2.2%
Gas and power base	6.5%	4.3%	1.9%	4.3%	4.3%	3.1%	2%	0.9%	2%	2%
Gas	6.1%	4%	1.8%	4%	4%	2.9%	1.9%	0.9%	1.9%	1.9%
Gas, power and carbon	8.2%	5.5%	2.6%	5.4%	5.4%	3.9%	2.6%	1.5%	2.6%	2.6%

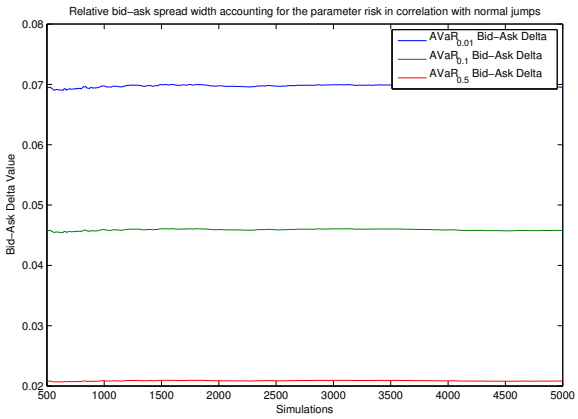
Table 5.2: Resulting values for the relative width of the bid-ask spread for various model risk sources in %. $\alpha_1 = 0.01$ (the highest risk-aversion), $\alpha_2 = 0.1$, $\alpha_3 = 0.5$, $\xi = 1\%$ (the smallest change value), $\xi_2 = 5\%$, $\xi_3 = 15\%$. When we did the case $\xi = 1\%$, we took $\alpha = \alpha_2$, therefore we omit here the column with $\xi = 1\%$.

Furthermore, we want to assess the numerical problems associated with approximating the sensitivity with the central finite difference as described in Equation (5.10).

- Correlation. We assess the parameter risk in correlation for both cases where the jump sizes follow either a Normal or a Laplace distribution. Figures 5.4 and 5.5 show that the parameter-risk implied bid-ask spread w.r.t. the correlation is in the order of few percentage points, mainly due to the large sample size.

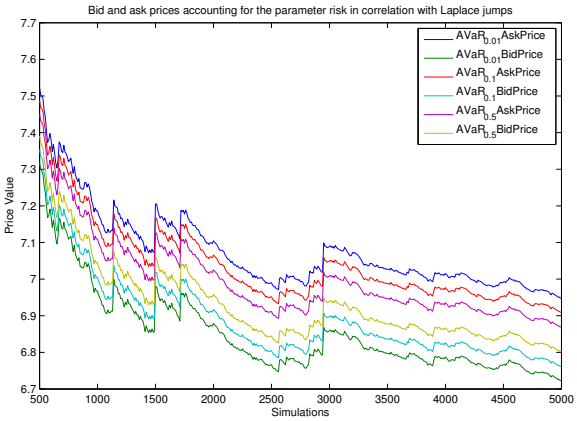


(a) Risk-captured bid and ask prices for different α levels when changing the **correlation** parameter value with a shift size of 1%, normal jumps.

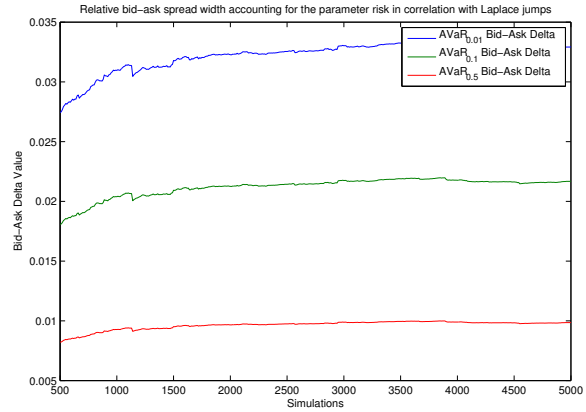


(b) Relative width of the risk-captured bid-ask spread for different α values when changing the **correlation** parameter value with a shift size of 1%, normal jumps.

Figure 5.4: Parameter-risk implied bid-ask spread w.r.t. the correlation.



(a) Risk-captured bid and ask prices for different α levels when changing the **correlation** parameter value with a shift size of 1%, Laplace jumps.

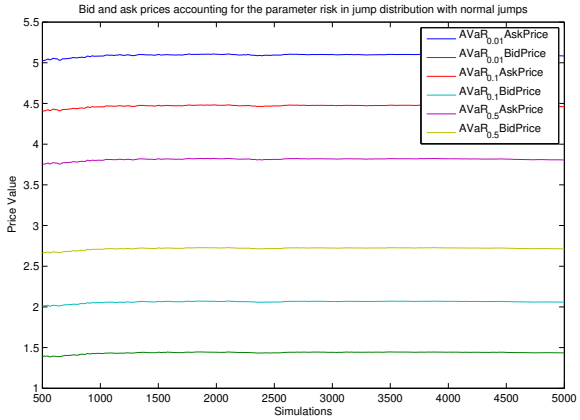


(b) Relative width of the risk-captured bid-ask spread for different α values when changing the **correlation** parameter value with a shift size of 1%, Laplace jumps.

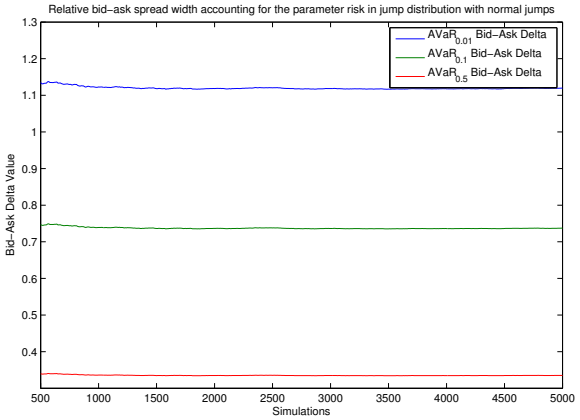
Figure 5.5: Parameter-risk implied bid-ask spread w.r.t. the correlation.

- Jump size distribution parameters. We assess the parameter risk in the jump size distributions for both cases where the jump sizes follow either a normal or a Laplace distribution. This effect is depicted in Figures 5.6 and 5.7, we check the robustness of the estimate of the sensitivity using different central differences. One can see that the parameter-risk im-

plied bid-ask spread w.r.t. the jump size distribution is crucial – the correct determination of the jump size distribution is the major driver for the power plant price. Furthermore, convergence is much slower in the Laplace case and the results become much less stable.

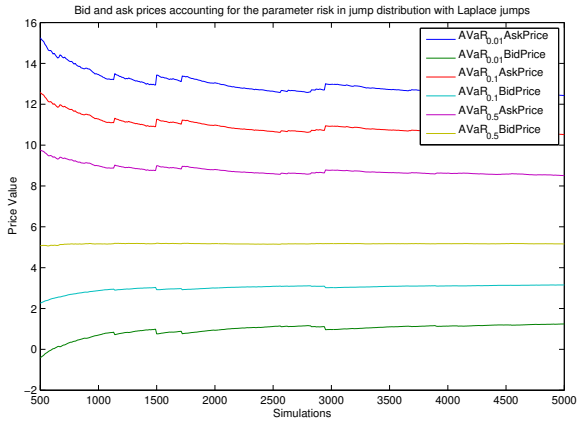


(a) Risk-captured bid and ask prices for different α levels when changing the **jump size** distribution parameter values with a shift size of 1%, normal jumps.

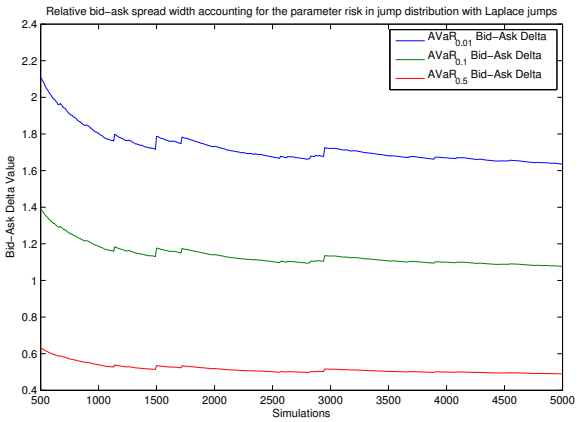


(b) Relative width of the risk-captured bid-ask spread for different α levels when changing the **jump size** distribution parameter values with a shift size of 1%, normal jumps.

Figure 5.6: Parameter-risk implied bid-ask spread w.r.t. the jump size distribution.



(a) Risk-captured bid and ask prices for different α levels when changing the **jump size** distribution parameter values with a shift size of 1%, Laplace jumps.



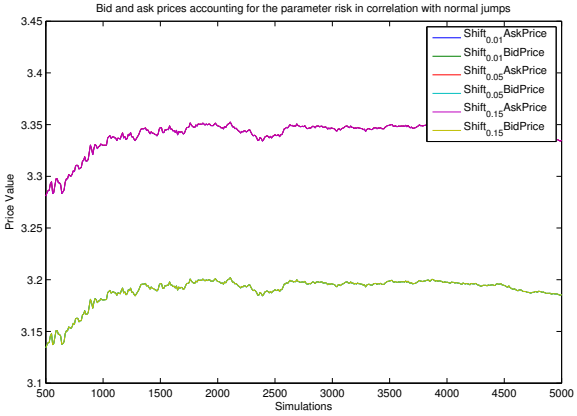
(b) Relative width of the risk-captured bid-ask spread for different α values when changing the **jump size** distribution parameter values with a shift size of 1%, Laplace jumps.

Figure 5.7: Parameter-risk implied bid-ask spread w.r.t. the jump size distribution.

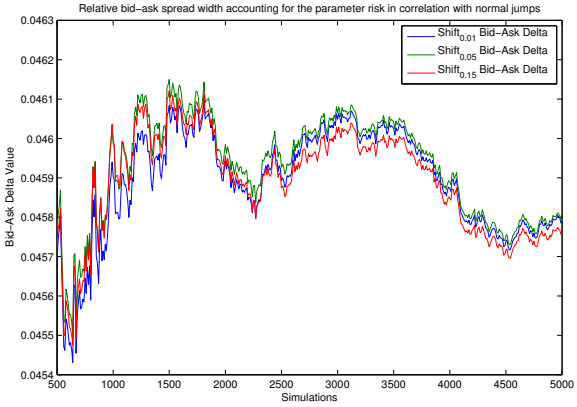
Sensitivity impact

- Correlation. The impact of the choice of numerical approximation of the sensitivity is minor, as can be seen in Figures 5.8 and 5.9;

- Jumps size distribution parameters. The choice of numerical approximation of the sensitivity has minor impact which is demonstrated in Figures 5.10 and 5.11. Compared to the normal case, one sees considerably more problems in estimating the sensitivity w.r.t. the jump size parameters.

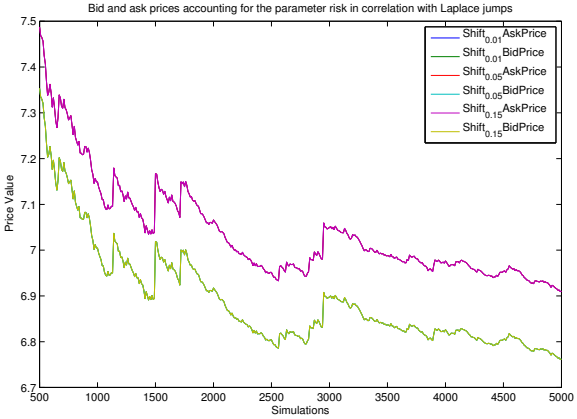


(a) Risk-captured bid and ask prices for different shift values ξ when changing the **correlation** parameter value with a significance level $\alpha = 10\%$, normal jumps.

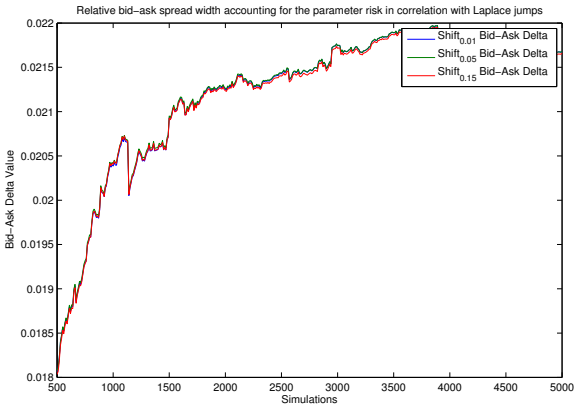


(b) Relative width of the risk-captured bid-ask spread for different shift values ξ when changing the **correlation** parameter value with a significance level $\alpha = 10\%$, normal jumps.

Figure 5.8: Parameter-risk implied bid-ask spread w.r.t. the sensitivity value (correlation).



(a) Risk-captured bid and ask prices for different shift values ξ when changing the **correlation** parameter value with a significance level $\alpha = 10\%$, Laplace jumps.

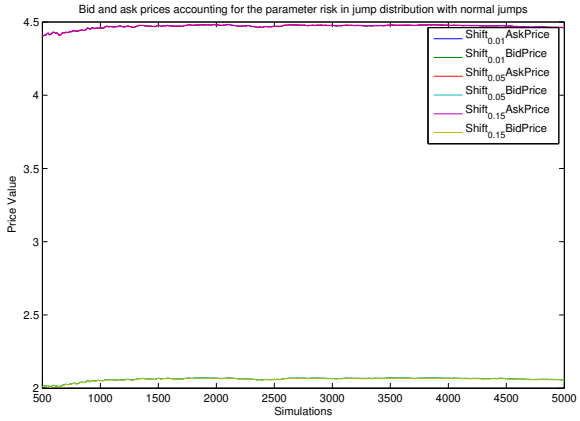


(b) Relative width of the risk-captured bid-ask spread for different shift values ξ when changing the **correlation** parameter value with a significance level $\alpha = 10\%$, normal jumps.

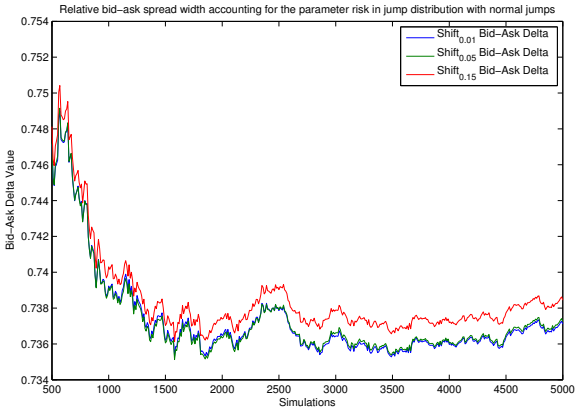
Figure 5.9: Parameter-risk implied bid-ask spread w.r.t. the sensitivity value (correlation).

5.6 Conclusion

We studied the model risk inherent in power plant valuation within a framework of risk-capturing functionals. Our study reveals that spike risk is the most important source of model risk. While

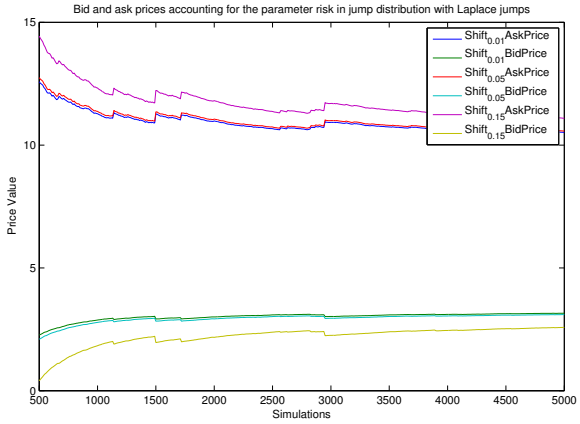


(a) Risk-captured bid and ask prices for different shift sizes when changing the **jump size** distribution parameter values with a significance level $\alpha = 10\%$, normal jumps.

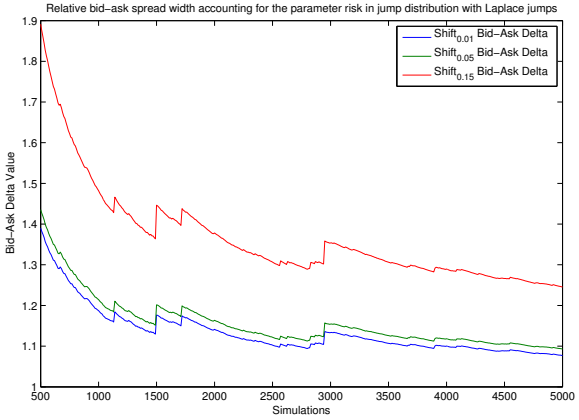


(b) Relative width of the risk-captured bid-ask spread for different shift sizes when changing the **jump size** distribution parameter values with a significance level $\alpha = 10\%$, normal jumps.

Figure 5.10: Parameter-risk implied bid-ask spread w.r.t. the sensitivity value (jump size distribution).



(a) Risk-captured bid and ask prices for different shift sizes when changing the **jump size** distribution parameter values with a significance level $\alpha = 10\%$, Laplace jumps.



(b) Relative width of the risk-captured bid-ask spread for different shift sizes when changing the **jump size** distribution parameter values with a significance level $\alpha = 10\%$, Laplace jumps.

Figure 5.11: Parameter-risk implied bid-ask spread w.r.t. the sensitivity value (jump size distribution).

this is not surprising in itself we were able to quantify the magnitude by which spike risk dominates all other sources of model risk.

Given that the increasing impact of renewable energies will make power prices more volatile and jumpy we see the importance of addressing the spike behaviour for valuation purposes.

We plan to apply our methodology for further applications in the energy markets such as generation of an hourly power forward curve and valuation procedures for storages.

CHAPTER 6

CONCLUSIONS

In this study we explored various aspects of stochastic modelling energy-related markets.We now summarise all the findings of our work and state the importance of them to practical applications.

We investigated statistical properties of power price modelling on the example of critical comparison of widely used three continuous-time electricity spot price models. We found that power price dynamics exhibits several types of behaviour: small/average daily fluctuations and spikes. Modelling power price as a jump-diffusion is not the best option due to the only one mean-reversion force, since it leads to a too slow mean-reversion for the spikes and a too fast mean-reversion for the base signal. The result of this is higher volatility. The better way to capture this mean-reverting nature is to introduce two mean-reversion speeds: one for the base and one for the spike signals as does the factor model, which in turn results in better modelling of the path behaviour in terms of the descriptive statistics. For the modelling of daily fluctuations Brownian component is rather satisfactory than an OU process with a background driving Lévy process such that OU process has the same stationary distribution. The latter one does not capture the variability of the paths appropriately and underestimates the noise in the base signal.

We also studied analytical tractability for all three models on the example of pricing forward prices. The factor model outperforms the jump-diffusion and the threshold models, since it allows for a straightforward derivatives calculation. The jump-diffusion model has a limited ability to compute forward prices in closed form, since it is only possible for some classes of distributions assumed for a jump size. The threshold model does not allow for explicit forward formula due to the presence of the state-dependent function h. There are quite a few possibilities to tackle this issue. One of them is to use efficient numerical and simulation-based Monte-Carlo evaluations. Comparing obtained forward prices for all three models we observe that the factor model converges towards the predicted spot price when we approach delivery time, whereas the threshold and the jump-diffusion models drive apart, which is also reflected in a risk premium with increasing absolute value for the jump-diffusion and threshold models, and whereas decreasing to zero for the factor model.

To continue our pricing investigation we derived an integro-PDE for the threshold model to pur-

sue its derivative pricing ability and consequently to allow for further hedging purposes. This equation gives the desired forward price dynamics important for hedging purposes which we compare with forwards obtained for the jump-diffusion model. When numerically solving this integro-PDE we implemented finite difference scheme and derived estimates for the truncation errors (domain truncation and integral truncation). We also investigated what happens if the x (log of the price) goes to infinity and how does it effect the forward prices? Having numerical scheme at hand and assuming some boundary conditions, we observe that the threshold forward price dynamics is insignificantly different from the jump-diffusion forwards dynamics. The bigger difference arises when we use Laplace distribution compared to Normal distribution for a jump size. We also found that in contrast to modelling spot prices, the effect of function h is minor, i.e. it does not demonstrate a prominent difference between forward prices when they are below the threshold.

We the continued with examining the stochastic modelling of energy markets and proposed a new storage model to complement existing approach to storage value modelling. The main finding here is that we consider storage level as a bounded stochastic process which lives between boundary l and u. We mean here a physical storage facility, hydro or gas, which is naturally bounded by a total capacity. A power producer and/or storage owner regularly takes a decision to withdraw and produce power, to inject (in case of gas), or to wait for a more favourable market conditions (higher power prices). Usually, this policy is a result of an optimal stochastic control problem. In our case, we looked at the problem differently and assumed that this policy is already given by a dynamics of storage level. With this at hand, we were able to derive a transition probability density for this storage level process and constructed various payoffs useful for hedging purposes in case of too high or too low prices and too high or too low reservoir levels. The main benefit of such approach is that it allows for a direct calculation of the storage value overcoming various numerical difficulties associated with an implementation of searching for an optimal control policy.

This topic offers a variety of possible further research problems which goes beyond the scope of this thesis. Among these, for example, would be a modification of the assumption on the constant mean-reverting level m and its substitution to a time-dependent component $m(t)$ to capture the seasonality of the precipitation inflow. Our approach also allows for a reservoir-dependent rate of production. For the sake of simplicity we considered only constant value for it. It would be beneficial to investigate it role and understand what is its impact on the storage value. Finally, it would be crucial to benchmark our model against a classical stochastic optimal control approach and to compare results. This will be the scope of our future research.

The last part of our conclusive chapter overviews results on detection of the model risk of energy markets. We considered a gas-driven power plant and studied the model risk within a framework of risk-capturing functionals. Our findings proved a quite natural assumption that spike risk is the most important source of model risk. We were able to quantify the magnitude by which

spike risk dominates all other sources of model risk. We also found that the second significant source of risk belongs to a correlation between energy commodities. Given that the increasing impact of renewable energies will make power prices more volatile and jumpy we especially emphasise the significance of addressing the spike behaviour for valuation purposes. A future research plan is to apply our methodology for further applications in the energy markets such as generation of an hourly power forward curve and valuation procedures for storages.

Appendices

APPENDIX A

ELECTRICITY MODELLING

A.1 Lévy processes

We start with the following definitions given in Chapter 3 in Cont & Tankov [2004].
Definition 4 (Lévy process). *A cádlág stochastic process $(L_t)_{t\geq 0}$ on $(\Omega, \mathbb{P}, \mathcal{F}, \{\mathcal{F}_t\})$ with values in $\mathbb{R}$ such that $L_0 = 0$ is called a Lévy process if it satisfies the following properties:*

1. *Independent increments: for every increasing sequence of times $t_0, \ldots, t_n$ the random variables $X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent;*
2. *Stationary increments: the law of $X_{t+h} - X_t$ does not depend on t;*
3. *Stochastic continuity: $\forall \epsilon > 0$, $\lim_{h\to 0} \mathbb{P}\left(|X_{t+h} - X_t| \geq 0\right) = 0$.*

A Lévy process is associated with its Lévy measure ν
Definition 5 (Lévy measure). *Let $(L_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}$. The measure ν on $\mathbb{R}$ defined by:*

$$\nu(A) = \mathbb{E}\left[\#\{t \in [0,1] : \Delta L_t \neq 0, \quad \Delta L_t \in A\}\right], \quad A \in \mathcal{B}(\mathbb{R}) \tag{A.1}$$

is called the Lévy measure of L_t: $\nu(A)$ is the expected number, per unit time, of jumps whose size belongs to A.

Every Lévy process is characterised by its characteristic triplet (γ, b, ν), where $\gamma \in \mathbb{R}$ is the drift term, $b \in \mathbb{R}_{\geq 0}$ is the diffusion coefficient and ν is the Lévy measure. Next useful theorem is a celebrated Lévy-Khinchin decomposition.
Theorem 3 (Lévy triplet). *Let $(L_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}$ with a characteristic triplet (γ, b, ν). Then*

$$\mathbb{E}\left[e^{izL_t}\right] = e^{t\,\psi(z)}, \quad z \in \mathbb{R}, \tag{A.2}$$

with

$$\psi(z) = i\gamma z - \frac{z^2 b}{2} + \int_{\mathbb{R}} \left(e^{izx} - 1 - izx\mathbf{1}_{|x|<1}\right)\, \nu(\mathrm{d}x). \tag{A.3}$$

Proof. The proof can be found in various sources, for example in Chapter 3 in Cont & Tankov [2004] and in Chapter 4 in Sato [1999]. □

According to this theorem a Lévy process can always be decomposed into several processes: a deterministic linear process (drift) with parameter γ, a Brownian motion with coefficient $\sqrt{b}$, a compound Poisson process with arrival rate $\lambda := \nu(\mathbb{R}\backslash(-1,1))$ and jump size distribution given by its cumulative distribution function $\mathbb{F}(\mathrm{d}x) := \frac{\nu(\mathrm{d}x)}{\nu(\mathbb{R}\backslash(-1,1))}\mathbf{1}_{|x|\geq 1}$, and the last component: pure jump martingale process.

Now let us recall a few facts about the Lévy measure ν:

- the Lévy measure ν on $\mathbb{R}$ satisfies $\nu(\{0\}) = 0$ and $\int_{\mathbb{R}} \left(1 \wedge (|x|)^2\right) \nu(\mathrm{d}x) < \infty$.
- if ν is a finite measure, i.e. $\lambda = \nu(\mathbb{R}) = \int_{\mathbb{R}} \nu(\mathrm{d}x) < \infty$, then $\mathbb{F}(\mathrm{d}x) := \frac{\nu(\mathrm{d}x)}{\lambda}$ is a probability measure. Then λ is interpreted as the expected number of jumps and $\mathbb{F}(\mathrm{d}x)$ is the distribution of the jump size x. It is also said the the Lévy process L_t has finite activity. (Theorem 21.3 in Sato [1999]).
- if $b \neq 0$ or $\int_{|x|\leq 1} |x| \ \nu(\mathrm{d}x) = \infty$, then almost all the paths of the Lévy process L_t have infinite variation. (Theorem 21.9 in Sato [1999]).

Definition 6 (Subordinator)**.** *A Lévy process $(L_t)_{t\geq 0}$ on $\mathbb{R}$ is said to be increasing if L_t is increasing as a function of t, a.s. An increasing Lévy process is called a subordinator.*

Proposition 5 (Characteristic triplet of a subordinator)**.** *Let $(L_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}$. The following conditions are equivalent:*

1. *$L_t \geq 0$ a.s. for some $t > 0$.*
2. *$L_t \geq 0$ a.s. for every $t > 0$.*
3. *Sample paths of L_t are a.s. nondecreasing: $t \geq s \quad \Rightarrow \quad L_t \geq L_s$ a.s..*
4. *The characteristic triplet of L_t satisfies $\gamma = 0$, $\nu((-\infty, 0]) = 0$, $\int_0^\infty (x \wedge 1) \ \nu(\mathrm{d}x) < \infty$ and $b \geq 0$, L_t has only positive jumps of finite variation and positive drift.*

Proof. Proposition 3.10 in Cont & Tankov [2004]. □

As example, Poisson process is a subordinator.

Proposition 6 (Exponential moments of a Lévy process)**.** *Let $(L_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}$ with a characteristic triplet (γ, b, ν). Then*

1. *$\mathbb{E}[|L_t|^p] < \infty$ if and only if $\int_{|x|\geq 1} |x|^p \ \nu(\mathrm{d}x) < \infty$;*
2. *$\mathbb{E}[e^{pL_t}] < \infty$ if and only if $\int_{|x|\geq 1} e^{px} \ \nu(\mathrm{d}x) < \infty$.*

Proof. Theorem 25.3 in Sato [1999]. □

We now continue with a few more known facts and properties of a Lévy process and its connection to the martingale theory.

- a *semimartingale* is a stochastic process $(X_t)_{0\le t\le T}$ that can be represented as

$$X = X_0 + M + A,$$

  where X_0 is finite and $\mathcal{F}$-measurable, M is a local martingale with $M_0 = 0$ and A is a finite variation process with $A_0 = 0$;

- a semimartingale X is a *special semimartingale*, if the process A is predictable;

- every Lévy process is a semimartingale due to its Lévy-Khinchin decomposition;

- every Lévy process with its finite first moment (i.e. if and only if $\int_{|x|\ge 1} |x|\ \nu(\mathrm{d}x) < \infty$;) is also a special semimartingale;

- the following three assertions are equivalent:

  1. a Lévy process L_t is a special semimartingale;
  2. $\int_{\mathbb{R}} \left(|x| \wedge |x|^2\right)\ \nu(\mathrm{d}x) < \infty$,
  3. $\int_{\mathbb{R}} \left(|x|\,\mathbf{1}_{|x|\ge 1}\right)\ \nu(\mathrm{d}x) < \infty$,

  this is a consequence of Lemma 2.8 in Kallsen & Shiryaev [2002].

Theorem 4 (g-moments of supremum of a Lévy process)**.** *Let $(L_t)_{t\ge 0}$ be a Lévy process on $\mathbb{R}$. Define*

$$L_t^* = \sup_{s\in[0,t]} |L_s|\,.$$

Let $g(r)$ be a nonnegative continuous submultiplicative function on $[0,\infty)$, increasing to ∞ as $r\to\infty$. Then the following four statements are equivalent:

1. $\mathbb{E}[g(L_t^*)] < \infty$ *for some* $t > 0$*;*
2. $\mathbb{E}[g(L_t^*)] < \infty$ *for every* $t > 0$*;*
3. $\mathbb{E}[g(|L_t^*|)] < \infty$ *for some* $t > 0$*;*
4. $\mathbb{E}[g(|L_t^*|)] < \infty$ *for every* $t > 0$*.*

Proof. Theorem 25.18 in Sato [1999]. □

A.2 Prediction-based estimating functions method

We follow the steps in Sørensen [2000]. We denote by $Z_1, Z_2, \dots Z_n$ the stochastic processes of the base signal, which we will model by Y_1 with a Gamma stationary distribution. There are three parameters to be estimated, $\theta = (\lambda_1, \alpha, \nu)$, where the last two are the parameters of the distribution, and λ_1 is the speed of mean-reversion of Y_1. Assume that $f_j, j = 1 \dots N$ are one-dimensional functions, defined on the state space of Y_1, such that $\mathbb{E}_\theta\{f_j(Z_i)^2\} < \infty$ for all $\theta \in \Theta$, the parameter space, and $j = 1, \dots N, i = 1, \dots, n$. For given θ write expectation as $\mathbb{E}_\theta$. Let $\mathcal{F}_i$ be the σ-algebra generated by $Z_1, Z_2, \dots Z_i$, and $\mathcal{H}_i^\theta$ the L^2-space of square integrable $\mathcal{F}_i$-measurable one-dimensional random variables given that θ is the true parameter value. We denote the set of square-integrable predictors of $f_j(Z_{i+1})$ given $Z_1, Z_2, \dots Z_i$ by $\mathcal{P}_{i,j}^\theta, j = 1, \dots N$. Observe that this is a closed linear subspaces of $\mathcal{H}_i^\theta$. In the case of the factor model we use $N = 2$, and $f_1(y) = y$ and $f_2(y) = y^2$. These will be our prediction-based estimating functions.

Further one needs to choose an appropriate number q_{ij} of lags under which the prediction-based estimating function will be constructed. These numbers represent the available information "required" to predict the consecutive value, and in the case of a stationary process they do not need to be too large. We let $q_{ij} = 4$ for all i, j, i.e., four observations are taken to predict the following one. A space of predictors is specified as $U_j^{i-1} = (Z_{i-1}, \dots Z_{i-4})$ and $U_j^{i-1} = (Z_{i-1}^2, \dots Z_{i-4}^2)$ for $j = 1, 2$, resp.

Define the estimating function $G_n(\theta)$ as

$$G_n(\theta) = \sum_{i=1}^{n} \sum_{j=1}^{N} \Pi_j^{i-1}(\theta)\{f_j(Y_i) - \hat{\pi}_j^{(i-1)}(\theta)\}\,, \tag{A.4}$$

where $\Pi_j^{(i-1)}(\theta) = \{\pi_{1,j}^{(i-1)}(\theta), \pi_{2,j}^{(i-1)}(\theta), \pi_{3,j}^{(i-1)}(\theta)\}^T$ is a stochastic vector of weights, which belong to $\mathcal{P}_{i-1,j}^\theta$. The terms $\hat{\pi}_j^{(i-1)}(\theta)$ are the minimum mean-square error predictors of $f_j(Z_i)$ in $\mathcal{P}_{i-1,j}^\theta$. This predictor $\hat{\pi}_j^{(i-1)}(\theta)$ is the orthogonal projection of $f_j(Z_i)$ on $\mathcal{P}_{i-1,j}^\theta$ with respect to the inner product in $\mathcal{H}_i^\theta$. This projection exists and is uniquely determined by the normal equations

$$\mathbb{E}_\theta[\pi\{f_j(Z_i) - \hat{\pi}_j^{(i-1)}(\theta)\}] = 0\,, \tag{A.5}$$

for all $\pi \in \mathcal{P}_{i-1,j}^\theta$. From Equation (A.5) it follows that $G_n(\theta)$ is an unbiased estimating function. An estimator is obtained by solving $G_n(\theta) = 0$. We remark that the weights in $\Pi_j^{(i-1)}(\theta)$ serve the purpose of improving the efficiency of the estimator. The optimal choice of weights is a separate task considered in Bibby *et al.* [2010]. In our case these weights did not contribute much and we did not choose any since the algorithm reached convergence without them.

It is assumed that $\mathcal{P}_{i-1,j}^\theta$ is spanned by $U_{j_0}^{(i-1)}, \dots, U_{j_q}^{(i-1)}$ $j_k = 1, 2$, which are linearly independent in $\mathcal{H}_i^\theta$. By Equation (A.5) the minimum mean-square error predictor of $f_j(Z_i)$ in $\mathcal{P}_{i-1,j}^\theta$ is

given by

$$\hat{\pi}_j^{(i-1)}(\theta) = \hat{a}_{j0}^{(i-1)}(\theta) + \hat{a}_j^{(i-1)}(\theta)^T U_j^{(i-1)}\,, \tag{A.6}$$

where

$$\hat{a}_j^{(i-1)}(\theta) = C_{i-1,j}(\theta)^{-1} b_j^{(i-1)}(\theta)$$

and

$$\hat{a}_{j_0}^{(i-1)}(\theta) = \mathbb{E}_\theta\{f_j(Z_j)\} - \hat{a}_j^{(i-1)}(\theta)^T \mathbb{E}_\theta\{U_j^{(i-1)}\}\,.$$

Here $C_{i-1,j}(\theta)$ denotes the covariance matrix of $U_j^{(i-1)}$ when θ is the true parameter value, while

$$b_j^{(i-1)} = [\mathrm{Cov}_\theta\{U_{j_1}^{(i-1)}, f_j(Z_i)\}, \ldots, U_{j_q}^{(i-1)}, f_j(Z_i)\}]^T.$$

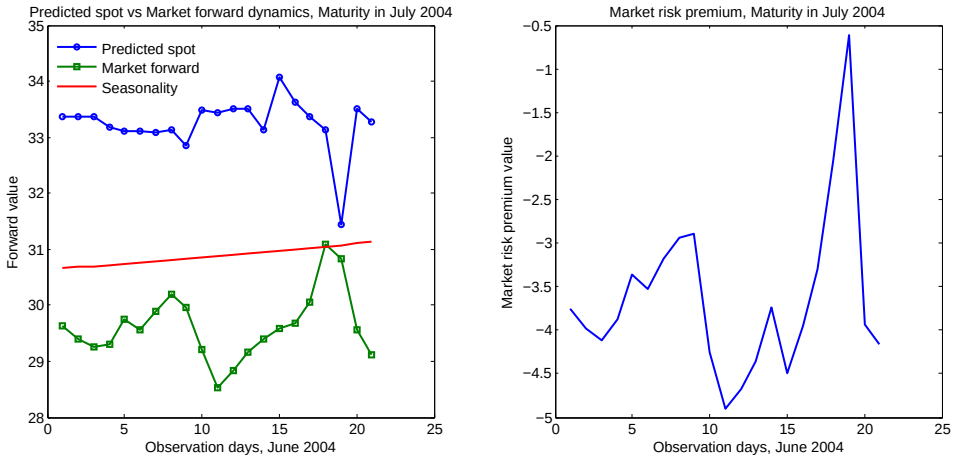
In conclusion, a prediction-based estimating function can be calculated provided that covariances in $C_{i-1,j}(\theta)$ and $b_j^{(i-1)}(\theta)$ can be computed. Since $\hat{\pi}_j^{(i-1)}(\theta)$ depends only on the first- and the second-order moments of the random vector $\{f_j(Z_i), U_{j_1}^{(i-1)}, \ldots, U_{j_q}^{(i-1)}\}$, only parameters appearing in these moments can be estimated using Equation (A.4). Observe that the characteristic function of $Y_1(t)$ is

$$\mathbb{E}[e^{uY_1(t)}] = e^{uY_1(0)e^{-\lambda t}} e^{\nu t(\alpha-\lambda)} \Big(\frac{\alpha - ue^{-\lambda t}}{\alpha - u}\Big)^{\nu}\,. \tag{A.7}$$

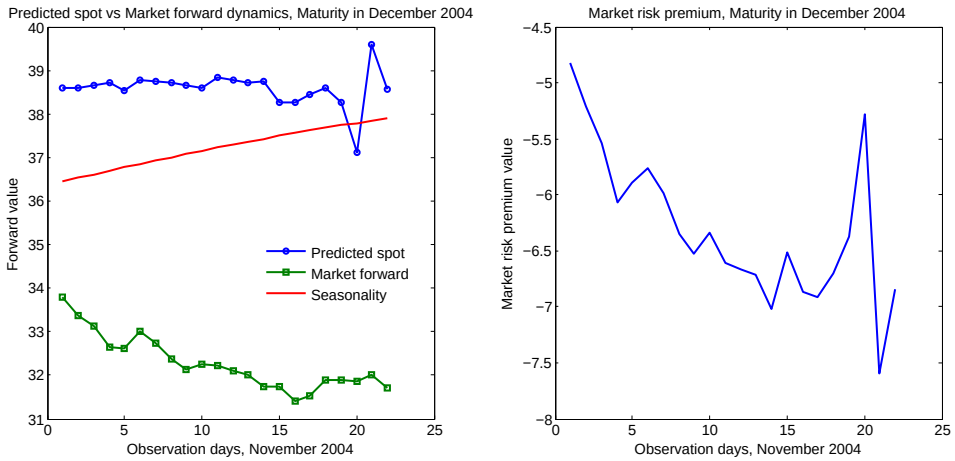
Therefore the moments can be obtained by taking the respective derivative of the characteristic function at $u = 0$. In order to apply the prediction-based estimating functions method based on the functions $f_1(y) = y$ and $f_2(y) = y^2$, it is necessary to obtain the four first moments, which can be computed explicitly for the Gamma stationary OU process. Besides expressions for moments, one needs to derive the covariance between the two first moments and the covariance between the two second moments, $\mathrm{Cov}(Z_t, Z_{t+s})$ and $\mathrm{Cov}(Z_t^2, Z_{t+s}^2)$.

A.3 Forward dynamics for various data sets

A.3.1 Jump-diffusion model and its forward modelling

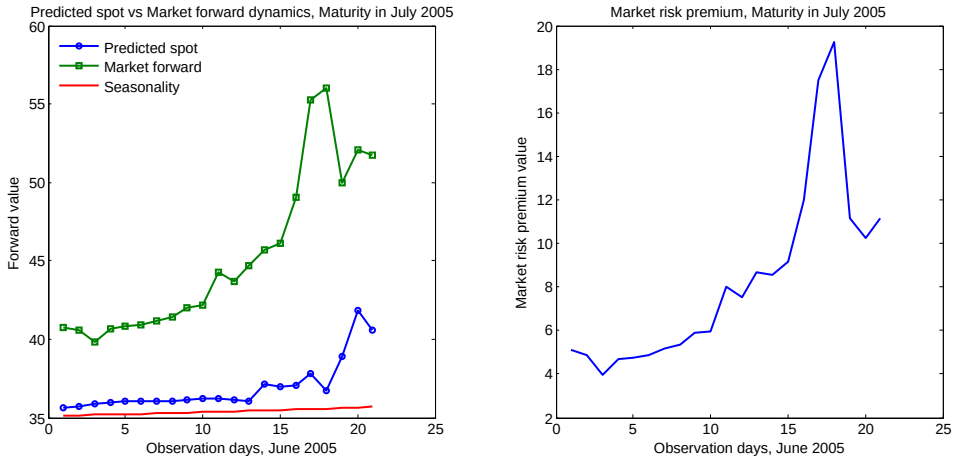


(a) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2004.

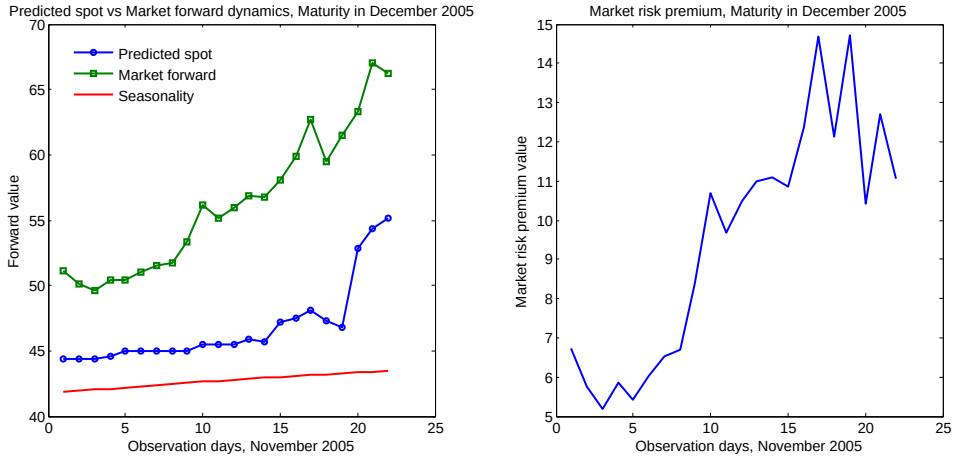


(b) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2004.

Figure A.1: The predicted spot, observed forward dynamics and market risk premium for the jump-diffusion model.

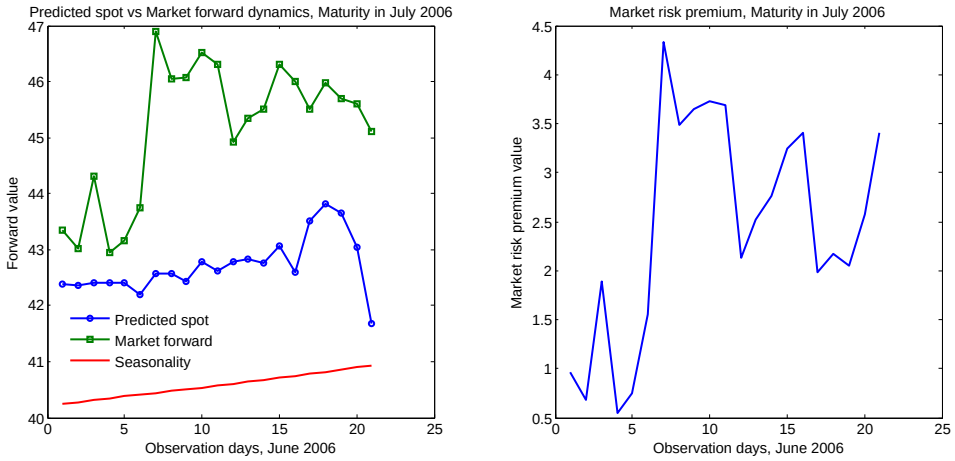


(c) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2005.

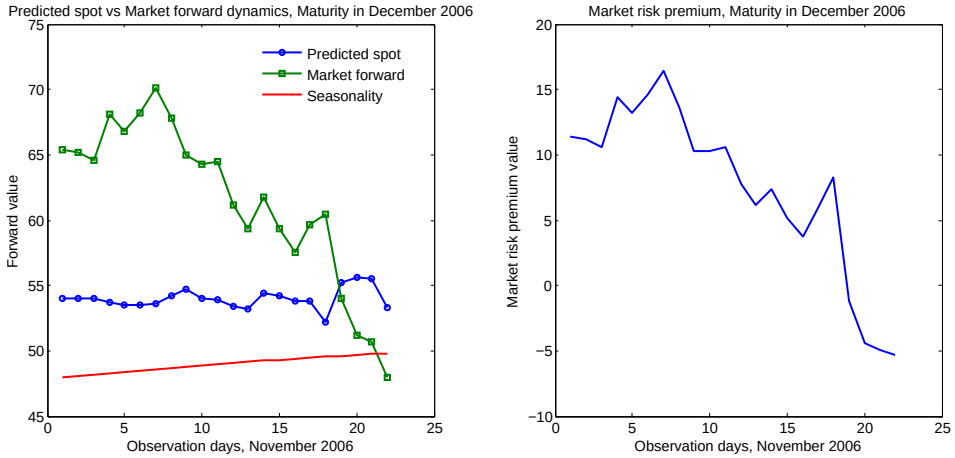


(d) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2005.

Figure A.0: The predicted spot, observed forward dynamics and market risk premium for the jump-diffusion model.



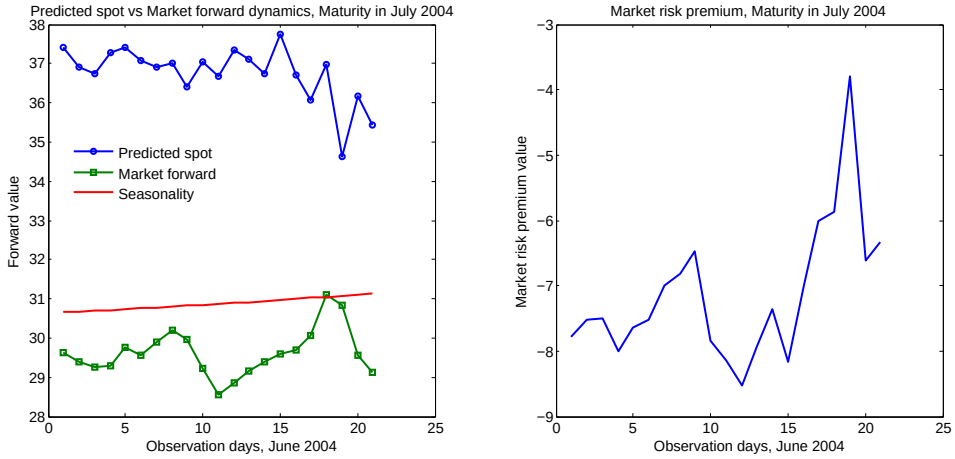
(e) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2006.



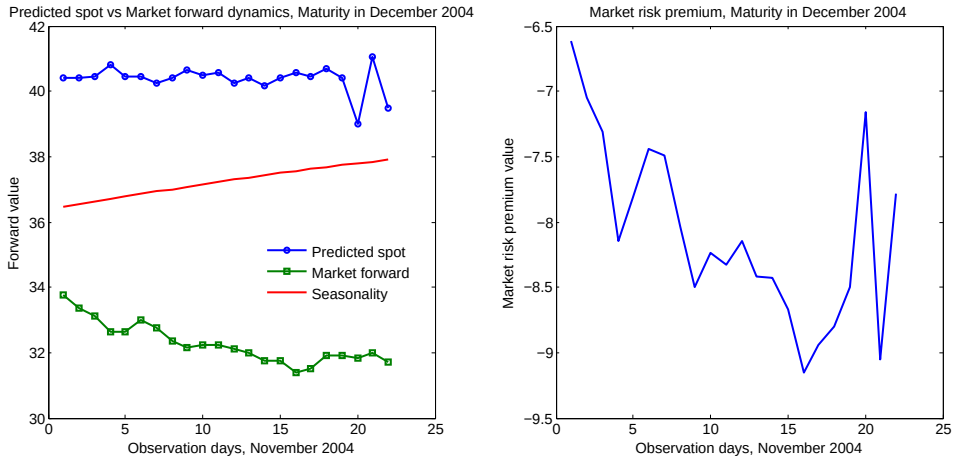
(f) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2006.

Figure A.-1: The predicted spot, observed forward dynamics and market risk premium for the jump-diffusion model.

A.3.2 Threshold model and its forward modelling

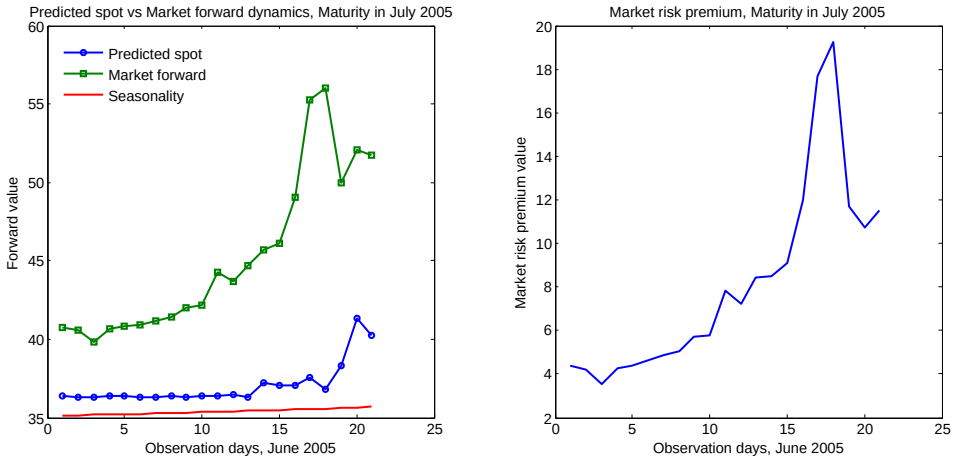


(a) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2004.

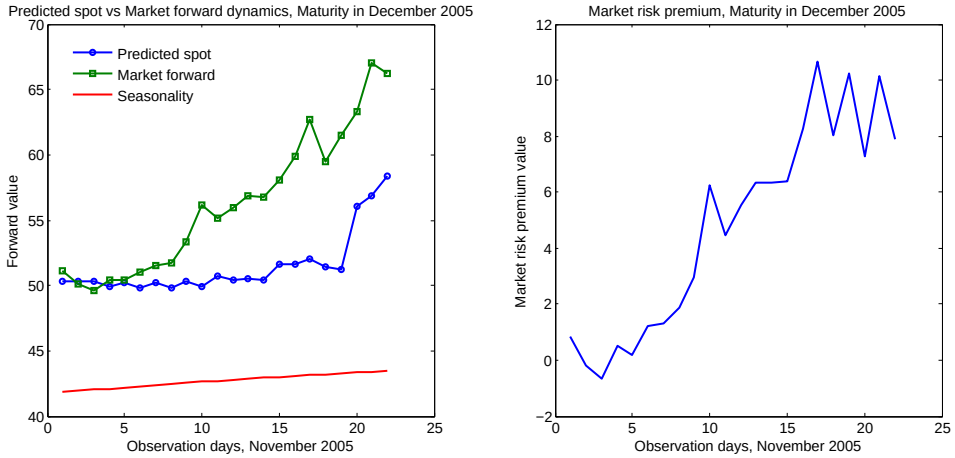


(b) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2004.

Figure A.0: The predicted spot, observed forward dynamics and market risk premium for the threshold model.

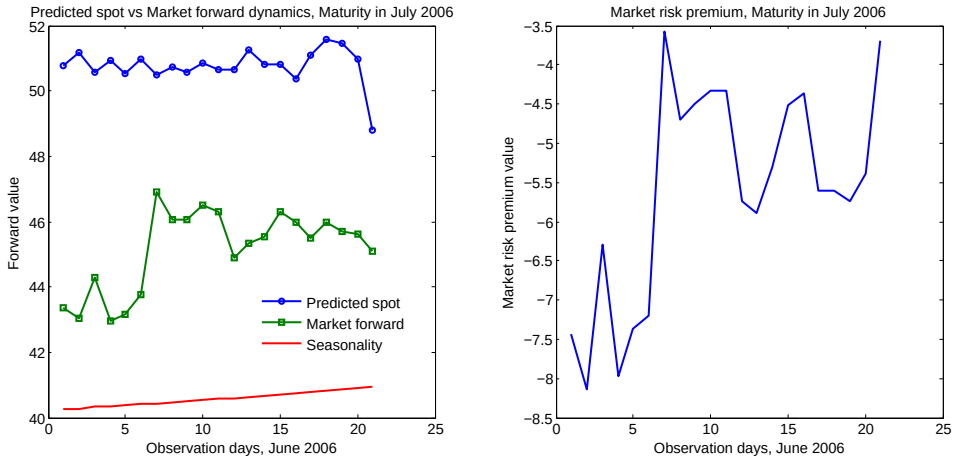


(c) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2005.

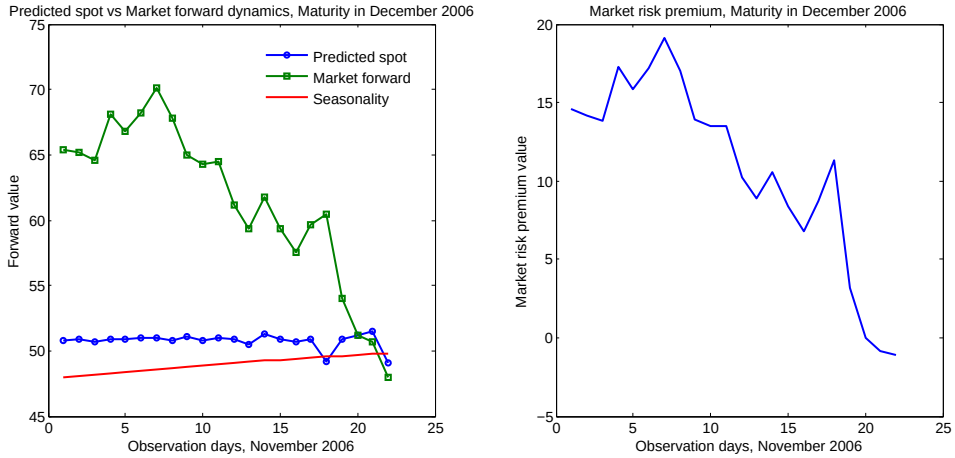


(d) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2005.

Figure A.-1: The predicted spot, observed forward dynamics and market risk premium for the threshold model.



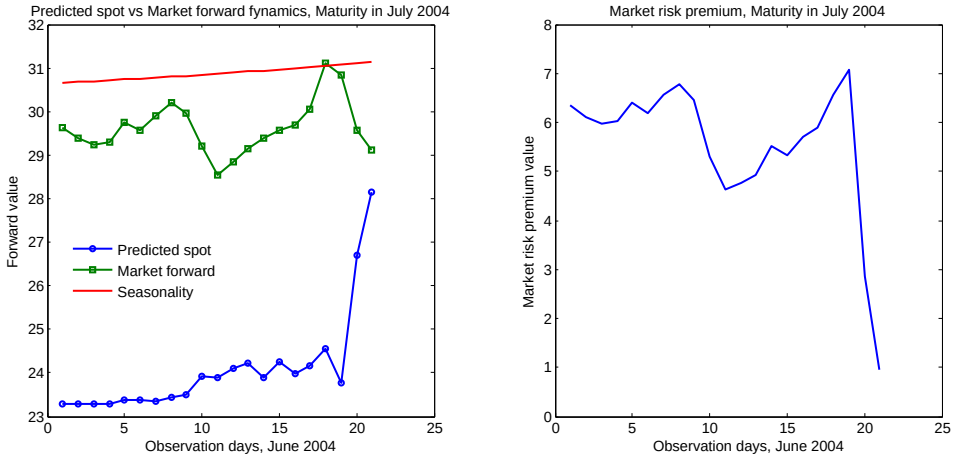
(e) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2006.



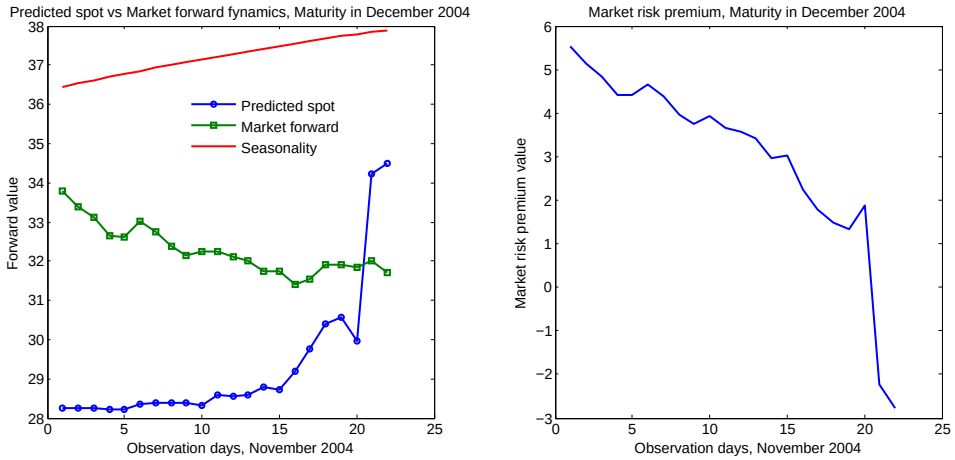
(f) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2006.

Figure A.-2: The predicted spot, observed forward dynamics and market risk premium for the threshold model.

A.3.3 Factor model and its forward modelling

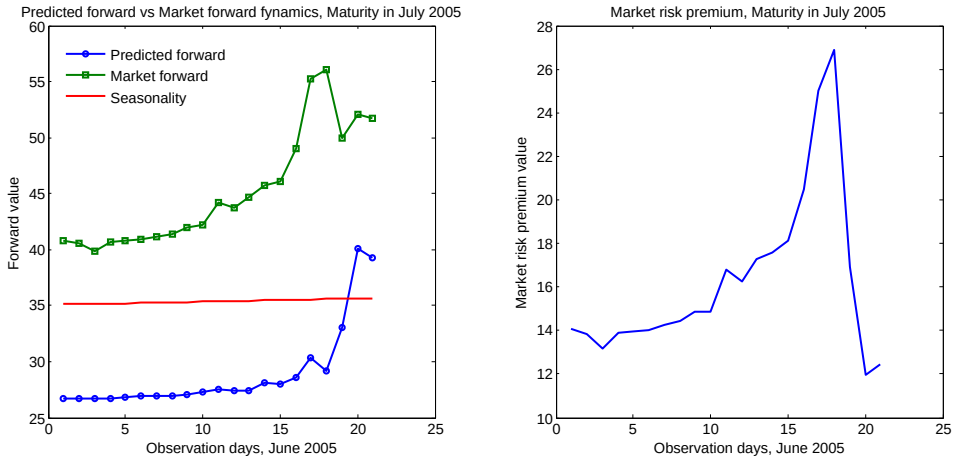


(a) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2004.

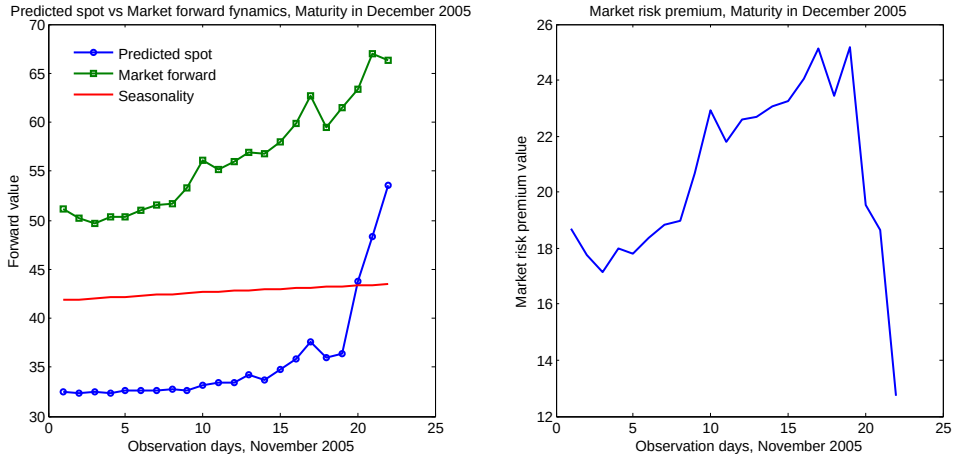


(b) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2004.

Figure A.-1: The predicted spot, observed forward dynamics and market risk premium for the factor model.

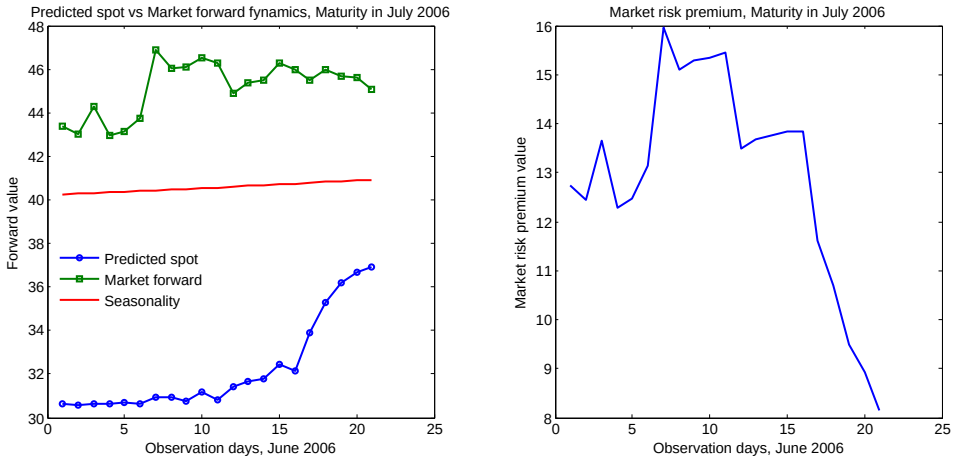


(c) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2005.

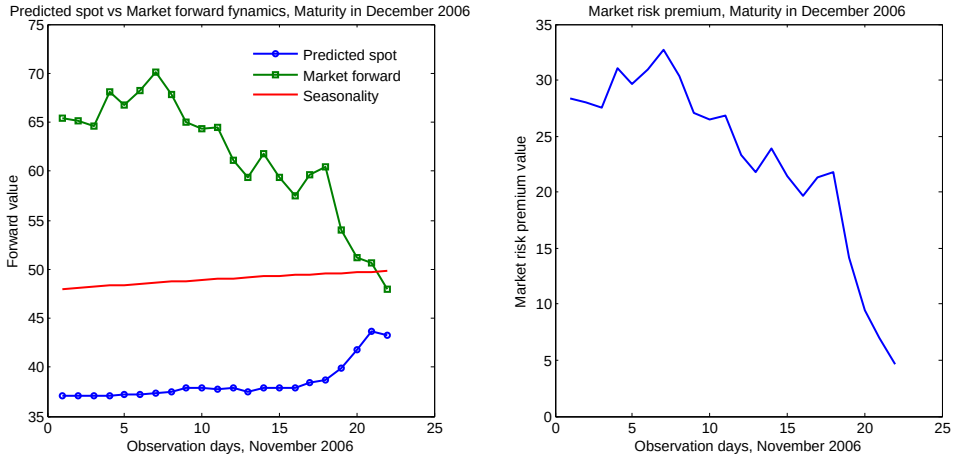


(d) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2005.

Figure A.-2: The predicted spot, observed forward dynamics and market risk premium for the factor model.



(e) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in July 2005.



(f) The predicted spot vs observed forward dynamics and market risk premium for forwards with maturity in December 2006.

Figure A.-3: The predicted spot, observed forward dynamics and market risk premium for the factor model.

APPENDIX B

MODEL RISK

B.1 Fisher information matrices

We can also compute the Fisher information matrix for the emissions process:

$$I(\alpha^E, \sigma^E) = \begin{bmatrix} \frac{1}{(\sigma^E)^2} & 0 \\ 0 & \frac{1}{2(\sigma^E)^4} \end{bmatrix} = \begin{bmatrix} 4747.5962 & 0 \\ 0 & 14259.9278 \end{bmatrix}.$$

Fisher information matrix for correlated power and gas processes:

$$I(\alpha^G, \alpha^P, \sigma^G, \sigma^P, \rho) = \begin{bmatrix} 1.7068 & -0.2340 & 0 & 0 & 0 \\ -0.2340 & 11.3785 & 0 & 0 & 0 \\ 0 & 0 & 1819.22 & -1015.81 & 141.80 \\ 0 & 0 & -1015.81 & 27667.79 & -7645.39 \\ 0 & 0 & 141.80 & -7645.39 & 103051.77 \end{bmatrix}.$$

Fisher information matrix for spike size with Laplace distributed jumps (using the usual parametrization (μ, σ), as in [Kotz *et al.*, 2001, p. 65]):

$$I(\mu_s, \sigma_s) = \begin{bmatrix} \frac{1}{\sigma_s^2} & 0 \\ 0 & \frac{1}{\sigma_s^2} \end{bmatrix} = \begin{bmatrix} 2.6226 & 0 \\ 0 & 2.6226 \end{bmatrix}.$$

Fisher information matrix for spike size process with normal distributed jumps (using the usual parametrization (μ, σ^2)):

$$I(\mu_s, \sigma_s) = \begin{bmatrix} \frac{1}{\sigma_s^2} & 0 \\ 0 & \frac{1}{2\sigma_s^4} \end{bmatrix} = \begin{bmatrix} 1.7074 & 0 \\ 0 & 1.4576 \end{bmatrix}.$$

B.2 Absolute and relative bid-ask prices for other sources of risk

- Gas signal parameters. We assess the parameter risk in the gas signal for both cases where the jump sizes follow either a Normal or a Laplace distribution. This effect is depicted in Figures B.1 and B.2. One observes that the parameter-risk implied bid-ask spread w.r.t. the gas price process is not too large, mainly due to the large sample size.

- Gas and power base signal parameters. We assess the parameter risk in the joint gas and power base signals for both cases where the jump sizes follow either a Normal or a Laplace distribution. This effect is depicted in Figures B.3 and B.4. One can see that the parameter-risk implied bid-ask spread w.r.t. the gas and power base price process is not too large, mainly due to the large sample size.

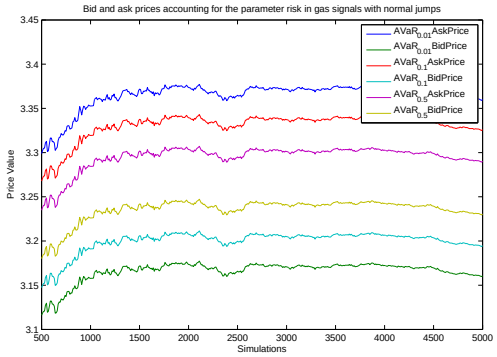
- Everything except of the jump size. We assess the parameter risk in the jump size distributions for both cases where the jump sizes follow either a Normal or a Laplace distribution. This effect is depicted in Figures B.5 and B.6.

  One can see that the parameter-risk implied bid-ask spread w.r.t. the jump size distribution is crucial – the correct determination of the jump size distribution is the major driver for the power plant price. Furthermore, convergence is much slower in the Laplace case and the results become much less stable.

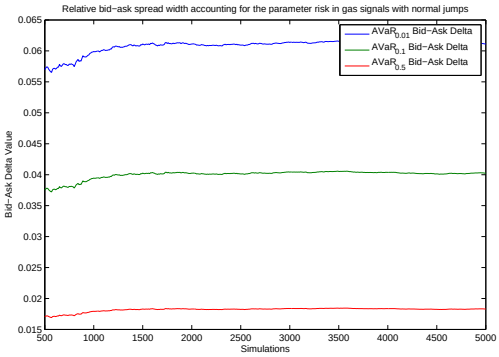
  By doing this risk assessment, we check the hypothesis that the major source of risk comes from the jump distribution.

B.2.1 Sensitivity impact

The choice of numerical approximation of the sensitivity is of minor importance for all the cases considered in Figures B.7, B.8, B.9, B.10, B.11 and B.12.

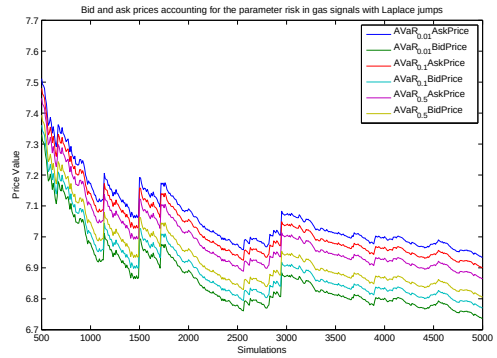


(a) Risk-captured bid and ask prices for different α levels when changing the **gas signal** parameter values with a shift size of 1%, normal jumps.

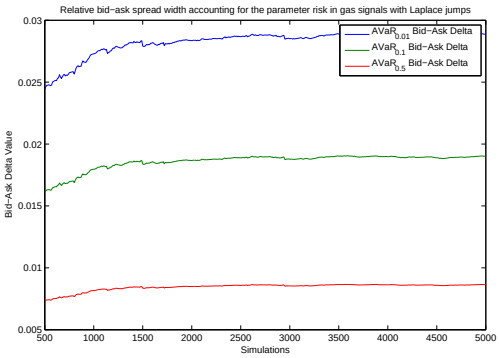


(b) Relative width of the risk-captured bid-ask spread for different α levels when changing the **gas signal** parameter values with a shift size of 1%, normal jumps.

Figure B.1: Parameter-risk implied bid-ask spread w.r.t. the gas price process.

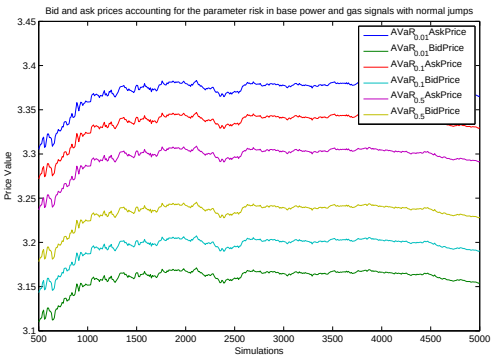


(a) Risk-captured bid and ask prices for different α levels when changing the **gas signal** parameter values with a shift size of 1%, Laplace jumps.

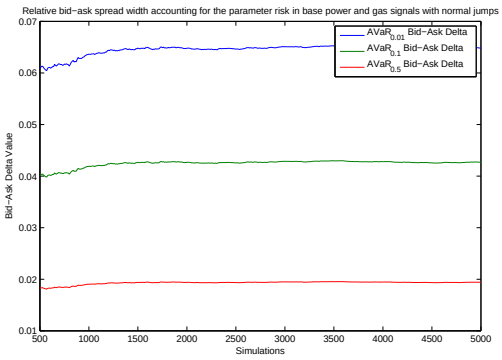


(b) Relative width of the risk-captured bid-ask spread for different α values when changing the **gas signal** parameter values with a shift size of 1%, Laplace jumps.

Figure B.2: Parameter-risk implied bid-ask spread w.r.t. the gas price process.

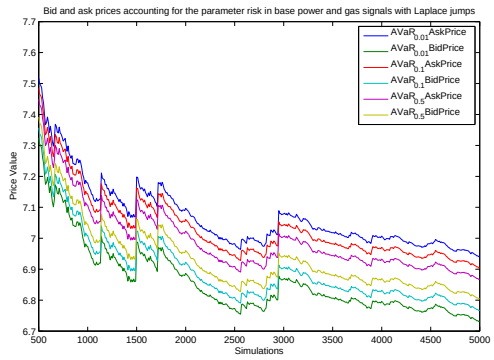


(a) Risk-captured bid and ask prices for different α levels when changing the **gas and power base** signal parameter values with a shift size of 1%, normal jumps.

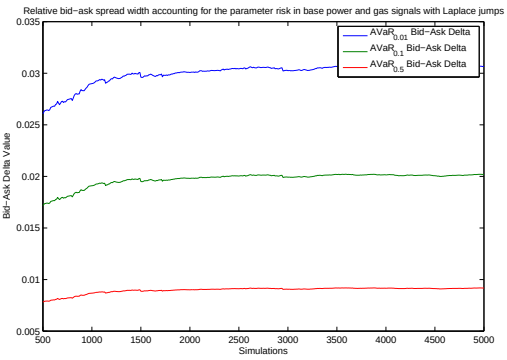


(b) Relative width of the risk-captured bid-ask spread for different α levels when changing the **gas and power base** signal parameter values with a shift size of 1%, normal jumps.

Figure B.3: Parameter-risk implied bid-ask spread w.r.t. the gas and power base price processes.

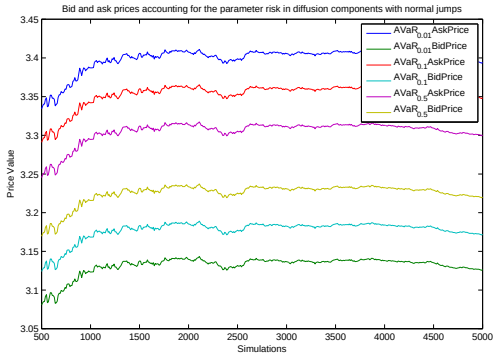


(a) Risk-captured bid and ask prices for different α levels when changing the **gas and power base** signal parameter values with a shift size of 1%, Laplace jumps.

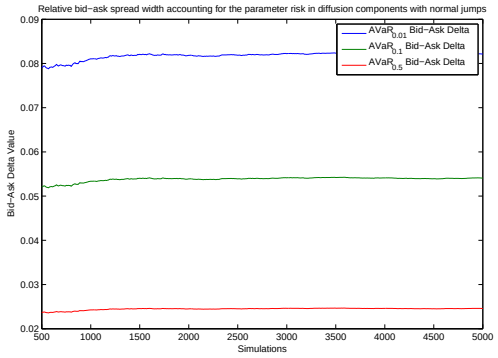


(b) Relative width of the risk-captured bid-ask spread for different α values when changing the **gas and power base** signal parameter values with a shift size of 1%, Laplace jumps.

Figure B.4: Parameter-risk implied bid-ask spread w.r.t. the gas and power base price processes.

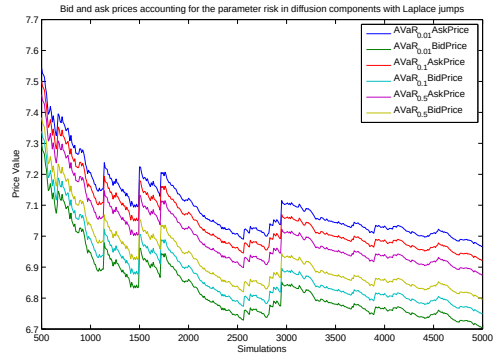


(a) Risk-captured bid and ask prices for different α levels when changing **all the parameters** values except of jump size with a shift size of 1%, normal jumps.

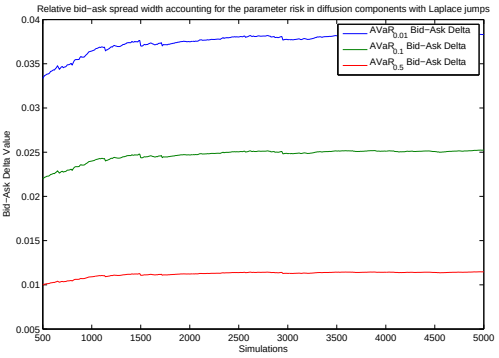


(b) Relative width of the risk-captured bid-ask spread for different α levels when changing **all the parameters** values except of jump size with a shift size of 1%, normal jumps.

Figure B.5: Parameter-risk implied bid-ask spread w.r.t. all the parameters, except of the jump size.

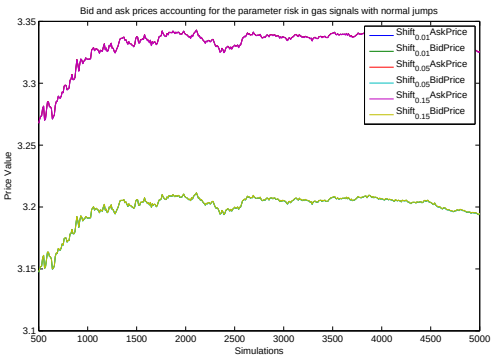


(a) Risk-captured bid and ask prices for different α levels when changing **all the parameters** values except of jump size with a shift size of 1%, Laplace jumps.

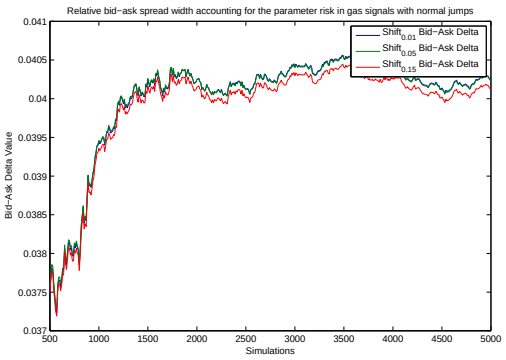


(b) Relative width of the risk-captured bid-ask spread for different α values when changing **all the parameters** values except of jump size with a shift size of 1%, Laplace jumps.

Figure B.6: Parameter-risk implied bid-ask spread w.r.t. all the parameters, except of the jump size.

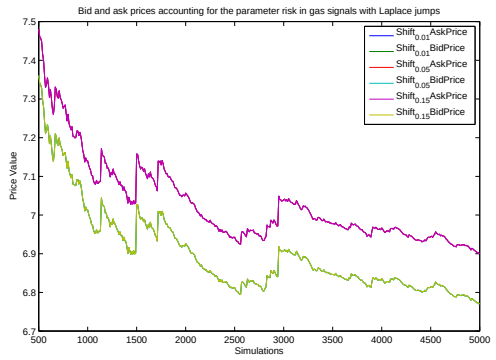


(a) Risk-captured bid and ask prices for different shift sizes when changing the **gas signal** parameter values with a significance level $\alpha = 10\%$, normal jumps.

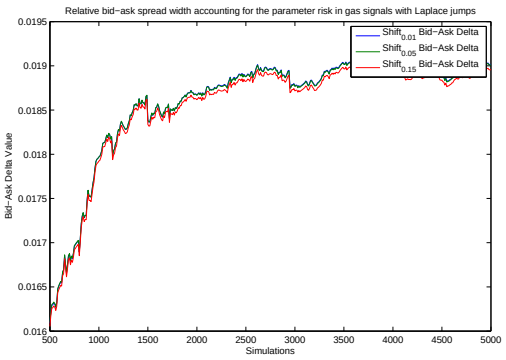


(b) Relative width of the risk-captured bid-ask spread for different shift sizes when changing the **gas signal** parameter values with a significance level $\alpha = 10\%$, normal jumps.

Figure B.7: Parameter-risk implied bid-ask spread w.r.t. the sensitivity value (gas price process).

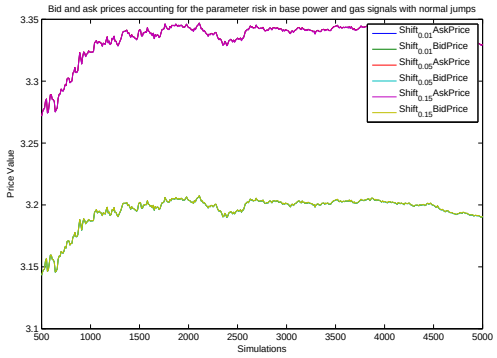


(a) Risk-captured bid and ask prices for different shift sizes when changing the **gas signal** parameter values with a significance level $\alpha = 10\%$, Laplace jumps.

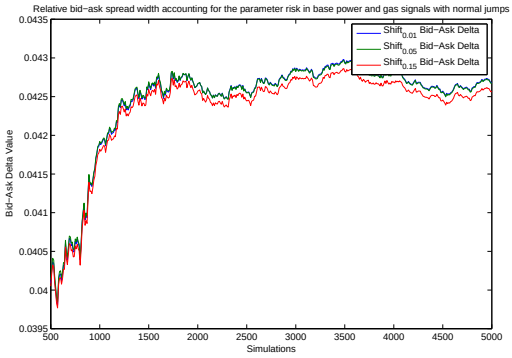


(b) Relative width of the risk-captured bid-ask spread for different shift sizes when changing the **gas signal** parameter values with a significance level $\alpha = 10\%$, Laplace jumps.

Figure B.8: Parameter-risk implied bid-ask spread w.r.t. the sensitivity value (gas price process).

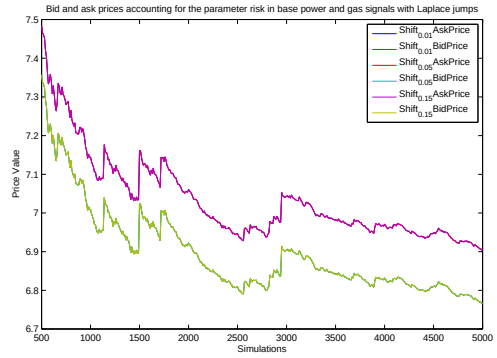


(a) Risk-captured bid and ask prices for different shift sizes when changing the **gas and power base** signal parameter values with a significance level $\alpha = 10\%$, normal jumps.

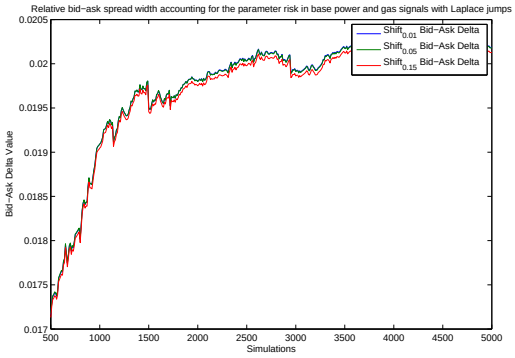


(b) Relative width of the risk-captured bid-ask spread for different shift sizes when changing the **gas and power base** signal parameter values with a significance level $\alpha = 10\%$, normal jumps.

Figure B.9: Parameter-risk implied bid-ask spread w.r.t. the sensitivity value (gas and power base price processes).

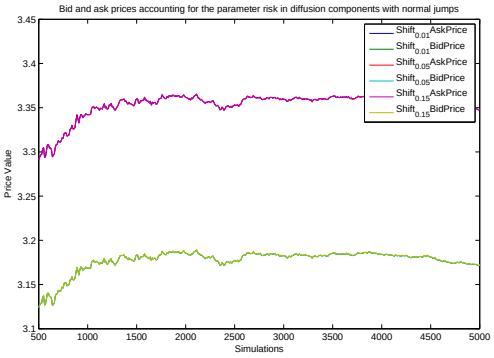


(a) Risk-captured bid and ask prices for different shift sizes when changing the **gas and power base** signal parameter values with a significance level $\alpha = 10\%$, Laplace jumps.

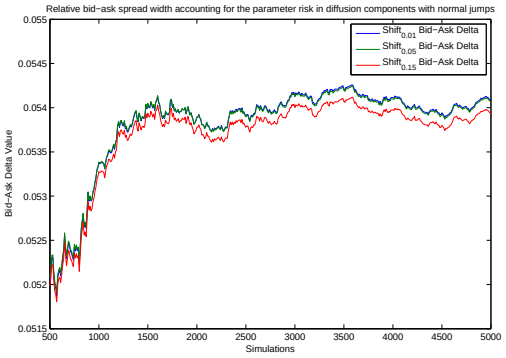


(b) Relative width of the risk-captured bid-ask spread for different shift sizes when changing the **gas and power base** signal parameter values with a significance level $\alpha = 10\%$, Laplace jumps.

Figure B.10: Parameter-risk implied bid-ask spread w.r.t. the sensitivity value (gas and power base price processes).

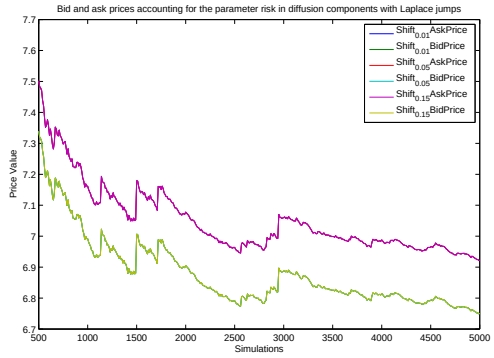


(a) Risk-captured bid and ask prices for different shift sizes when changing **all the parameters** values except of jump size with a significance level $\alpha = 10\%$, normal jumps.

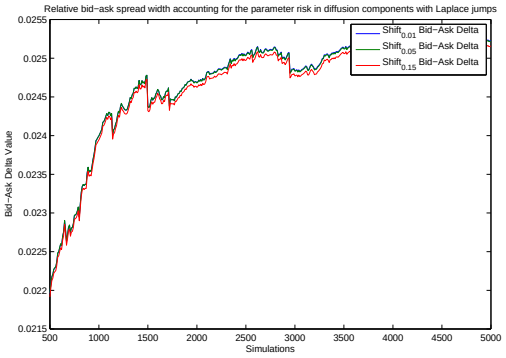


(b) Relative width of the risk-captured bid-ask spread for different shift sizes when **all the parameters** values except of jump size with a significance level $\alpha = 10\%$, normal jumps.

Figure B.11: Parameter-risk implied bid-ask spread w.r.t. the sensitivity value (all parameters, except of the jump size).



(a) Risk-captured bid and ask prices for different shift sizes when changing **all the parameters** values except of jump size with a significance level $\alpha = 10\%$, Laplace jumps.



(b) Relative width of the risk-captured bid-ask spread for different shift sizes when changing **all the parameters** values except of jump size with a significance level $\alpha = 10\%$, Laplace jumps.

Figure B.12: Parameter-risk implied bid-ask spread w.r.t. the sensitivity value (all parameters, except of the jump size).

REFERENCES

ABRAMOWITZ, M. & STEGUN, I. (1970). Handbook of mathematical functions. 94

ACERBI, C. & TASCHE, D. (2002). On the coherence of expected shortfall. *Journal of Banking and Finance*, **26**, 1487–1503. 123, 131

AHN, H., DANILOVA, A. & SWINDLE, G. (2002). Storing arb. *Wilmott*, **1**, 78–82. 4, 89

ALBANESE, C., LO, H. & TOMPAIDIS, S. (2008). A numerical method for pricing electricity derivatives based on continuous time lattices. Unpublished manuscript, Imperial College London and University of Texas at Austin. 4, 8

ARTZNER, P., DELBAEN, F., EBER, J.M. & HEATH, D. (1999). Coherent measures of risk. *Mathematical Finance*, **9**, 203–228. 122

AVELLANEDA, M., LEVY, A. & PARAS, A. (1995). Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, **2**, 73–88. 121

BANNÖR, K. & SCHERER, M. (2013). Capturing parameter risk with convex risk measures. *To appear in European Actuarial Journal*. 5, 122, 123, 124

BANNÖR, K., KIESEL, R., NAZAROVA, A. & SCHERER, M. (2014). Model risk and power plant valuation. *Submitted to Energy Economics*. 5

BARLOW, M. (2002). A diffusion model for electricity prices. *Mathematical Finance*, **12**, 287–298. 3

BARNDORFF-NIELSEN, O. & SHEPHARD, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, **63**, 167–241. 14, 28, 130

BELL, W.W. (2004). *Special functions for scientists and engineers*. Courier Dover Publications. 94

BENTH, F. & KOEKEBAKKER, S. (2008). Stochastic modeling of financial electricity contracts. *Energy Economics*, **30**, 1116–1157. 125

BENTH, F. & KUFAKUNESU, R. (2009). Pricing of exotic energy derivatives based on arithmetic spot models. *International Journal of Theoretical and Applied Finance*, **12**, 491–506. 8, 125

BENTH, F., KALLSEN, J. & MEYER-BRANDIS, T. (2007). A non-Gaussian Ornstein-Uhlenbeck process for electricity spot price modeling and derivatives pricing. *Applied Mathematical Finance*, **14**, 153–169. 3, 7, 8, 11, 14, 15, 28, 38, 47, 130

BENTH, F., BENTH, J. & KOEKEBAKKER, S. (2008a). *Stochastic modelling of electricity and related markets*. World Scientific Pub Co Inc. 36, 37, 42

BENTH, F., CARTEA, Á. & KIESEL, R. (2008b). Pricing forward contracts in power markets by the certainty equivalence principle: explaining the sign of the market risk premium. *Journal of Banking & Finance*, **32**, 2006–2021. 40

BENTH, F., LANGE, N. & MYKLEBUST, T.A. (2013). Pricing and hedging quanto options in energy markets. *Available at SSRN 2133935*. 100

BENTH, F.E., KIESEL, R. & NAZAROVA, A. (2012). A critical empirical study of three electricity spot price models. *Energy Economics*, **34**, 1589–1616. 5, 51, 67, 122, 125, 130

BESSEMBINDER, H. & LEMMON, M.L. (2002). Equilibrium pricing and optimal hedging in electricity forward markets. *The Journal of Finance*, **57**, 1347–1382. 2

BIBBY, B., JACOBSEN, M. & SØRENSEN, M. (2010). Estimating functions for discretely sampled diffusion-type models. *In Aït-Sahalia, Y. and Hansen, L.P. (eds.): Handbook of financial econometrics, North Holland, Oxford*, **1**, 203–268. 28, 152

BREIMAN, L. (1968). *Probability. Addison*. 73

BRIGO, D., DALESSANDRO, A., NEUGEBAUER, M. & TRIKI, F. (2007). A stochastic processes toolkit for risk management. *The Journal of Risk Management for Financial Institutions*. 130

BUNN, D., ANDRESEN, A., CHEN, D. & WESTGAARD, S. (2013). Analysis and forecasting of electricity price risks with quantile factor models. *forthcoming in Operational Research*. 131

BURGER, M., GRAEBER, B. & SCHINDLMAYER, G. (2008). *Managing energy risk: An integrated view on power and other energy markets*, vol. 425. Wiley. 122, 125

CARMONA, R. & LUDKOVSKI, M. (2010). Valuation of energy storage: An optimal switching approach. *Quantitative Finance*, **10**, 359–374. 4, 89

CARTEA, A. & FIGUEROA, M.G. (2005). Pricing in electricity markets: a mean reverting jump diffusion model with seasonality. *Applied Mathematical Finance*, **12**, 313–335. 3, 4, 7, 9, 15, 17, 18, 37, 50, 53, 126, 130, 131

CHEN, Z. & FORSYTH, P.A. (2007). A semi-lagrangian approach for natural gas storage valuation and optimal operation. *SIAM Journal on Scientific Computing*, **30**, 339–368. 4, 89

CLEWLOW, L. & STRICKLAND, C. (2000). *Energy derivatives: Pricing and risk management*. Lacima, London. 24

CONNOLLY, D., LUND, H., FINN, P., MATHIESEN, B.V. & LEAHY, M. (2011). Practical operation strategies for pumped hydroelectric energy storage (phes) utilising electricity price arbitrage. *Energy Policy*, **39**, 4189–4196. 107

CONT, R. (2006). Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical Finance*, **16**, 519–547. 5, 121, 122

CONT, R. & TANKOV, P. (2004). *Financial modelling with jump processes*. Chapman & Hall. 57, 58, 59, 64, 149, 150

CONT, R. & VOLTCHKOVA, E. (2005). A finite difference scheme for option pricing in jump diffusion and exponential lévy models. *SIAM Journal on Numerical Analysis*, **43**, 1596–1626. 61, 64

COOPER, R., HOARE, M.R. & RAHMAN, M. (1977). Stochastic processes and special functions: on the probabilistic origin of some positive kernels associated with classical orthogonal polynomials. *Journal of Mathematical Analysis and Applications*, **61**, 262–291. 79

CULOT, M.E.A. (2013). Practical stochastic modelling of electricity prices. *The Journal of Energy Markets*, **6**, 3–31. 122, 125

DE JONG, C. (2006). The nature of power spikes: a regime-switching approach. *Studies in Nonlinear Dynamics and Econometrics*, **10**. 122

DREES, H., DE HAAN, L. & RESNICK, S. (2000). How to make a Hill plot. *The Annals of Statistics*, **28**, 254–274. 30

DUFFY, D. (2001). *Green's functions with applications*. CRC Press. 81

EYDELAND, A. & GEMAN, H. (1999). Fundamentals of electricity derivatives. *Energy Modelling and the Management of Uncertainty*, 35–43. 3

EYDELAND, A. & WOLYNIEC, K. (2003). *Energy and power risk management: New developments in modeling, pricing, and hedging*. John Wiley & Sons Inc. 1, 2, 7, 17, 122

FEDERAL ENERGY REGULATORY COMMISSION AND OTHERS (2004). Current state of and issues concerning underground natural gas storage. *Energy, Washington DC*. 88

FLETEN, S.E. (2013). Is there power in derivatives? In *The Third Energy Christmas workshop*. 92

FÖLLMER, H. & SCHIED, A. (2002). Convex measures of risk and trading constraints. *Finance and Stochastics*, **6**, 429–447. 123

FRITTELLI, M. & SCANDOLO, G. (2006). Risk measures and capital requirements for processes. *Wiley: Mathematical Finance*, **16**, 589–612. 123

FUSAI, G. & RONCORONI, A. (2008). *Implementing models in quantitative finance: methods and cases*. Springer Verlag. 39, 57

GEMAN, H. & KOUROUVAKALIS, S. (2008). A Lattice-Based Method for Pricing Electricity Derivatives Under the Threshold Model. *Applied Mathematical Finance*, **15**, 531–567. 8

GEMAN, H. & RONCORONI, A. (2006). Understanding the Fine Structure of Electricity Prices. *The Journal of Business*, **79**, 1225–1261. 3, 7, 9, 10, 11, 12, 15, 19, 21, 22, 29, 36, 39, 49, 50, 130

GENON-CATALOT, V. & JACOD, J. (1993). On the estimation of the diffusion coefficient for multi-dimensional diffusion processes. *Annales de l'I. H. Poincaré Probabilités et statistiques*, **29**, 119–151. 23

GIHMAN, I. & SKOROKHOD, A. (1972). *Stochastic Differential Equations*. Springer-Verlag, Berlin. 73, 76

GLASSERMAN, P. (2004). *Monte Carlo methods in financial engineering*. Springer Verlag. 39

GLASSERMAN, P. & XU, X. (2012). Robust Risk Measurement and Model Risk. 121

GUPTA, A., REISINGER, C. & WHITLEY, A. (2010). Model uncertainty and its impact on derivative pricing. *in Rethinking Risk Measurement and Reporting*. 122

HAMBLY, B., HOWISON, S. & KLUGE, T. (2009). Modelling spikes and pricing swing options in electricity markets. *Quantitative Finance*, **9**, 937–949. 126

INTERNATIONAL ENERGY AGENCY (2013). World energy outlook 2013. Tech. rep., International Energy Agency. 1

ITO, K. & MCKEAN, H. (1974). Diffusion processes and their sample paths, 2nd printing. 75

JANCZURA, J.E. (2013). Identifying spikes and seasonal components in electricity spot price data: A guide to robust modeling. *Energy Economics*, **38**, 96–110. 130

KALLSEN, J. & SHIRYAEV, A.N. (2002). The cumulant process and esscher's change of measure. *Finance and Stochastics*, **6**, 397–428. 151

KAMINSKI, V. (2005). *Energy modelling: Advances in the management of uncertainty*. Risk Books. 2

KARATZAS, I. & SHREVE, S.E. (1991). *Brownian motion and stochastic calculus*, vol. 113. Springer. 72

KARLIN, S. & TAYLOR, H. (1981). *A second course in stochastic processes*. Academic press. 71, 73

KAUPPI, O. & LISKI, M. (2008). An empirical model of imperfect dynamic competition and application to hydroelectricity storage. *MIT Center for Energy and Environmental Policy Research*. 88

KJAER, M. & RONN, E.I. (2008). Valuation of natural gas storage facility. *Journal of Energy Markets*, **1**, 3–22. 4, 89

KNIGHT, F. (1921). *Risk, uncertainty, and profit*. Hart, Schaffner & Marx. 121, 122

KNITTEL, C. & ROBERTS, M. (2005). An empirical examination of restructured electricity prices. *Energy Economics*, **27**, 791–817. 130

KOTZ, S., KOZUBOWSKI, T. & PODGÓRSKI, K. (2001). *The Laplace distribution and generalizations: A revisit with applications to communications, economics, engineering, and finance*. Birkhäuser. 163

KOZUBOWSKI, T. & PODGORSKI, K. (2003). A log-laplace growth rate model. *The Mathematical Scientist*, **28**, 49–60. 53

KUSUOKA, S. (2001). On law invariant coherent risk measures. *Advances in Mathematical Economics*, **3**, 83–95. 123

LINDSTRÖM, E. (2010). Implication of parameter uncertainty on option prices. *Advances in Decision Sciences*, 15 pages. 122

LUCIA, J. & SCHWARTZ, E. (2002). Electricity prices and power derivatives: Evidence from the Nordic Power Exchange. *Review of Derivatives Research*, **5**, 5–50. 3, 7, 126

MEYER-BRANDIS, T. & TANKOV, P. (2008). Multi-Factor Jump-Diffusion Models Of Electricity Prices. *International Journal of Theoretical and Applied Finance (IJTAF)*, **11**, 503–528. 8, 12, 27, 29, 30, 32, 122, 130, 131

NAZAROVA, A. (2008). Lévy-Based Electricity Spot Price Modelling. *Master thesis, University of Ulm*. 27

NOWOTARSKI, J.W.R., J.; TOMCZYK (2013). Robust estimation and forecasting of the long-term seasonal component of electricity spot prices. *Energy Economics*, **39**. 130

OYEBANJI, G. (2007). Implementation of the Geman and Roncoroni Threshold Model. *Master thesis, University of Zürich*. 35

PIRRONG, C. & JERMAKYAN, M. (1999). Valuing power and weather derivatives on a mesh using finite difference methods. *Energy Modelling and the Management of Uncertainty*, 59–69. 3

REBONATO, R. (2010). *Coherent stress testing: A Bayesian approach to the analysis of financial stress*. John Wiley and sons. 121

REVUZ, D. & YOR, M. (1999). *Continuous martingales and Brownian motion*, vol. 293. Springer. 74

RONCORONI, A. (2002). Essays in quantitative finance: Modelling and calibration in interest rate and electricity markets. *PhD diss., University Paris Dauphine*. 7, 10, 23, 50

RUEHLICKE, R.E.A. (2013). Robust Risk Measurement. 121

SATO, K. (1999). *Lévy processes and infinitely divisible distributions*. Cambridge Univ Pr. 150, 151

SCHOUTENS, W., SIMONS, E. & TISTAERT, J. (2004). A perfect calibration! Now what? *Wilmott Magazine*, **3**, 66–78. 122

SCHWARTZ, E. (1997). The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance*, **52**, 923–973. 7

SCHWARTZ, E. & SMITH, J.E. (2000). Short-term variations and long-term dynamics in commodity prices. *Management Science*, **46**, 893–911. 3

SHAO, J. (1999). *Mathematical Statistics*. Springer Verlag. 133

SHREVE, S.E. (2004). *Stochastic calculus for finance II: Continuous-Time Modelsl*, vol. 2. Springer. 64

SØRENSEN, M. (2000). Prediction-based estimating functions. *Econometrics Journal*, **3**, 123–147. 8, 28, 130, 152

STAKGOLD, I. (1979). *Green's Functions and Boundary Value Problems, 1979*. A Wiley-Interscience Series of Texts. 80

SUPATGIAT, C., ZHANG, R.Q. & BIRGE, J.R. (2001). Equilibrium values in a competitive power exchange market. *Computational Economics*, **17**, 93–121. 2

SZEGO, G. (1939). Orthogonal polynomials, 23. In *American Mathematical Society Colloquium Publi*. 79

THOMPSON, M., DAVISON, M. & RASMUSSEN, H. (2009). Natural gas storage valuation and optimization: A real options application. *Naval Research Logistics (NRL)*, **56**, 226–238. 4, 89

WHITTAKER, E.T. & WATSON, G.N. (1996). *A course of modern analysis*. Cambridge university press. 94