

# Estimating forward price curve in energy markets with MCMC method

by

Jeroen De Maeseneire

## **Advisor**

Professor Christiane Lemieux

A research paper  
presented to the University of Waterloo  
in partial fulfillment of the  
requirements for the degree of  
Master of Mathematics  
in  
Computational Mathematics

Waterloo, Ontario, Canada, 2010

I hereby declare that I am the sole author of this paper. This is a true copy of the paper, including any required final revisions, as accepted by my examiners.

I understand that my paper may be made electronically available to the public.

## Abstract

In this essay, modeling of natural gas price forward curves is discussed. A two-factor, regime switching model is proposed to model forward price dynamics. The model is a Markov Switching model: the parameters, found in the respective volatility functions take different values for different states of the world. Further, the state transitions are driven by a finite state Markov Chain. A Markov Chain Monte Carlo (MCMC) technique is used for estimation of the parameter set and the transition probability matrix. A study is performed for a number of 2 states using data from NBP. Compared to a previous study it is found that, while the volatility and attenuation parameters are similar, the Markov Chain stays in regime 2 more often. That is, over the span of the data period, more time is spent in a high-volatility regime. Although, for the better part of the time horizon the Markov Chain still settles in regime 1. A higher volatile environment for natural gas could be explained by the global financial crisis which unfolded during the period of observation.

# Contents

<b>List of Tables</b>	<b>vi</b>
<b>List of Figures</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Stylized facts of energy forward price dynamics . . . . .	2
1.2 Literature review . . . . .	9
<b>2 Forward price curve model</b>	<b>13</b>
2.1 Multi-factor single-commodity model . . . . .	13
2.2 Spot - forward price relationship . . . . .	18
2.3 Seasonality adjustment . . . . .	18
2.4 Multi-factor multi-commodity model . . . . .	19
2.5 Two-factor regime switching forward price curve model . . . . .	20
<b>3 Estimation</b>	<b>22</b>
3.1 Data transformation . . . . .	22
3.2 Bayesian estimation for a Markov switching model . . . . .	23
3.3 Prior distributions . . . . .	24
3.4 Gibbs sampler . . . . .	25
3.4.1 Full conditional posterior distributions . . . . .	25
3.4.2 Sample the transition matrix . . . . .	26
3.4.3 Sample the model parameters . . . . .	27
3.4.4 Markov Chain state simulation . . . . .	29
<b>4 Implementation</b>	<b>32</b>
4.1 Data . . . . .	32
4.2 Choosing the prior . . . . .	32
4.3 Running the Gibbs sampler . . . . .	33
4.3.1 One-factor model . . . . .	34
4.3.2 Two-factor model . . . . .	35

4.4	Numerical Results . . . . .	36
4.4.1	One-factor model . . . . .	36
4.4.2	Two-factor model . . . . .	38
4.4.3	Historical NBP forward data . . . . .	45
<b>5</b>	<b>Conclusion</b>	<b>47</b>
<b>A</b>	<b>nbp_1fac.m</b>	<b>48</b>
<b>B</b>	<b>mcmc.m</b>	<b>50</b>
<b>C</b>	<b>nbp_2fac.m</b>	<b>54</b>
<b>D</b>	<b>mcmc.m</b>	<b>56</b>
<b>E</b>	<b>mcmc.m</b>	<b>60</b>
	<b>Bibliography</b>	<b>65</b>

# List of Tables

# List of Figures

1.1	Forward price curves from Oct-2007 to Mar-09, $F(t, T)$ . . . . .	5
1.2	Log differences of forward prices $\mathbf{y}_T$ . . . . .	6
1.3	Different forward price curves, $F(t, T)$ . . . . .	8
2.1	Annualized covariance surface for NBP data $F(t, T)$ . . . . .	16
2.2	First three volatility functions for NBP data $F(t, T)$ . . . . .	17
4.1	Generated forward price dynamics, $\mathbf{y}_T$ . . . . .	34
4.2	Path of parameter estimates for $\sigma_{1j}$ and $\alpha_{1j}$ , ( $j = 1, 2$ ) . . . . .	37
4.3	Generated forward price dynamics, $\mathbf{y}_T$ . . . . .	38
4.4	Path of parameter estimates for $\sigma_{ij}$ , ( $i, j = 1, 2$ ) . . . . .	40
4.5	Path of parameter estimates for $\alpha_{ij}$ , ( $i, j = 1, 2$ ) . . . . .	41
4.6	Path of parameter estimates for $\sigma_{ij}$ , ( $i, j = 1, 2$ ) . . . . .	43
4.7	Path of parameter estimates for $\alpha_{ij}$ , ( $i, j = 1, 2$ ) . . . . .	44



# Chapter 1

## Introduction

This paper considers the problem of modeling and estimating the forward price curve in natural gas markets. A common modeling framework is provided that is consistent not only with the forward curve prices observed in the market but also the volatilities and correlations of forward price returns. The model discussed and implemented is a two-factor regime switching model suggested by Chiarella *et al.* [4] where the parameters of the volatility functions are allowed to take different values in different states of the world. The respective states are then governed by a Markov Chain. The dynamics of the two volatility functions are thus driven by shifts in the Markov Chain.

The paper is built as follows. First, in Chapter 1, the properties of energy forward prices are discussed. Energy markets differ from non-physical, financial markets. It is then important to incorporate those differences in the model to better explain the forward price dynamics of energy assets. Also, a brief overview of theoretical and practical work by other authors is provided. Next, in Chapter 2 the general multi-factor multi-commodity model by Breslin *et al.* [2] is introduced after which a two-factor regime switching model for the natural gas forward price curve is described. The model was initially proposed by Chiarella *et al.* [4].

The parameter estimation procedure is then explained in Chapter 3. A Bayesian approach is taken by means of a Markov Chain Monte Carlo (MCMC) method provided by Frühwirth-Schnatter [14] and Hahn *et al.* [18]. The algorithms implemented are extensively discussed and code of the respective routines is supplied. In a following chapter, Chapter 4, a number of numerical examples and calibration results from simulated data are considered. Subsequently, the estimation procedure is applied to real market data. Here, the analysis by Chiarella *et al.* [4] is replicated with updated natural gas forward price data. Finally, in Chapter 5 conclusions are drawn and suggestions for further research are discussed.

## 1.1 Stylized facts of energy forward price dynamics

Energy markets, to which the market for natural gas belongs, differ from non-physical markets such as the markets of interest rates and equities. Most of these differences can be traced back to supply and demand fundamentals, dynamics of production, consumption and storage. All drivers which are not found in the traditional markets. A comprehensive overview of the various supply and demand drivers of energy markets and their impacts on the prices and volatilities of the respective markets can be found in Pilipovic [25]. Below, the main features of energy forward price dynamics are presented. These market characteristics should be captured to model efficiently the evolution of natural gas forward curves.

- **Seasonality**

One of the major fundamental price drivers in energy markets is weather. On the demand side, there are considerable seasonality effects of residential users. Natural gas is consumed mostly during the winter. Therefore, one expects natural gas prices to peak during winter and drop to lows in summer period. Further, the magnitude of these peaks and troughs, or the degree of seasonality, will vary according to the region. The heating demand in the state of Massachusetts is expected to surpass the heating demand for Florida. These seasonality effects are recurring and, hence, incorporated in observed forward prices (Pilipovic [25]).

- **Mean reversion**

Mean reversion is the process of a market returning to its equilibrium price level. This is witnessed in most financial and physical markets where the equilibrium level, called *mean* level, could be a return on equity, a commodity price or a historical interest rate. In financial markets the actual rate of mean reversion is rather weak and tied to economic cycles. On the other hand, in energy markets mean reversion is much stronger and related to various, different drivers. In the case of energy commodities mean-reversion depends on how fast the market supply side responds to certain events. Or, also, how fast events go away (Pilipovic [25]).

Events that produce supply-demand imbalances could be wars, hurricanes or other headline-grabbing events. In 2005, for example, the hurricanes Katrina and Rita damaged a number of natural gas processing facilities on the US Gulf Coast. The loss has delayed the recovery of natural gas production in the area. As a result, price spikes occurred in the spot and short-term forward markets, but longer-term contracts did not move dramatically. The

mean reversion, as revealed by the forward prices, thus depended on how fast the production side could restore the supply imbalance (Pilipovic [25]).

- **High volatility**

Production and storage are two supply drivers which are nonexistent in financial markets. Expectations of market production and costs in the long run will be expressed in the levels and yields of long-term forward prices. Likewise, overcapacity in natural gas markets will result in distressed prices and its impact will depend on how long the overcapacity is expected to last. This so-called *storage limitation* problem is causing day-to-day volatility which is high compared to financial markets. Especially spot and short forwards exhibit high levels of volatility. For example, when there is no natural gas to go around, prices will easily reach very elevated levels (Pilipovic [25]).

- **Decreasing term structure of forward curve volatility** Storage limitation is thus a major volatility driver. Although spot prices show extremely high volatility, the volatilities of forward prices decrease significantly with increasing maturities. It can be expected that in the long run supply and demand will be balanced and, hence, long-term forward prices reveal a rather stable equilibrium price level. As a result, the short- and long-term portions of the energy forward price curves tend to be driven by different market factors with usually very little or no relationship (Pilipovic [25]).

The latter two characteristics demonstrate that the issue of storage causes energy prices to display a *split personality*. Where short-term forward prices are the result of energy in storage, and long-term forward prices reflect behavior of future potential energy. So while energy markets share some properties with traditional financial markets, they have unique and challenging workings. And these should be taken into account when modeling natural gas forward dynamics (Pilipovic [25]).

The natural gas prices modeled are forward prices for the *National Balancing Point* (NBP), which is the British virtual trading location operated by TSO National Grid, covering all entry and exit points in mainland Britain. It is the most liquid gas trading point in Europe. Natural gas at NBP trades on a forward month, forward quarter, forward season or year forward basis and prices are expressed in pence (GBP) per therm. For parameter estimation of the chosen model, daily observations for the forward prices were obtained with maturities  $T$  ranging from July 2009 to September 2012. The observations span over a year with prices at times  $t$  from 23/10/2007 to 12/03/2009. This data is represented by  $F(t, T)$ . Further, in order to obtain the forward dynamics  $y_{t,T}$  for every maturity  $T$ , the log difference transform is applied to the original data  $F(t, T)$ . Throughout the essay,

the notation  $\mathbf{y}_T$  is used to represent the forward dynamics  $y_{t,T}$  for each maturity  $T$  along all points in time  $t$ .

Figures 1.1 and 1.2 below show the behavior of  $F(t, T)$  and  $y_T$  respectively.

Figure 1.1: Forward price curves from Oct-2007 to Mar-09,  $F(t, T)$

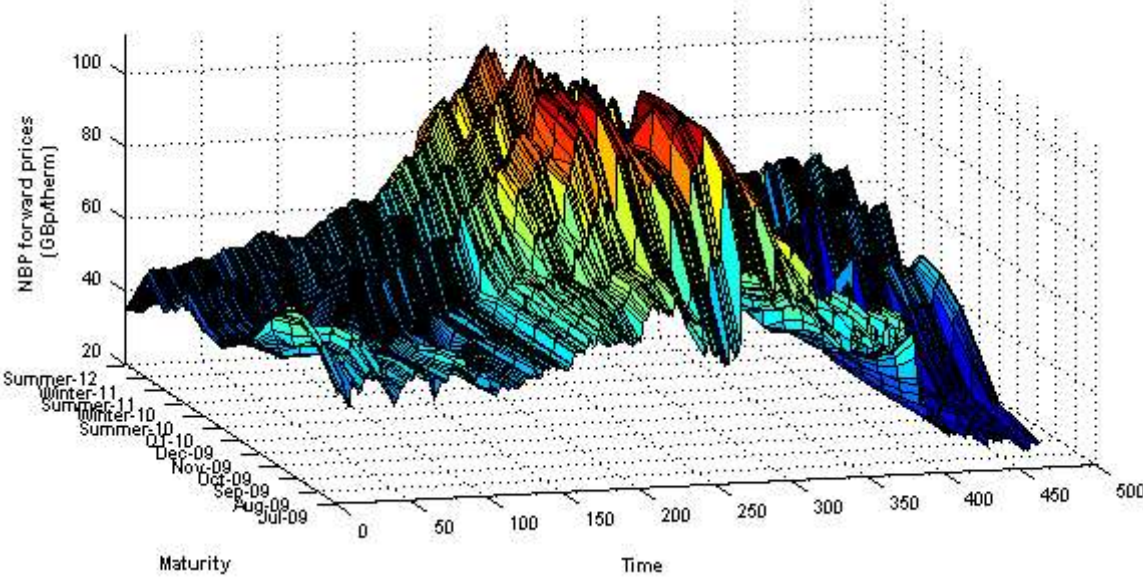
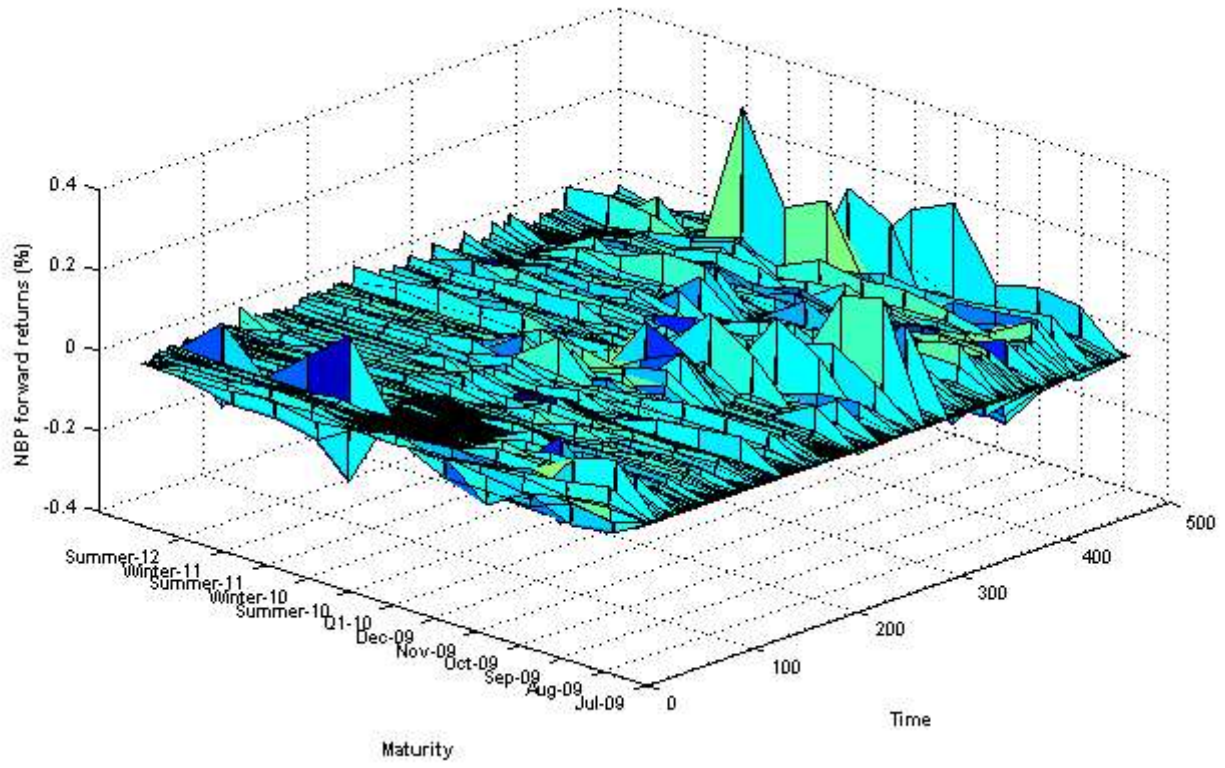


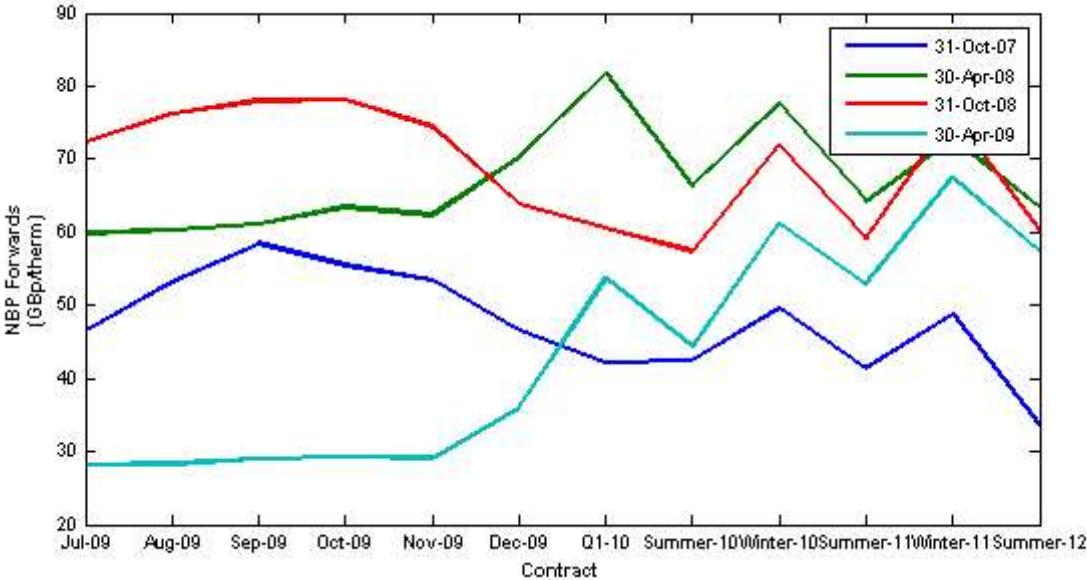
Figure 1.2: Log differences of forward prices  $y_T$



At every point in time  $t$  a seasonal pattern can be distinguished with higher prices for forward deliveries in Winter compared to deliveries in Summer. Also, the forward curves spike during a certain period, after which things turn back to normal and a more stable price pattern emerges. It is therefore easy to see that a regime switching model could be suited to approximate this behavior. Finally, natural gas for delivery further out in time is usually more expensive than delivery in the near time. The difference is explained by the *carry*, that is expenses incurred for storage and financing when the commodity is bought at the spot price and delivered at a later point in time. Such an increasing price curve is called *contango*. Although different shapes could occur such as a decreasing path for the forward curve which is called *backwardation*. Or any combination of contango and backwardation could occur along the forward price curve.

Figure 1.3 depicts some of the forward price curves observed.

Figure 1.3: Different forward price curves,  $F(t, T)$





## 1.2 Literature review

In order to model the gas forward curve, a regime switching model is proposed. It is considered a useful approach because many economic time series are known to exhibit dramatic breaks in their behavior. Therefore, when modeling, different behavior is assumed in one subsample or regime to another. Since the prevailing regime is not directly observable, a process that governs the transition has to be defined. For example, a process where the current state depends only on the state before, is called a *Markov* process. A regime switching model where the transitions are governed by a Markov process is called a *Markov Switching* model.

There has been an enormous amount of theoretical and practical work on regime switching models. In a seminal paper, Goldfeld and Quandt [17] propose a regression system of demand and supply functions thought to be switching between the two equations or regimes. An unobserved Markov process is assumed to govern the regime. This set-up is applied to a model for a housing market in disequilibrium. Two regimes are distinguished: if there is excess demand, the observed point lies on the supply function and if there is an excess demand, it lies on the demand function. Hamilton [19] and [20] extends the approach to asset prices.

In Hamilton [19] the approach by Goldfeld and Quandt [17] is used to describe exchange rate dynamics. Episodes of dollar appreciation relative to other currencies are followed by episodes of dollar depreciation. Therefore, in a regime switching model the dollar rate is allowed to switch, with a certain probability, between an appreciation and depreciation regime. Changes in asset prices are driven by specific identifiable events. Further, major exogenous economic events such as OPEC oil shocks affect financial time series. They also could be regarded as events of a certain duration in which the dynamics of financial time series might behave differently from what is seen outside these periods.

In Hamilton [20] a vector autoregression model is put forward subject to occasional discrete shifts, where a discrete-valued Markov process governs the shifts. Next, one has to find out when the shifts occurred and estimate parameters characterizing the different regimes and the probabilities for the transition between regimes. For maximum likelihood estimates of the respective parameters, an EM algorithm is applied. Applications of Markov regime switching processes to asset price volatility modeling can be found in contributions by Naik [24] and Di Masi *et al.* [11].

The volatility of risky assets is affected by random and discontinuous shifts over time. Those jumps are driven by various unsystematic and systematic events. Examples are mergers and acquisitions, shifts in corporate strategy and changes in economic policy. In order to hedge claims contingent on risky assets, a volatility model that captures the outcomes of such events is proposed. For example, Di Masi *et al.* [11] consider the problem of hedging an European call option for a diffusion model where drift and volatility are functions of a Markov jump process.

Fixed income models have been an important source for applications to energy price modeling. The energy forward price curve exhibit features similar to forward interest rate dynamics, especially the Heath-Jarrow-Morton (HJM) model by Heath *et al.* [22]. One could take two approaches for stochastic modeling of fixed income markets. First, one starts out with a stochastic model for the spot interest rate and derives bond prices from this short rate based on no-arbitrage principles. Alternatively, one could specify the complete yield curve dynamics directly which is suggested by the HJM approach. It means that forward rates are modeled directly. Since commodities trade both in spot and forward markets, these modeling approaches for fixed income markets could be applied to energy markets.

Markov Chain models for the instantaneous short-term interest rate were introduced by, amongst others, Hansen and Poulsen [21] and Elliott and Wilson [13]. In the first paper, only the drift rate of the short rate was modeled by a continuous Markov chain. Whereas Elliott and Wilson [13] extended this approach to the case when both the drift rate and the volatility of the short rate are driven by a common Markov Chain. The work of Valchev [29] builds on that of Elliott and Wilson [13], but a Markov Chain stochastic volatility HJM model is introduced. Also, the evolution of the whole term structure of volatilities is explored, not merely the volatility of the short rate, which is only one point of the volatility curve.

The Markov Chain specification allows for jump discontinuities and captures the various shapes of the term structure of volatilities as well. The model is an extension of the class of deterministic volatility HJM models to the wider class of HJM models with piecewise-deterministic volatility. The class of piecewise-deterministic processes, introduced by Davis [8], provides a suitable framework for modeling the dynamics of the volatility term structure. In between jumps the volatility follows an almost deterministic process, that is the volatility function is non-stochastic. The volatility modeled is stochastic only through its dependence

on the Markov Chain. A process in the class of piecewise-deterministic processes is the continuous-time homogeneous Markov Chain with a finite number of jump times. This process only approaches the actual jumps with jumps over a finite set of values but permits the use of stochastic calculus for continuous Markov chains Elliot *et al.* [12].

In energy markets modeling, most of existing literature focuses on developing realistic spot price markets. Based on a stochastic model for the time evolution of the spot price, forward dynamics are derived calling arbitrage theory. Therefore, various regime-switching models are calibrated to electricity spot prices (see, for example, Deng [10] and de Jong & Huisman [9]). The approach taken in this essay, based on the framework provided by Chiarella *et al.* [4], is in the spirit of Davis [8], Elliot *et al.* [12] and Valchev [29]. The result is a stochastic model that specifies directly the dynamics of the forward contracts traded in the natural gas markets. The class of piecewise deterministic processes is retained where the volatility functions are specified explicitly and the coefficients switch values dependent on the Markov Chain. Both the volatility functional form and Markov switching were introduced for electricity markets by Benth & Koekebakker [1] where various different volatility dynamics of forward curve models were specified based on the HJM approach.

Finally, for the estimation of the parameters of the proposed volatility functions a Bayesian solution is proposed. The Bayesian way to tackle the inference problem is to come up with the distribution of parameters and latent variables conditional on observed data. Next, MCMC techniques are used to explore these distributions. There has been a fair amount of theoretical and practical work on regime-switching models using MCMC methods. There were early contributions by Carlin and Polson [3] and Chib [5] & [6]. Johannes & Polson [23] provide a general algorithm based on the work by Scott [27]. A forward filtering backward sampling (FFBS) algorithm is applied to the case of regime-switching models. Frühwirth-Schnatter [14] and Frühwirth-Schnatter & Sass [18] further extend the approach and found that the MCMC framework outperformed the corresponding EM algorithm.

An application to energy markets was first found in Rambharat *et al.* [26] which estimates a discrete time model for electricity prices using a Markov Chain Monte Carlo (MCMC) approach. Also Seifert & Uhrig-Homburg [28] follow a similar approach to model European electricity spot prices. In this essay, the MCMC set-up and respective algorithms suggested by Frühwirth-Schnatter [14] and Hahn, Frühwirth-Schnatter & Sass [18] are implemented to estimate the parameters of the

model including the transition probabilities of the Markov Chain from all available forward curves on the market. This is also the framework adopted by Chiarella *et al.* [4] for the estimation of natural gas forward prices which is explained thoroughly and deployed in this essay.

# Chapter 2

## Forward price curve model

In this chapter a general multi-factor, multi-commodity (MFMC) model is introduced. Subsequently, in the next section the process of parameter estimation from historical data is explained.

### 2.1 Multi-factor single-commodity model

From the stylized facts detailed above in Section 1.1 it is clear that forward prices of different maturities are not perfectly correlated. Although the separate price curves usually move up and down together, the front end of the curve displays much more volatility than the end of the curve. Further, both parts of curve change shape in different ways (see Breslin *et al.* [2]). Therefore, a single factor, as incorporated in (2.1), is not sufficient to explain the behavior witnessed in the markets.

$$\frac{dF(t, T)}{F(t, T)} = \sigma(t, T) dz(t) \quad (2.1)$$

Here, the term  $F(t, T)$  stands for the forward price at time  $t$  with expiry  $T$ . And,  $z(t)$  is an independent Brownian motion. It is the sole source of uncertainty and is associated with a volatility function  $\sigma(t, T)$ . The volatility function could be interpreted as the *the time  $t$  volatility of the  $T$  maturity forward price return*. It determines the magnitude and direction of the single random shock that drives changes in the forward curve. However, more than one single factor of uncertainty needs to be introduced to capture forward curve changes other than parallel moves along the curve. This can be done by adding independent Brownian motions (Breslin *et al.* [2]).

The stochastic differential equation (SDE) below represents a general multi-factor model of the forward curve:

$$\frac{dF(t, T)}{F(t, T)} = \sum_{i=1}^n \sigma_i(t, T) dz_i(t) \quad (2.2)$$

The evolution of the forward price is now driven by  $n$  independent Brownian motions,  $dz_i(t)$ . For every source of risk there is a correspondent volatility function  $\sigma_i(t, T)$ . The respective volatility functions now define the magnitude and direction of multiple shocks that trigger changes along all points on the forward curve. As a result, (2.2) will represent changes in shapes where different parts of the forward curve will move in different ways (Breslin *et al.* [2]).

A typical example of a multi-factor model is a model with three risk factors. Three independent Brownians,  $dz_1(t)$ ,  $dz_2(t)$  and  $dz_3(t)$  correspond to the functions  $\sigma_1(t, T)$ ,  $\sigma_2(t, T)$  and  $\sigma_3(t, T)$  which respectively act to *shift*, *tilt* and *bend* the forward curve. The first function is positive for all maturities which implies that a positive shock to the system causes all prices to shift up. The second most important factor is a tilt factor which causes the short and long ends of the curve to move in opposite directions. Finally, the third factors is a bending factor which causes the long and short ends to move opposite to the middle. In practice, it is found that a number of three factors is usually adequate to capture the forward price dynamics (Breslin *et al.* [2]).

In our data sample, represented by  $F(t, T)$ , 12 different contracts are included. Therefore, there are 12 factors that can explain the variance of the evolution of the curve. However, only a few of these will be significant for explaining the variation in the forward curve. The respective volatility functions can easily be obtained from time series analysis. One method that can be applied is *principal components analysis* (PCA) or eigenvector decomposition of the covariance matrix of the forward returns. Sample covariance values between pairs of forward price returns are calculated. Then the eigenvectors of the covariance matrix yield approximate the factors governing changes in the forward curve (Breslin *et al.* [2]).

In order to start estimating the volatility functions, Equation (2.2) for a single commodity with  $n$  factors is rewritten. Ito's lemma is applied to (2.2), which renders the following forward curve dynamics:

$$\Delta \log F(t, t + \tau_j) = -\frac{1}{2} \sum_{i=1}^n \sigma_i(t, t + \tau_j)^2 \Delta t + \sum_{i=1}^n \sigma_i(t, t + \tau_j) \Delta z_i \quad (2.3)$$

Therefore, the difference of the log of forward prices with respective maturities  $\tau_j$  for  $j = 1, \dots, m$  are jointly normally distributed with mean  $-\frac{1}{2} \sum_{i=1}^n \sigma_i(t, t+\tau_j)^2 \Delta t$  and standard deviation  $\sqrt{\sum_{i=1}^n \sigma_i(t, t+\tau_j)^2 \Delta t}$ .

Next, the annualized sample covariance matrix of these forward prices  $F(t, T)$  is calculated and the eigen-decomposition is computed:

$$\Sigma = \Gamma \Lambda \Gamma^T \quad (2.4)$$

where

$$\Gamma = \begin{bmatrix} \nu_{11} & \nu_{12} & \dots & \nu_{1n} \\ \nu_{21} & \nu_{22} & \dots & \nu_{2n} \\ \dots & \dots & \dots & \dots \\ \nu_{n1} & \nu_{n2} & \dots & \nu_{nn} \end{bmatrix} \quad (2.5)$$

and

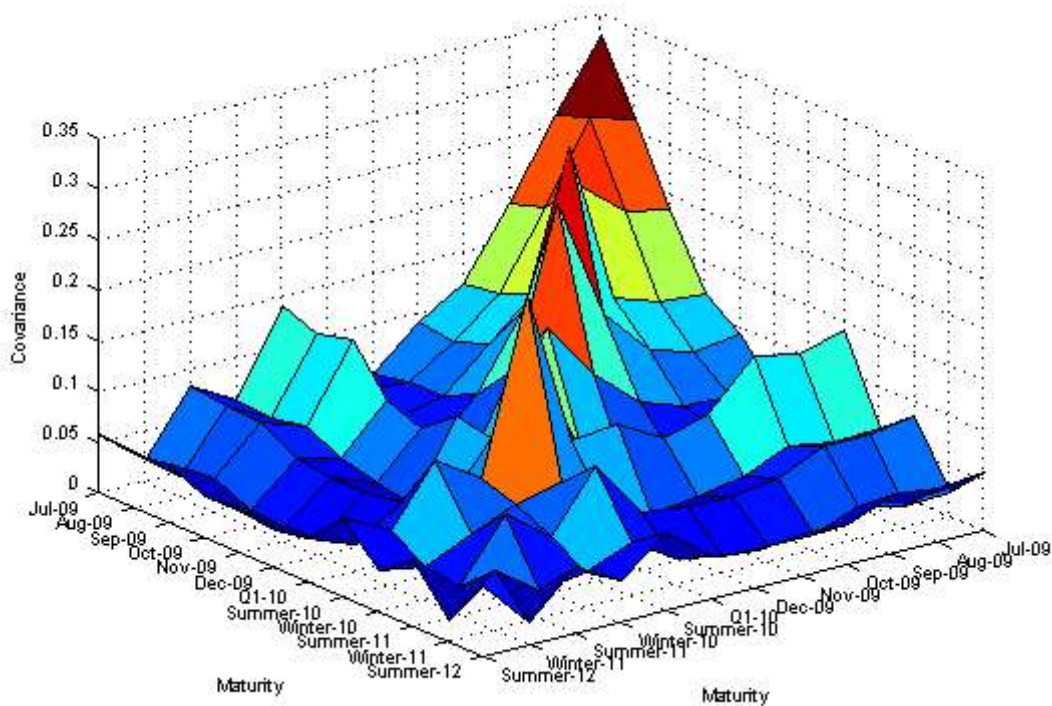
$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda_n \end{bmatrix} \quad (2.6)$$

. The eigenvectors are found in the columns of  $\Gamma$ . Then the variances of the factors that govern the points on the forward curve are the eigenvalues weighed by the corresponding eigenvectors (Breslin *et al.* [2]). The discrete volatility functions become thus,

$$\sigma_i(t, t+\tau_j) = \nu_{ji} \sqrt{\lambda_i} \quad (2.7)$$

Figure 2.1 shows the covariance surface of the data  $F(t, T)$ .

Figure 2.1: Annualized covariance surface for NBP data  $F(t, T)$

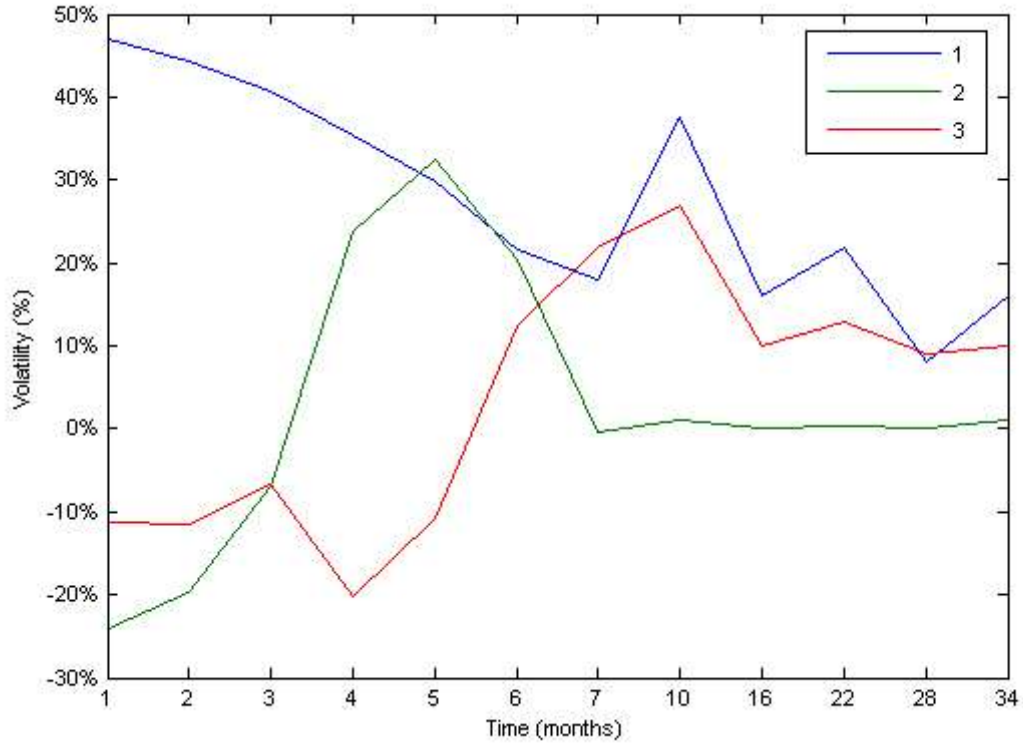


One observes that higher covariance values are found at the short end of the forward curve. For longer-dated contracts the surface will decay. The surface is not smooth due to illiquidity in longer-dated contracts and changes in market dynamics. However, another important source of *noise* is seasonality. It is therefore advised to repeat the PCA analysis for the different seasons defined (Breslin *et al.* [2]).



Next, from the covariance matrix the form of the volatility functions is obtained as described by (2.7). Figure 2.2 plots the first three volatility functions  $\sigma_1(t, T)$ ,  $\sigma_2(t, T)$  and  $\sigma_3(t, T)$  which respectively shift, tilt and bend the forward curve.

Figure 2.2: First three volatility functions for NBP data  $F(t, T)$



## 2.2 Spot - forward price relationship

The dynamics of the forward curve are related to spot price dynamics. The spot price is the price quoted at which a particular commodity can be bought or sold for *immediate* delivery. In the case of natural gas the delivery will take place the next day. This is called *day-ahead delivery*. Hence, the spot price is essentially the front end of the forward curve. Since the forward curve and spot dynamics are related, an equation for the evolution of the spot price can be written down which is consistent with forward price dynamics.

At the front end of the curve the maturity date is equal to the current date  $t$ . Further, (2.2) is integrated. Following a further differentiation, an SDE is obtained which describes the spot price dynamics. The spot price is given by,  $F(t, t) = S(t)$  (see Breslin *et al.* [2]).

$$\begin{aligned} \frac{dS(t)}{S(t)} = & \left[ \frac{\partial \ln F(0, t)}{\partial t} - \sum_{i=1}^n \left( \int_0^t \sigma_i(u, t) \frac{\partial \sigma_i(u, t)}{\partial t} du + \int_0^t \frac{\partial \sigma_i(u, t)}{\partial t} dz_i(u) \right) \right] dt \\ & + \sum_{i=1}^n \sigma_i(t, t) dz_i(t) \end{aligned} \quad (2.8)$$

The *drift* term of (2.8) is found between square brackets. It contains integrals over the Brownian motions from time zero to time  $t$ . That is, all random shocks from the start of the process up till the current date are incorporated in the spot price. The spot price process described is thus *non-Markovian*.

## 2.3 Seasonality adjustment

Further, as mentioned in Section 1.1, an important stylized fact of energy markets is seasonality in the forward price volatilities. Therefore, seasonality should be taken into account when modeling evolution of the forward curve. One could estimate a different volatility function for each season, where seasons are obtained by splitting the data. Examples are *summer/winter* and *summer/autumn/winter/spring*. Alternatively, the functions of (2.2) could be rewritten as a product of a time-dependent spot volatility function and maturity-dependent volatility functions (Breslin *et al.* [2]). (2.2) thus becomes:

$$\frac{dF(t, T)}{F(t, T)} = \sigma_s(t) \sum_{i=1}^n \sigma_i(T-t) dz_i(t) \quad (2.9)$$

where the spot price volatility at time  $t$  is described by  $\sigma_s(t)$  and  $\sigma_i(T-t)$ ,  $i = 1, \dots, n$  who represent the  $n$  maturity-dependent volatility functions.

## 2.4 Multi-factor multi-commodity model

Equation (2.2) could be further generalized to incorporate multiple commodities. Practitioners in energy risk management often have to deal with joint modeling of different commodities. For example, the *spark spread* is the spread between the power price and the price for natural gas used to produce the electricity. Spark spread modeling is critical in order to value power plants. Then, (2.10) describes the joint forward curve dynamics of multiple commodities as:

$$\frac{dF_c(t, T)}{F_c(t, T)} = \sum_{i=1}^{n_c} \sigma_{c,i}(t, T) dz_{c,i}(t) \quad (2.10)$$

where  $c = 1, \dots, m$  stands for the different commodities and  $i = 1, \dots, n_c$  points out the particular volatility function that corresponds to each commodity. The correlation between the different commodities is fixed by a correlation matrix for the Brownian motions. Obviously, the correlations between the Brownians governing a commodity,  $dz_i(t)$  for a given commodity  $c$ , are still zero. But now the correlations between the Brownian motions driving different commodities  $c$ ,  $dz_{c,i}(t)$ , represent the *inter-commodity* correlations (Breslin *et al.* [2]).

Equivalent to the approach taken for a single commodity, as described by equation (2.9), the discrete-time evolution of the forward curve can be written in terms of the estimated volatility functions. This results in the following matrix representation:

$$\bar{x}(t) = \bar{\mu}(t) + \tilde{\Sigma} \bar{\varepsilon} \quad (2.11)$$

where

- $\bar{x}(t)$  is the vector of changes in the natural logarithms of the forward prices for each maturity at the specified time step;
- $\bar{\mu}(t)$  is the vector of drift terms over the time step;

- $\tilde{\Sigma}$  is the matrix of discrete volatility function terms;
- $\bar{\varepsilon}$  is the unknown vector of standard normally distributed random shocks.

## 2.5 Two-factor regime switching forward price curve model

Building on the general framework for commodity forward price curves introduced in Section 2.1, the following two-factor regime switching model for the natural gas forward curve is proposed (see Chiarella *et al.* [4]). As indicated in Section 1.2 it is modeling a process in the class of piecewise-deterministic processes which provides a suitable framework for modeling the dynamics of the volatility term structure. In between two random jump times an almost deterministic process is followed by the volatility. Further, Markov Switching of the coefficients is imposed. Therefore, the model belongs to the class of Markov Switching models (Chiarella *et al.* [4]).

$$\frac{dF(t, T)}{F(t, T)} = \sigma_1(t, T)dz_1(t) + \sigma_2(t, T)dz_2(t) \quad (2.12)$$

where

- the natural gas forward price at time  $t$  with maturity  $T$  is denoted by  $F(t, T)$ .
- both volatility functions, that is for  $i = 1, 2$ , have different volatility and attenuation parameters depending on the state of the Markov Chain  $X_t$ . That is,

$$\sigma_i = (\sigma_{i1}, \sigma_{i2}, \dots, \sigma_{iN}), \alpha_i = (\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{iN}),$$

- further, the notation  $\langle \cdot, \cdot \rangle$  is introduced and stands for the scalar product in  $\mathbb{R}^N$  if  $u = (u_1, \dots, u_N)$  then

$$\langle u, X_t \rangle = \sum_{i=1}^N u_i \mathbf{I}_{(X_t=e_i)}$$

where the indicator function is represented by  $\mathbf{I}_{(X_t=e_i)}$ :

$$\mathbf{I}_{(X_t=e_i)} = \begin{cases} 1 & \text{if } X_t = e_i \\ 0 & \text{otherwise.} \end{cases}$$

$$\begin{aligned}
\sigma_1(t, T) &= \langle \sigma_1, X_t \rangle c(t) (e^{-\langle \alpha_1, X_t \rangle (T-t)} (1 - \sigma_{l1}) + \sigma_{l1}) \\
\sigma_2(t, T) &= \langle \sigma_2, X_t \rangle c(t) (\sigma_{l2} - e^{-\langle \alpha_2, X_t \rangle (T-t)}) \\
c(t) &= c + \sum_{j=1}^J (d_j (1 + \sin(f_j + 2\pi j t)))
\end{aligned} \tag{2.13}$$

where

- switching between parameter values is governed by a finite state Markov Chain  $X_t$ , with the state space defined by  $X = \{e_1, e_2, \dots, e_N\}$ , with  $e_i$  a vector of length  $N$  that has a value of 1 at the  $i$ -th position and 0 elsewhere, that is

$$e_i = (0, \dots, 0, 1, 0, \dots, 0)^T \in \mathbb{R}^N.$$

- the transition probability matrix of the Markov Chain  $X_t$  is defined by  $\xi = (\xi_{ij})_{N \times N}$ . Where  $\xi_{ij}$  is the conditional probability that the Markov Chain  $X_t$  switches from state  $e_i$  at time  $t$  to state  $e_j$  at time  $t + 1$ , for all  $i = 1, \dots, N, j = 1, \dots, N$ , that is

$$\xi_{ij} = \pi(X_{t+1} = e_j | X_t = e_i)$$

- $c(t)$  represents the seasonal part which is modeled as a truncated Fourier series

Therefore seasonality is introduced into the model by making the parameters periodic functions of time, with a periodicity of one year. For the application below, the Fourier series will consist of only one series, that is  $J = 1$ . The parameters  $c, d_1$  and  $f_1$  are then estimated.

- $z_1(t)$  and  $z_2(t)$  are independent Brownian motions.
- since the long term volatility behaves differently from short term behavior, parameters  $\sigma_{l1}$  and  $\sigma_{l2}$  are defined to describe long run volatility.

# Chapter 3

## Estimation

In this chapter a Bayesian approach is presented to estimate the parameters of the forward curve model described by (2.12) and (2.13). The model is a Markov Switching model: the parameters, found in the respective volatility functions  $\sigma_1(t, T)$  and  $\sigma_2(t, T)$ , take different values for different states of the world. Further, the state transitions are driven by a finite state Markov Chain,  $X$ . The set of parameters is defined as  $\theta$ :

$$\theta = \{\sigma_i, \alpha_i, \sigma_{ii}(i = 1, 2), c, d_j, f_j, (j = 1, 2, \dots, J)\} \quad (3.1)$$

The elements of the transition probability matrix  $\xi$  are also unknown and have to be inferred from the historical data as well.

For the econometric estimation of the states for drift and volatility,  $\theta$ , and the transition probability matrix of the underlying Markov Chain,  $\xi$ , a Markov Chain Monte Carlo (MCMC) method is suggested and implemented. In particular, a discrete-time Gibbs sampler is developed to estimate  $\theta$  and  $\xi$  given the forward prices at fixed observation times  $\Delta t, 2\Delta t, \dots, M\Delta t = t$ , where  $t$  is the present time. It is thus assumed that the state process can jump only at the discrete observation times (Chiarella *et al.* [4]).

### 3.1 Data transformation

So, for the different maturities the forward prices  $F(k\Delta t, T), k = 1, \dots, M$  are observed. Then in discrete-time the forward dynamics of (2.12) can be rewritten as

$$\begin{aligned}
y_{t,T} &= \log F(t\Delta t, T) - \log F(t, T) \\
&= -\frac{1}{2} \sum_{i=1}^2 \sigma_i(t, T)^2 \Delta t + \sum_{i=1}^2 \sigma_i(t, T) \Delta z_i(t).
\end{aligned} \tag{3.2}$$

One can see that difference of the log of forward price,  $y_{t,T}$  is normally distributed with mean

$$\mu(t, T) = -\frac{1}{2} \sum_{i=1}^2 \sigma_i(t, T)^2 \Delta t \tag{3.3}$$

and standard deviation

$$\sqrt{\sum_{i=1}^2 \sigma_i(t, T)^2 \Delta t}. \tag{3.4}$$

The notation for the forward dynamics  $y_{t,T}$  is abbreviated to  $\mathbf{y}_{\mathbf{T}}$ .

## 3.2 Bayesian estimation for a Markov switching model

In the Bayesian estimation for a Markov switching model, the latent Markov Chain  $X$  is introduced as missing data and samples for  $\theta$  and  $X$  are generated given  $\mathbf{y}_{\mathbf{T}}$  for parameter estimation. The technique of constructing sampling algorithms via the introduction of latent variables is called *data augmentation*. The sampling distribution is the posterior distribution,  $\pi(X_t, \theta | \mathbf{y}_{\mathbf{T}})$ , and the sampler applied here is a *Gibbs sampler* (Frühwirth-Schnatter [14]).

The posterior distribution  $\pi(X, \theta | \mathbf{y}_{\mathbf{T}})$  is *proportional to* the complete-data likelihood distribution  $\pi(\mathbf{y}_{\mathbf{T}} | X, \theta)$  which expresses the forward dynamics  $\mathbf{y}_{\mathbf{T}}$  given the states of the Markov Chain  $X$  and parameter values  $\theta$ . Further, the conditional distribution  $\pi(X | \theta)$  indicates that realizations of  $X$  are dependent on the parameter values  $\theta$ , and initial knowledge about the model parameters is expressed by defining a prior distribution for  $\theta$ ,  $pr(\theta)$ .

$$\pi(X, \theta | \mathbf{y}_{\mathbf{T}}) \propto \pi(\mathbf{y}_{\mathbf{T}} | X, \theta) \pi(X | \theta) pr(\theta) \tag{3.5}$$

Before drawing from this augmented complete-data likelihood distribution the prior distributions have to be chosen (Chiarella *et al.* [4]).

### 3.3 Prior distributions

Prior distributions are chosen for  $\theta$  and  $X_0$ , the state of the Markov Chain at  $t = 0$ . Under the assumption that the set of parameters  $\theta$  given by (3.1),  $\xi$  and  $X_0$  are independent, the priors can be written as follows

$$\pi(\theta, \xi, X_0) = pr(\xi) \prod_{i=1}^2 (pr(\sigma_i) \cdot pr(\alpha_i) \cdot pr(\sigma_{li})) pr(c) \prod_{j=1}^J (pr(d_j) pr(f_j)) pr(X_0).$$

Where the following distributions are assumed for the priors:

- The priors of volatility and attenuation parameters are  $\sigma_{il} \sim U(a_{il}^\sigma, b_{il}^\sigma)$ , and  $\alpha_{il} \sim U(a_{il}^\alpha, b_{il}^\alpha)$  where  $a_{il}^\sigma$  and  $a_{il}^\alpha$  are discrete values which define the parameters of the uniform distribution. As a result, the volatility and attenuation parameters are bounded which eases convergence.
- The priors of the long run volatility parameters,  $\sigma_{li}$ , are uniformly distributed as well,  $\sigma_{li} \sim U(0, 1)$
- The rows  $\xi_l$  of the transition probability matrix  $\xi$ , are assumed to be independent and to follow a Dirichlet distribution  $\xi_l \sim D(g_{l1}, \dots, g_{lN})$  where the vector  $g_l$  equals the prior expectation of  $x_{il}$  times a constant that determines the variance. If  $\xi_0$  denotes our prior expectation of  $\xi$ , we may set  $g_l = \xi_l^0 c_p$ . Then  $c_p$  can be interpreted as the number of observations of jumps out of state  $l$  in the prior distribution. If the number of states,  $N$ , is restricted to 2, the Dirichlet distribution simplifies to a Beta distribution.
- The priors for the parameters of the seasonal part,  $c, d, f$ , are uniformly distributed  $c, d, f \sim U(a^c, b^c)$
- The prior of the initial state of the Markov Chain is following a discrete uniform distribution,  $X_0 \sim U(\{1, \dots, N\})$  where  $N$  is equal to the number of states defined for the Markov Chain.

All parameters  $g_{l1}, \dots, g_{lN}, a_{il}^\sigma, b_{il}^\sigma, a_{il}^\alpha, b_{il}^\alpha, a^c, b^c$  in the prior distributions are defined before running the Gibbs sampler (Chiarella *et al.* [4]).



## 3.4 Gibbs sampler

Starting from the prior distributions specified in Section 3.2, the unknowns are partitioned into three blocks  $\theta$ ,  $\xi$  and  $X$ . Next, a three-step MCMC sampler is used to sample from the augmented posterior distribution  $\pi(\theta, \xi, X | \mathbf{y}_T)$  given the observed data  $\mathbf{y}_T$  where each part is drawn from the proper conditional density (Chiarella *et al.* [4]).

The particular scheme is explained in Algorithm 1 below. Conditional on knowing the state of the finite Markov Chain  $X$ ,  $\theta$  is sampled and  $X$  is sampled conditional on knowing  $\theta$  (Frühwirth-Schnatter [14]).

**Algorithm 1: Unconstrained MCMC for a Markov Switching Model**

Start with some state process  $X^{(0)}$  and repeat the following steps for  $k = 1, \dots, K_0, \dots, K + K_0$ . The notation  $K_0$  represents the number of iterations for the *burn-in*. These samples are thrown away for the calculation of the estimators. Clearly, the length of burn-in depends on the initial state of the Markov Chain state,  $X_0$ , and the rate of convergence.

1. **Parameter simulation** conditional on the states  $X^{(k-1)}$ 
  - (a) Sample the transition matrix  $\xi^{(k)}$  from the complete-data posterior distribution  $\pi(\cdot | X^{(k-1)})$ .
  - (b) Sample the model parameters  $\theta^{(k)}$  from the complete-data posterior  $\pi(\theta | \mathbf{y}_T, X^{(k-1)})$ . Store the actual values of all parameters  $\theta^{(k)}$  and  $\xi^{(k)}$ .
2. **State simulation** conditional on knowing  $\theta^{(k)}$  by sampling a path  $X$  of the hidden Markov Chain from the conditional posterior  $\pi(X | \theta^{(k)}, \mathbf{y}_T)$  using Algorithm 3, which will be discussed below.

### 3.4.1 Full conditional posterior distributions

Next, the full conditional posteriors required for the three sampling steps are derived. The *complete-data likelihood function* is a product of normal densities, described by  $\phi(\mathbf{y}_T, \mu, \sigma)$  with the with mean  $\mu_{k\Delta t, T}$  and standard deviation  $\sigma_{k\Delta t, T}$ , given by the moments of ( 3.5 ) (Chiarella *et al.* [4]).

$$\pi(\mathbf{y}_T | \theta, \xi, X) = \prod_{k=1}^M \phi(y_{k\Delta t, T}, \mu_{k\Delta t, T}, \sigma_{k\Delta t, T}) \quad (3.6)$$

Further the full conditional of  $\theta$ , the *conditional joint distribution of the parameters associated with drift and volatility* is given by

$$\pi(\theta|\mathbf{y}_T, \xi, X) \propto \pi(\mathbf{y}_T|\theta, X)pr(\theta) \quad (3.7)$$

Finally, the *full conditional posterior of the state process*  $X_{l\Delta_t}$ ,  $l = 1, 2, \dots, M$ , is

$$\pi(X|\mathbf{y}_T, \theta, \xi) \propto \pi(\mathbf{y}_T|\theta, X)\pi(X|\xi) \quad (3.8)$$

where the prior distribution  $\pi(X|\xi)$  is not dependent on  $\theta$  but rather specified by the distribution of  $X_0$  and the transition matrix  $\xi$ . The conditional probability  $\pi(X|\xi)$  in ( 3.8 ) is provided by

$$\pi(X|\xi) = \pi(X_0|\xi) \prod_{k=1}^M \pi(X_k|X_{k-1}, \xi) = \pi(X_0|\xi) \prod_{l,q=1}^N \xi_{lq}^{N_{lq}} \quad (3.9)$$

where the notation  $N_{lq}$  stands for the number of transitions from state  $l$  to  $q$ , that is  $N_{lq} = \sum_{k=1}^M \mathbf{I}_{(X_{k-1}=l, X_k=q)}$  (Chiarella *et al.* [4]).

### 3.4.2 Sample the transition matrix

The MCMC routine Algorithm 1 was introduced in Section 3.4 to infer the parameters of the volatility functions. In the next sections the several steps involved will be explained in more detail. First, from the complete-data posterior distribution  $\pi(\xi|X^{(k-1)})$ , the transition matrix  $\xi^{(k)}$  is sampled (Frühwirth-Schnatter [14]).

Given the transition matrix  $\xi^{(k)}$  of the hidden Markov Chain, each row is an unknown probability distribution that has to be estimated from the data. Therefore, complete-data estimation is carried out for a given path  $X$ .

The complete-data likelihood  $\pi(X|\xi)$  is given by,

$$\pi(X|\xi) = \prod_{j=1}^2 \prod_{k=1}^2 \xi_{jk}^{N_{jk}(X)} \quad (3.10)$$

where 2 states are defined for the Markov Chain and  $N_{jk}$  counts the number of transitions from  $j$  to  $k$  (Frühwirth-Schnatter [14]).

In order to determine the particular algorithm, assumptions regarding the distribution  $\pi_0$  of the initial value  $X_0$  have to be made. In this essay the initial

distribution  $\pi_0$  is assumed to be independent of  $\xi$ . Now a classical Bayesian inference problem, Gibbs sampling from the conditional posterior  $\pi(\xi|X)$  can be applied. The rows  $\xi_j$  of  $\xi$ , independent a posteriori, are drawn from the following Dirichlet distribution (Frühwirth-Schnatter [14]).

$$\xi_j \sim \mathcal{D}(e_{j1} + N_{j1}(X), \dots, e_{jK}(X) + N_{jK}(X)) \quad j = 1, \dots, N \quad (3.11)$$

where  $N_{jk}(X) = \#\{X_{t-1} = j, X_t = k\}$  counts the numbers of transitions from  $j$  to  $k$  for the actual draw of  $X$ .

Since only two distinct states are defined for the transition matrix  $\xi$ , that is  $N = 2$ , the persistence probabilities  $\xi_{11}$  and  $\xi_{22}$ , both independent, are each following a Beta distribution.

$$\xi_{11} \sim \mathcal{B}(e_{11} + N_{11}(X), e_{12} + N_{12}(X)) \quad \xi_{22} \sim \mathcal{B}(e_{21} + N_{21}(X), e_{22} + N_{22}(X)) \quad (3.12)$$

These densities generalize to the Dirichlet distribution for the case  $N > 2$  (Frühwirth-Schnatter [14]).

### 3.4.3 Sample the model parameters

A single-component *Metropolis-Hastings* algorithm is set up in order to update the parameter set  $\theta$ . The candidate  $\theta^*$  is generated by a so-called *candidate generating density*  $q(\theta, \theta^*)$ . Here, a normal random walk is used to draw  $\theta^*$ .

$$\theta^* = \theta + r^\theta \psi \quad (3.13)$$

where  $\psi$  is a vector of independent standard normal random variables and  $r^\theta$  is a scalar which represents a fixed value for the step width.

Further, instead of updating the elements of  $\theta$  all at once, it is more straightforward and computationally efficient to split  $\theta$  into its components  $\{\theta_1, \theta_2, \dots, \theta_d\}$ , with  $d$  the number of parameters, and update one by one, that is sequentially.

**Algorithm 2: single-component Metropolis-Hastings**

1. **Initialize vector**  $\theta^0$  at some values for its components.

Different starting values could be chosen. One could, for example, look at parameter values found in previous research.

2. Let  $\theta_i^{(k)}$  denote the value of  $\theta_i$  at end of iteration  $k$ . For step  $i$  of iteration  $k + 1$ ,  $\theta_i$  is updated.

The **candidate is generated from proposal distribution**  $q_i(\theta_i|\theta_i^{(k)}, \theta_{-i}^{(k)})$  where  $\theta_{-i}^{(k)}$  denotes the value of  $\theta_{-i}$  after completing step  $i - 1$  of iteration  $k + 1$ :  $\theta_{-i}^{(k)} = \{\theta_1^{(k+1)}, \theta_2^{(k+1)}, \dots, \theta_{i-1}^{(k+1)}, \theta_{i+1}^{(k)}, \dots, \theta_h^{(k)}\}$

where components  $1, 2, \dots, i - 1$  have already been updated. So, only for the  $i$ -th component of  $\theta$  a candidate is generated and this candidate is dependent on the current values of the other elements of  $\theta$ .

3. Next, the **candidate is accepted** with probability  $\alpha(\theta_{-i}^{(k)}, \theta_i^{(k)}, \theta_i^*)$  where

$$\alpha(\theta_{-i}, \theta_i, \theta_i^*) = \min \left[ 1, \frac{\pi(\theta_i^*|\theta_{-i})q_i(\theta_i|\theta_i^*, \theta_{-i})}{\pi(\theta_i|\theta_{-i})q_i(\theta_i^*|\theta_i, \theta_{-i})} \right] \quad (3.14)$$

where the full conditional distribution for  $\theta_i$  under  $\pi(\cdot)$  is denoted by  $\pi(\theta_i|\theta_{-i})$ .

4. **Generate a random number**  $\sim U(0, 1)$ .

If this random number is less than  $\alpha(\theta_{-i}, \theta_i, \theta_i^*)$  then the move is accepted. Otherwise, the random number is greater than the acceptance probability, thus the proposed move for component  $i$  of parameter set  $\theta$  is rejected. So, if  $\theta_i^*$  is accepted,  $\theta_i^{(m+1)} = \theta_i^*$ ; otherwise  $\theta_i^{(m+1)} = \theta_i^{(m)}$ . At this step  $\theta_{-i}^{(m)}$  is left untouched.

Notice that  $q_i(\theta_i|\theta_i^*, \theta_{-i}) = q_i(\theta_i^*|\theta_i, \theta_{-i})$  by the symmetry of the normal distribution. Therefore, the factor

$$\frac{q_i(\theta_i^{(k)}|\theta_i^*, \theta_{-i}^{(k)})}{q_i(\theta_i^*|\theta_i^{(k)}, \theta_{-i}^{(k)})}$$

drops out of the Metropolis probability calculation (3.14). This greatly simplifies the calculation of  $\alpha$ . This special case of the Metropolis-Hastings algorithm is often referred to as the *random-walk Metropolis-Hastings* (Frühwirth-Schnatter [14]).

When the random-walk Metropolis algorithm is used as candidate generating density  $q(\cdot, \cdot)$ , the standard deviation of the normal random variable is the step size  $r^\theta$ . If that step size is too high, candidate draws  $\theta_i^*$  will be very far away from the current draw, and in a region of the parameter space that has low probability. This would lead to a high rejection rate, and a Markov Chain that doesn't move very often. Whereas, when  $r^\theta$  is chosen to be too low, then candidate draws will be very close to the current draw. Hence, the algorithm accepts frequently, but the chain will move slowly around the parameter space. So, it is a good idea to experiment with different values for step size. And, in general, it is wise to run multiple chains from disperse starting values, to help assess convergence (Frühwirth-Schnatter [14]).

Finally, the updating order for the elements of  $\theta$  in above Algorithm 2 is assumed to be fixed. But random permutations of the updating order are acceptable as well.

### 3.4.4 Markov Chain state simulation

In a final step of the Gibbs sampler, as described by Algorithm 1, a path  $X$  of the hidden Markov Chain from the conditional posterior  $\pi(X | (\theta^{(k)}, \xi^{(k)}), \mathbf{y}_T)$  is sampled conditional on  $(\theta^{(k)}, \xi^{(k)})$ . In this section this sampling method for the hidden Markov Chain  $X = (X_0, \dots, X_T)$ , is worked out in great detail.

#### Sampling Posterior Paths of the Hidden Markov Chain

A resourceful way to sample is called *multi-move sampling* where the states of the whole path  $X$  are jointly sampled from the conditional posterior distribution  $\pi(X | \mathbf{y}_T, \theta)$ . First, the joint posterior  $\pi(X | \mathbf{y}_T, \theta)$  is written as follows

$$\pi(X | y, \theta) = \left[ \prod_{t=0}^{T-1} \pi(X_t | X_{t+1}, \dots, X_T, \theta, \mathbf{y}_T) \right] \pi(X_T | \mathbf{y}_T, \theta) \quad (3.15)$$

where the filtered probability distribution at  $t = T$  is denoted by  $\pi(X_T | \mathbf{y}_T, \theta)$ .

Further,

$$\pi(X_t | X_{t+1}, \dots, X_T, \theta, \mathbf{y}_T) \propto \xi_{X_t, X_{t+1}} \pi(X_T | \mathbf{y}_T^t, \theta) \quad (3.16)$$

where  $\pi(X_T | \mathbf{y}_T^t, \theta)$  is the filtered probability distribution at time  $t$ . Here, the assumption is taken that the Markov Chain  $X_t$  is homogeneous. As a consequence,  $\pi(X_{t+1} | X_t, \theta, \mathbf{y}_T^t)$  is simplified to  $\xi_{X_t, X_{t+1}}$ .

Given the above-mentioned distributions it is now possible to implement a *forward-filtering-backward-sampling algorithm*.

**Algorithm 3: Multi-Move Sampling of the States**

Given fixed values for parameter set  $\theta$ , execute the following steps to sample a path  $X^k$  of the hidden Markov Chain.

1. Obtain **Filtered state probabilities**

First, conditional on  $\theta$ , run the filter described in Algorithm 3.1. The resulting filtered state probability distribution  $\pi(X_t = j|y^t, \theta)$   $j = 1, 2$  for  $t = 1, \dots, T$  is then stored.

For a state space model, the filtering problem implies statistical inference about the state variable conditional on observations from the beginning up till time  $t$ . The complete filtering distribution is obtained by the discreteness of the support of the state variable  $X_t$ .  $\pi(X_t = l|\mathbf{y}_T^t, \theta)$  for all possible realizations  $l \in \{1, 2\}$  of  $X_t$ .

**Algorithm 3.1: Filtering the States**

For  $t = 1, 2, \dots, T$ , the following steps are executed sequentially.

(a) **One-step ahead prediction** of  $X_t$ :

$$\pi(X_t = l|y^{t-1}, \theta) = \sum_{k=1}^2 \xi_{kl} \pi(X_{t-1} = k|\mathbf{y}_T^{t-1}, \theta) \quad (3.17)$$

for  $l = 1, 2$  where  $\pi(X_t = l|X_{t-1} = k, \mathbf{y}_T^{t-1}, \theta)$  simplifies to the transition probability  $\xi_{kl}$  for homogeneous Markov Chains.

(b) **Filtering** for  $X_t$ :

$$\pi(X_t = l|y^t, \theta) = \frac{\pi(y_t|X_t = l, \mathbf{y}_T^{t-1}, \theta) \pi(X_t = l|\mathbf{y}_T^{t-1}, \theta)}{\pi(y_t|\mathbf{y}_T^{t-1}, \theta)} \quad (3.18)$$

where

$$\pi(\mathbf{y}_T^t|\mathbf{y}_T^{t-1}, \theta) = \sum_{k=1}^2 \pi(\mathbf{y}_T^t|X_t = k, \mathbf{y}_T^{t-1}, \theta) \pi(X_t = k|\mathbf{y}_T^{t-1}, \theta) \quad (3.19)$$

At  $t = 1$ , the filter is started with the initial distribution  $\pi(X_0 = k|\xi)$ .

So,  $\pi(X_1 = l|\mathbf{y}_T^0, \theta) = \sum_{k=1}^2 \xi_{kl} \pi(X_0 = k|\xi)$ .

As is typical for any filter, the discrete filter described in Algorithm 3 is an adaptive inference tool. At time  $t - 1$ , the filtered probabilities  $\pi(X_{t-1} = l | \mathbf{y}_{\mathbf{T}}^{t-1}, \theta)$  summarize, for a fixed value of  $\theta$ , all information the observations  $y_1, \dots, \mathbf{y}_{\mathbf{T}}^{t-1}$  contain about  $X_{t-1}$ . To obtain inference about  $X_t$  at time  $t$  in terms of the posterior distribution  $\pi(X_t = l | \mathbf{y}_{\mathbf{T}}^t, \theta)$ , knowledge of the posterior distribution  $\pi(X_{t-1} = l | \mathbf{y}_{\mathbf{T}}^{t-1}, \theta)$  at time  $t - 1$  and only the actual value of  $\mathbf{y}_{\mathbf{T}}^t$  are sufficient.

2. **Sample  $X_T^{(k)}$  from the filtered state probability distribution  $\pi(X_t = j | \mathbf{y}_{\mathbf{T}}^t, \theta)$**
3. For  $t = T-1, T-2, \dots, 0$  **sample  $S_t^{(k)}$  from the conditional distribution  $\pi(X_t = j | X_{t+1}^{(k)}, y^t, \theta)$  given by**

$$\pi(X_t = j | X_{t+1}^{(k)}, \mathbf{y}_{\mathbf{T}}^t, \theta) = \frac{\xi_{j,l_m} \pi(X_t = j | \mathbf{y}_{\mathbf{T}}^t, \theta)}{\sum_{k=1}^K \xi_{k,l_m} \pi(X_t = k | \mathbf{y}_{\mathbf{T}}^t, \theta)}$$

where  $l_m$  is equal to the state of  $S_{t+1}$  and  $\pi(X_{t+1} = l_m | X_t = j, \theta, \mathbf{y}_{\mathbf{T}}^t)$  reduces to  $\xi_{j,l_m}$  since the Markov Chain  $X_t$  is homogeneous.

For each  $t$ ,  $\pi(X_t | X_{t+1}^m, \mathbf{y}_{\mathbf{T}}^t, \theta)$  needs to be evaluated for all  $j = 1, \dots, K$ . This requires knowledge of the filtered state probabilities  $\pi(X_t = j | \mathbf{y}_{\mathbf{T}}^t, \theta)$ , which were stored in step 1 of Algorithm 3.

# Chapter 4

## Implementation

In this chapter, the details of the implementation are laid out and the numerical results for the routines are provided. Before applying the chosen methodology to historical gas forward prices, the algorithms are tested with simulated data. Further, estimators are given for both a one-factor model and the two-factor model described in Section 2.5.

### 4.1 Data

The natural gas prices modeled are forward prices for the *National Balancing Point* (NBP), which is the British virtual trading location operated by TSO National Grid, covering all entry and exit points in mainland Britain. It is the most liquid gas trading point in Europe. Natural gas at NBP trades on a forward month, forward quarter, forward season or year forward basis and prices are expressed in pence (GBP) per therm. For parameter estimation of the chosen model, daily observations for the forward prices were obtained with maturities ranging from July 2009 to September 2012. The observations span over a year with prices at times  $t$  from 23/10/2007 to 12/03/2009. This data is represented by  $F(t, T)$ . Further, in order to obtain the forward dynamics  $\mathbf{y}_T$  for every maturity  $T$ , the log difference transform is applied to the original data  $F(t, T)$ .

### 4.2 Choosing the prior

The complete-data likelihood and the full conditionals introduced in Chapter 3 contain prior knowledge about the parameter set  $\theta$ , and  $X_0$ . There is some degree of freedom in the choice of prior distribution but the prior distribution should be



independent of the data and the posterior distribution. The following priors are used in the implementation (Chiarella *et al.* [4]):

- For the transition matrix  $\xi$  the vector  $g_l$  equals the prior expectation of  $x_{i_l}$  times a constant that determines the variance. If  $\xi_0$  denotes the prior expectation of  $\xi$ , one may set  $g_l = \xi_l^0 c_p$ . Then  $c_p$  can be interpreted as the number of observations of jumps out of state  $l$  in the prior distribution. Hence, the matrix of parameters for the Dirichlet distribution is set to

$$g = \begin{bmatrix} 0.64 & 0.36 \\ 0.16 & 0.84 \end{bmatrix} \quad (4.1)$$

- The priors of  $\sigma$ s and  $\alpha$ s, ( $i, l = 1, 2$ ) are  $\sigma_{il} \sim U(0, 1.5)$ , and  $\alpha_{il} \sim U(0, 5)$
- The priors of  $\sigma_{i_i}$  are  $\sigma_{i_i} \sim U(0, 1)$ , ( $i = 1, 2$ )
- The priors of  $c, d, f(J = 1)$  are  $c, d_1, f_1 \sim U(0, 1)$
- The prior of the initial state of the Markov chain is  $X_0 \sim U(\{1, 2\})$

### 4.3 Running the Gibbs sampler

A *randomized* algorithm is used to calculate the complete-date likelihood function. This way, all observations along the forward price curve at each time  $t$  are taken into account. Before estimation a string  $\lambda$  is selected,  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_N)$  with  $\lambda_k \sim U(\{1, 2, \dots, 12\})$ ,  $k = 1, \dots, N$  where  $\lambda_k$  ranges from 1 to 12 to label 12 different maturities. At each time  $k \cdot \Delta t$  in the data series, the data  $y_{t,T}$  corresponding to the maturity  $T_j$  is picked if  $\lambda_k = j$ . The complete-data likelihood function thus becomes

$$\pi(y_{T_\lambda} | \theta, \xi, X) = \prod_{k=1}^N \phi(y_{k\Delta t, T_{\lambda_k}}, \mu_{k\Delta t, T_{\lambda_k}}, \sigma_{k\Delta t, T_{\lambda_k}}) \quad (4.2)$$

where  $\phi$ ,  $\mu$  and  $\sigma$  defined as in Section 3.1 (Chiarella *et al.* [4]).

Finally, to implement the Metropolis-Hastings algorithm, the scaling factors  $r^\theta$  have to be chosen. Fixing  $r^\theta$  around 2% to 5% seemed to work well (Chiarella *et al.* [4]).

### 4.3.1 One-factor model

First, the second volatility function,  $\sigma_2(t, T)$  is omitted in (2.12) such that one is left with a one-factor model. Then the parameters for the first volatility function  $\sigma_1(t, T)$  are defined as in Section 2.5.

$$\frac{dF(t, T)}{F(t, T)} = \sigma_1(t, T) dz_1(t) \quad (4.3)$$

Next, the implementation of the Gibbs sampler is verified by means of *simulated forward data*. The code in Appendix A (p.) generates forward price dynamics,  $\mathbf{y}_T$  following formula (3.2). After which the model parameters are inferred with the Gibbs sampler.

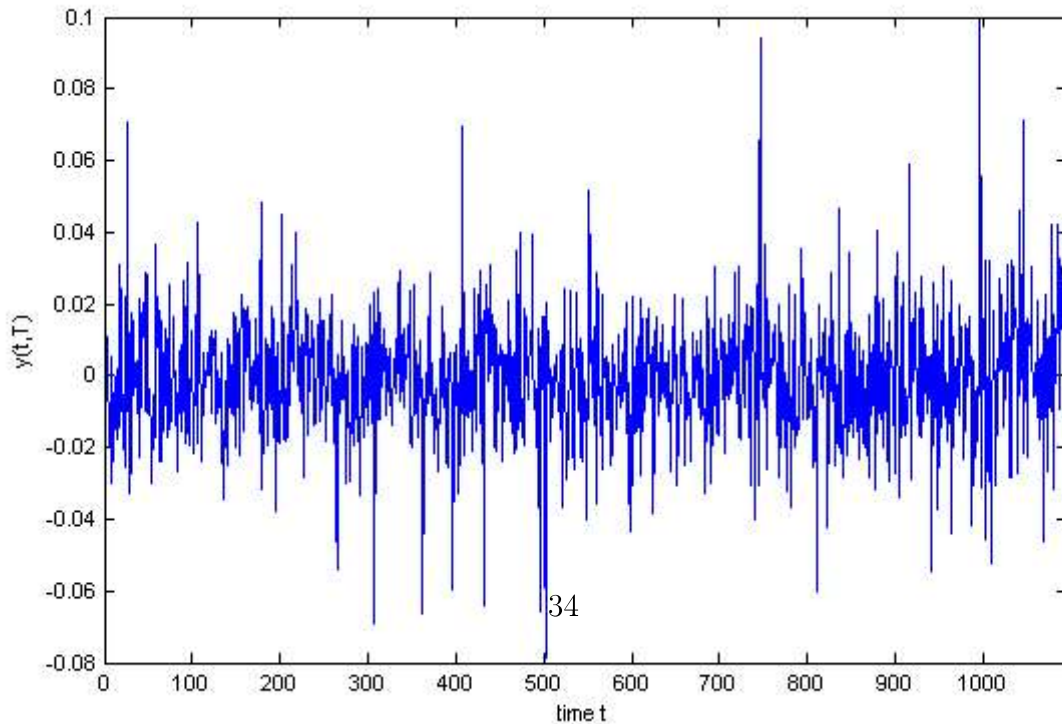
Figure 4.1 below shows the forward price dynamics  $\mathbf{y}_T$  of a one-factor regime switching model. The transition probabilities  $\xi$  were equal to

$$\xi = \begin{bmatrix} 0.9 & 0.1 \\ 0.7 & 0.3 \end{bmatrix}$$

and, the parameter set  $\theta$

$$\theta = \{\sigma_{11} = 0.5, \sigma_{12} = 1, \alpha_{11} = 2, \alpha_{12} = 1.75, \sigma_{11} = 0.6, c = 0.8, d = 0.08, f = 0.31\}$$

Figure 4.1: Generated forward price dynamics,  $\mathbf{y}_T$



Next, parameter estimation is carried out with the Gibbs sampler. For this very first stage of the implementation, the parameters for the seasonality function  $c(t)$  are not estimated. The respective parameters  $c$ ,  $d$  and  $f$  are fixed at their true values 0.8, 0.08 and 0.31.

In Appendix B (p.) one can find the routine for the Gibbs sampler which is applied to the generated forward price dynamics,  $\mathbf{y}_T$ .

### 4.3.2 Two-factor model

Forward price dynamics are simulated to check the implementation of the Gibbs sampler for the two-factor model as expressed by (2.12). The dynamics  $\mathbf{y}_T$  are generated calling the routine **nbp\_2fac.m**, which is found in Appendix C (p.)

Also in the case of the two-factor model, the parameters of the seasonality function,  $c(t)$  are initially fixed. Below one can find the routine for the Gibbs sampler which is applied to the generated forward price dynamics,  $\mathbf{y}_T$ . The code for this discrete-time sampler is provided in Appendix D (p.)

## 4.4 Numerical Results

### 4.4.1 One-factor model

For a two-state regime, one-factor model, about 100,000 steps were run with simulated data. A rate of 20% is chosen for the burnin which means that the first 20% of the output is thrown away and does not contribute to the estimated values.

At this stage of the implementation, the values for the seasonality function  $c(t)$  are kept fixed. That is, the parameters  $c, d$  and  $f$  are not included in the parameter set  $\theta$ . Parameters are respectively  $c = 0.81$ ,  $d = 0.08$ ,  $f = 0.31$ , which are seasonality parameters obtained from previous estimations by Chiarella *et al.* [4].

Below one can find the estimates of the transition probability matrix  $\xi$  and the parameter set  $\theta$ .

- Number of states = 2
- Transition probability matrix  $\xi$

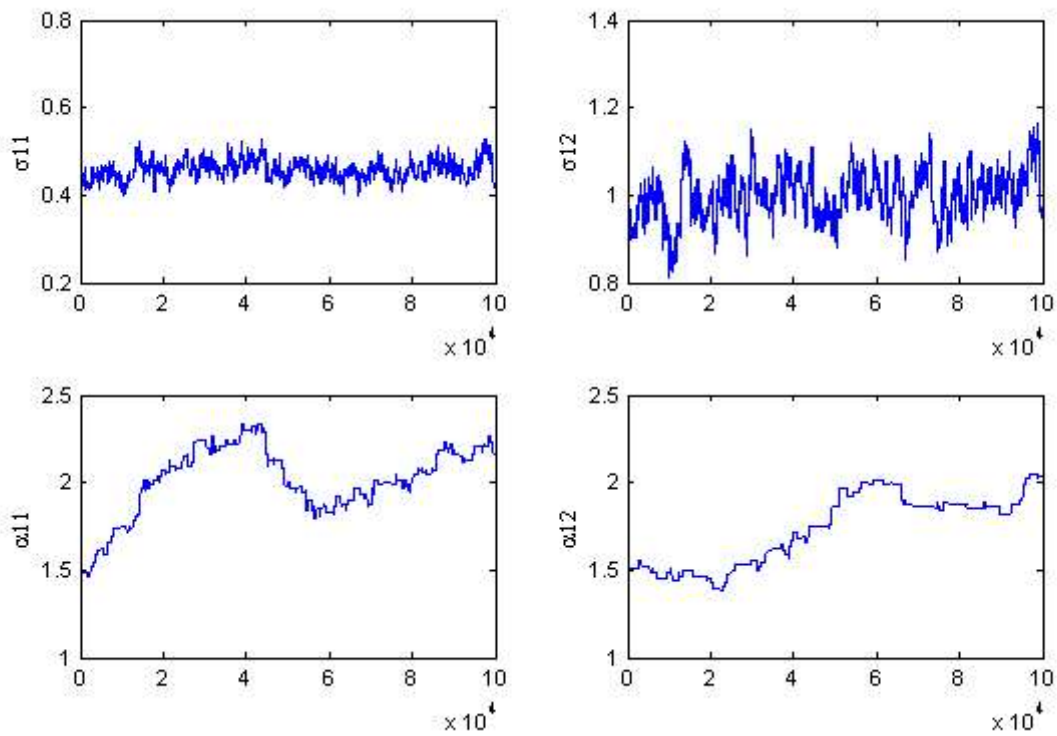
Simulated data				
	Actual		Estimated	
	1	2	1	2
1	0.9	0.1	0.88	0.12
2	0.7	0.3	0.72	0.28

- Regime-switching parameters  $\sigma_{1j}$  and  $\alpha_{1j}$ , ( $j = 1, 2$ )

Simulated data				
	Actual		Estimated	
$j$	1	2	1	2
$\sigma_{1j}$	0.5	1	0.47	1.01
$\alpha_{1j}$	2	1.75	2.06	1.78

Figure 4.2 below shows the parameter estimates throughout the runs. Convergence is reached fast for the volatility parameters  $\sigma_{1j}$ , ( $j = 1, 2$ ), but also the values for the attenuation parameters  $\alpha_{1j}$  reach the actual values of the underlying forward dynamics  $\mathbf{y}_T$ .

Figure 4.2: Path of parameter estimates for  $\sigma_{1j}$  and  $\alpha_{1j}$ , ( $j = 1, 2$ )



- estimator for  $\sigma_{l1}$ ,  $\sigma_{l1} = 0.62$ , with an actual value of 0.6

## 4.4.2 Two-factor model

For a two-state regime, two-factor model, about 100,000 steps were run with simulated data. Any parameters that lie within the boundaries set by the prior distributions could be chosen for the simulated forward dynamics. Here the following parameter values were used to generate  $\mathbf{y}_T$ .

The transition probabilities  $\xi$  were equal to

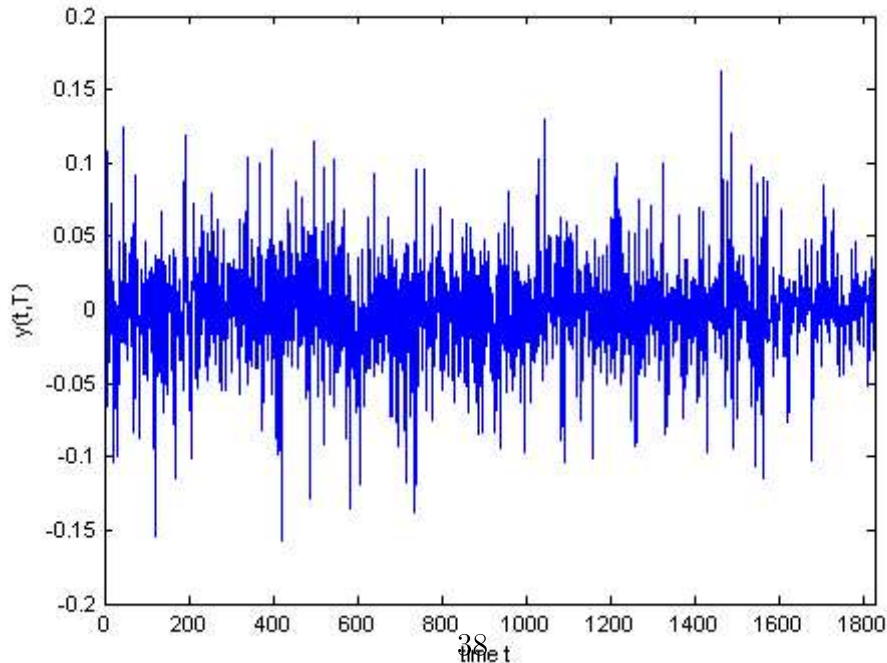
$$\xi = \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix}$$

and, the parameter set  $\theta$

$$\theta = \{ \sigma_{11} = 0.3, \sigma_{12} = 0.8, \alpha_{11} = 1.5, \alpha_{12} = 2, \sigma_{l1} = 0.5, \\ \sigma_{21} = 0.6, \sigma_{22} = 1.2, \alpha_{21} = 1, \alpha_{22} = 3, \sigma_{l2} = 0.9, c = 0.8, d = 0.08, f = 0.31 \}$$

Figure 4.3 below shows the forward price dynamics  $\mathbf{y}_T$  of a two-factor regime switching model.

Figure 4.3: Generated forward price dynamics,  $\mathbf{y}_T$



Next, the Gibbs sampler implemented for the two-factor regime should be able to retrieve those values. As previously stated, a rate of 20% is chosen for the burnin which means that the first 20% of the output is thrown away and does not contribute to the estimated values.

Below one can find the estimates of the transition probability matrix  $\xi$  and the parameter set  $\theta$ .

1. With fixed seasonality function  $c(t)$  Parameters are respectively  $c = 0.81$ ,  $d = 0.08$ ,  $f = 0.31$  Those seasonality parameters were obtained from previous estimations by Chiarella *et al.* [4].

- Number of states = 2
- Transition probability matrix  $\xi$

Simulated data				
	Actual		Estimated	
	1	2	1	2
1	0.8	0.2	0.83	0.17
2	0.6	0.4	0.63	0.37

- Regime-switching parameters  $\sigma_{1j}$  and  $\alpha_{1j}$ , ( $j = 1, 2$ )

Simulated data				
	Actual		Estimated	
$j$	1	2	1	2
$\sigma_{1j}$	0.3	0.8	0.37	0.78
$\sigma_{2j}$	0.6	1.2	0.67	1.31
$\alpha_{1j}$	1.5	2	1.67	2.69
$\alpha_{2j}$	1	3	1.01	3.59

Figures 4.4 and 4.5 below show the parameter estimates throughout the runs. Convergence is reached fast for the volatility parameters  $\sigma_{ij}$ , ( $i, j = 1, 2$ ). The values for the attenuation parameters  $\alpha_{ij}$  are near the actual values of the underlying forward dynamics  $\mathbf{y}_T$ . One could increase the number of runs.

Figure 4.4: Path of parameter estimates for  $\sigma_{ij}$ , ( $i, j = 1, 2$ )

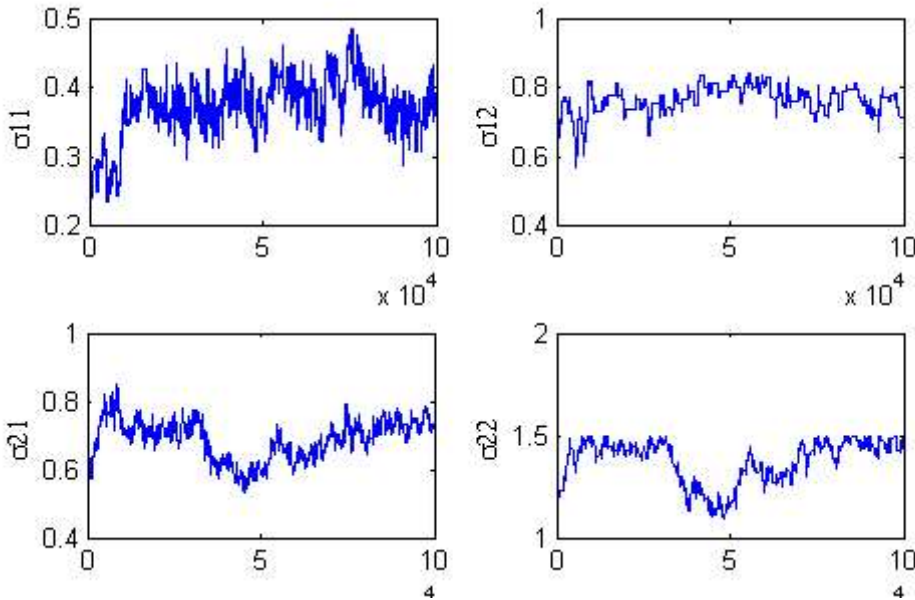
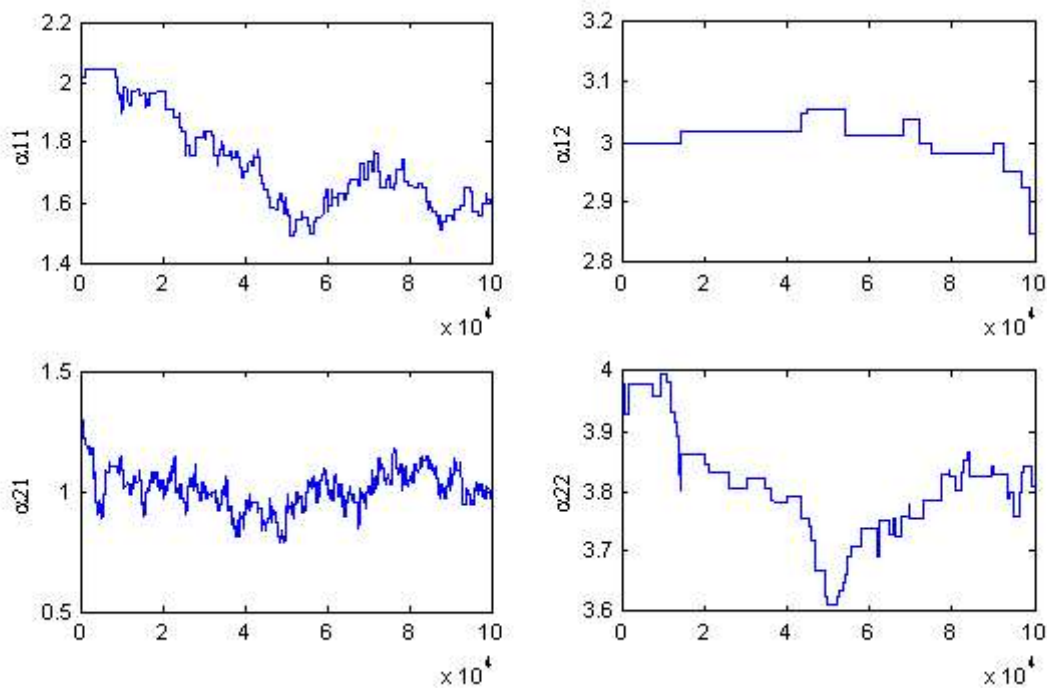




Figure 4.5: Path of parameter estimates for  $\alpha_{ij}$ , ( $i, j = 1, 2$ )



- estimator for  $\sigma_{l1}$ ,  $\sigma_{l1} = 0.66$ , with an actual value of 0.6 estimator for  $\sigma_{l2}$ ,  $\sigma_{l2} = 0.83$ , with an actual value of 0.9

The same exercise is repeated, but now the parameters for the seasonality function  $c(t)$  are inferred as well from the simulated forward data. The number of runs is increased to 200,000.

2. With varying seasonality function  $c(t)$

Now, the seasonality parameters  $c$ ,  $d$  and  $f$  are included in the parameter set  $\theta$

- Number of states = 2

- Transition probability matrix  $\xi$

Simulated data				
	Actual		Estimated	
	1	2	1	2
1	0.8	0.2	0.79	0.21
2	0.6	0.4	0.62	0.38

- Regime-switching parameters  $\sigma_{1j}$  and  $\alpha_{1j}$ , ( $j = 1, 2$ )

Simulated data				
	Actual		Estimated	
$j$	1	2	1	2
$\sigma_{1j}$	0.3	0.8	0.26	0.73
$\sigma_{2j}$	0.6	1.2	0.53	1.19
$\alpha_{1j}$	1.5	2	1.83	1.61
$\alpha_{2j}$	1	3	0.94	3.30

Figures 4.4 and 4.5 below show the parameter estimates throughout the runs. Convergence is reached fast for the volatility parameters  $\sigma_{ij}$ , ( $i, j = 1, 2$ ). The values for the attenuation parameters  $\alpha_{ij}$  are near the actual values of the underlying forward dynamics  $\mathbf{y}_T$ . One could increase the number of runs.

Figure 4.6: Path of parameter estimates for  $\sigma_{ij}$ , ( $i, j = 1, 2$ )

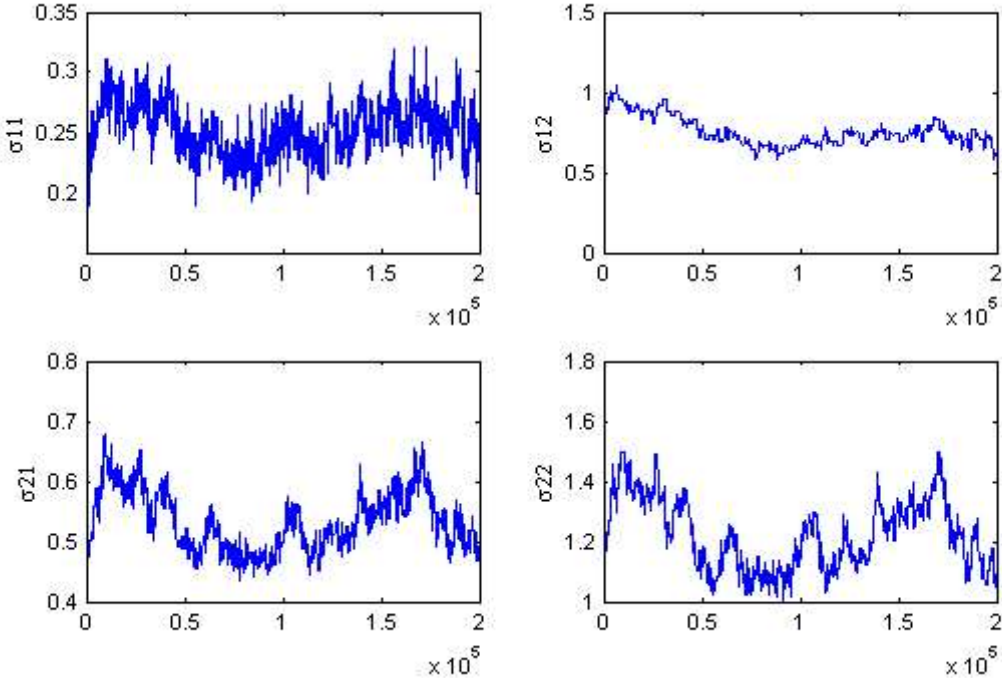
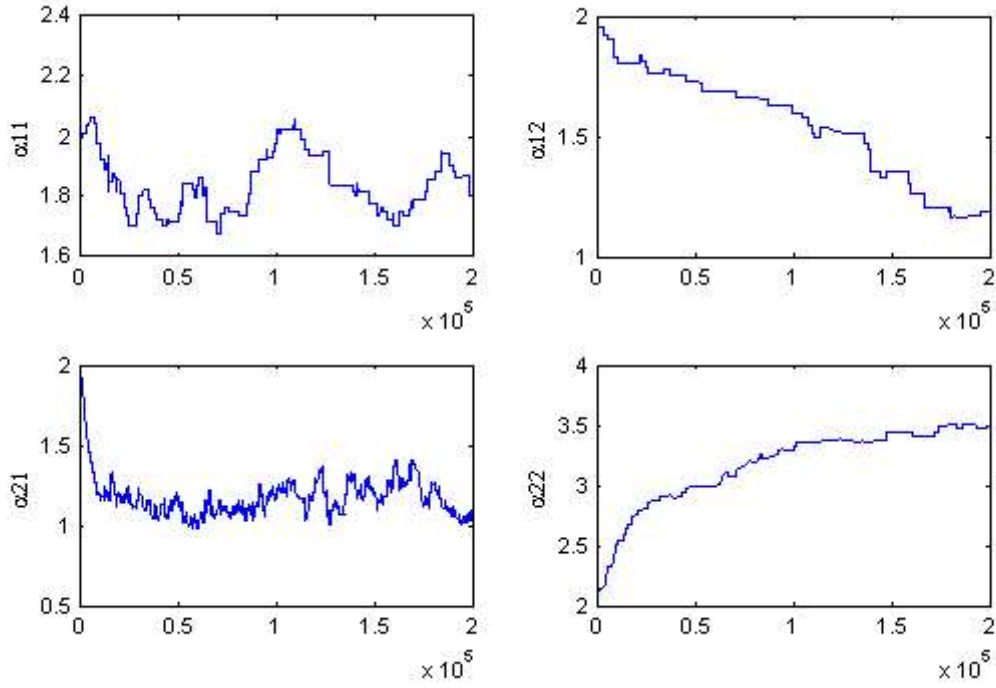


Figure 4.7: Path of parameter estimates for  $\alpha_{ij}$ ,  $(i, j = 1, 2)$



- estimator for  $\sigma_{l1}$ ,  $\sigma_{l1} = 0.53$ , with an actual value of 0.6 estimator for  $\sigma_{l2}$ ,  $\sigma_{l2} = 0.93$ , with an actual value of 0.9
- estimators for seasonality function are respectively,  $c = 0.92$ ,  $d = 0.04$  and  $f = 0.35$  with actual values of  $c = 0.8$ ,  $d = 0.08$  and  $f = 0.31$

### 4.4.3 Historical NBP forward data

Now all pieces have been put together and tested. Finally, the Gibbs sampler for the two-factor regime switching model is applied to historical forward NBP data. The data sample was introduced in Section 4.1. In Appendix E (p.) one can find the routine for the Gibbs sampler which is applied to the sample.

A number of MCMC estimation procedures were run with different initial values for the respective parameters, with rather similar results. The burnin rate is 20%. About 200,000 steps were run and the following are the estimates of the parameters. The estimates for the transition probability matrix  $\xi$  show that with high probability the Markov Chain will stay in regime 1 for the better part of the time horizon. Sporadically, a jump from regime 1 to regime 2 is witnessed.

- Number of states = 2
- Transition probability matrix  $\xi$

Historical data		
	Estimated	
	1	2
1	0.8185	0.185
2	0.4327	0.5673

- Regime-switching parameters  $\sigma_{1j}$  and  $\alpha_{1j}$ , ( $j = 1, 2$ )

Historical data		
	Estimated	
$j$	1	2
$\sigma_{1j}$	0.3777	0.6207
$\sigma_{2j}$	0.5020	1.0885
$\alpha_{1j}$	1.2080	2.0948
$\alpha_{2j}$	1.4115	2.9067

- estimator for  $\sigma_{l1}$  is 0.3654 estimator for  $\sigma_{l2}$ , is 0.7833

- estimators for seasonality function are respectively,  $c = 0.6922$ ,  $d = 0.3362$  and  $f = 0.2791$

These estimates can be compared to the estimates provided by Chiarella *et al.* [4] for NBP forward price data that spanned from 29/09/2006 to 19/02/2008. One could see that less time is spent in regime 2 here, which indicates that the environment is less volatile. It could be argued that volatility levels for natural gas have exceeded normal levels in 2008 as a result of the global financial crisis.

- Number of states = 2
- Transition probability matrix  $\xi$

Historical data		
	Estimated	
	1	2
1	0.8516	0.1484
2	0.7080	0.2920

- Regime-switching parameters  $\sigma_{1j}$  and  $\alpha_{1j}$ , ( $j = 1, 2$ )

Historical data		
	Estimated	
$j$	1	2
$\sigma_{1j}$	0.3057	0.8429
$\sigma_{2j}$	0.4762	1.0292
$\alpha_{1j}$	2.0464	1.8932
$\alpha_{2j}$	1.1533	3.2536

- estimator for  $\sigma_{l1}$  is 0.4869 estimator for  $\sigma_{l2}$ , is 0.6203
- estimators for seasonality function are respectively,  $c = 0.8121$ ,  $d = 0.0781$  and  $f = 0.3070$

# Chapter 5

## Conclusion

In this essay a two-factor regime switching model for the forward price curve in the natural gas market suggested by Chiarella *et al.* [4] was presented. For parameter estimation a Markov Chain Monte Carlo (MCMC) technique was applied, introduced first by Frühwirth-Schnatter (Frühwirth-Schnatter [14]) and Hahn *et al.* [18]. Episodes of higher volatility are followed by episodes of lower volatility driven by market structure changes or supply-demand imbalances. The respective changes in regime affect the whole forward curve structure.

Both the transition probabilities of these switches and the model parameters were inferred from natural gas forward price data, which included 12 forward curve series. The use of MCMC techniques makes the calibration of more complex volatility functions possible. The parameter estimates for the NBP sample were found considerably higher than the estimators found by Chiarella *et al.* [4]. The data in this essay spans from 23/10/2007 to 12/03/2009, whereas Chiarella *et al.* [4] presented data from 29/09/2006 to 19/02/2008. The highly volatile remainder of 2008 and also the beginning of 2009 was thus not included.

Finally, as mentioned by Chiarella *et al.* [4] the Gibbs sampler presented is an off line approach. When new data becomes available, the routine is slow to update the estimates. It is thus not suited in a live setting. Therefore, it would be recommended to implement an on line approach, like a particle filter, such that the volatility functions can be updated quickly from high-frequency data.

# Appendix A

## nbp\_1fac.m

### one-factor volatility model, nbp\_1fac.m

Generates forward dynamics of a one-factor regime switching model

```
% parameters
%
dt=1/365; % daily time-stepping
tT=(3:-dt:0)'; % time-to-maturity
t=flipud(tT); % time

p0=[.9 .1 .7 .3]; % transition probability matrix
th0=[.5 1 1.5 1 .5]; % initial values for both states 1,2
% volatility parameters:
% \sigma_11, \sigma_12
% attenuation parameters:
% \alpha_11, \alpha_12
% long-term volatility parameter
% \sigma_1_1

c=.8; d=.08; f=.31; % seasonality parameters
ct=c+d*(1+sin(f+2*pi.*t)); % seasonality function

% sample state vector x0
%
len=size(tT,1);
x0=ones(len,1); % state 1
x0(1)=unidrnd(2);
for it=2:len;
```



```

        if rand>p0(1,2*x0(it-1)-1)
            x0(it)=2;
        end
    end
    x=x0;

% volatility function
%
x0=[x0==1 x0==2];
sig=sum(x0.*th0(ones(1,len),1:2),2).*...
    (exp(-sum(x0.*th0(ones(1,len),3:4),2).*tT)*(1-th0(5))+th0(5)));
sig=ct.*sig;

% simulate path of forward returns
%
y=-0.5*dt*(sig.*sig)+sig.*randn(len,1)*sqrt(dt);

% plot
%
figure, plot(y)
title('forward data sample, one-factor model')
xlabel('time t'), ylabel('y(t,T)')

```

# Appendix B

## mcmc.m

### Gibbs sampler, mcmc.m

Parameter estimation of Markov switching model with MCMC

```
clc, clear all

% initialize
%
iter=1e4;
burnin=iter/5;

p=zeros(iter,4);           % allocate
th=zeros(iter,5);

% generate simulated data
%
nbp_1fac;                 % call nbp_1fac.m for data

% priors
%
p(1,:)=[.64 .36 .16 .84]; % guess transition probabilities
g=p(1,)*15;              % prior expectation of
                          % probability matrix

th(1,:)=[.5 1 1.5 1 .5]; % initial parameter values
rg=[zeros(1,5);1.5 1.5 5 5 1 ]; % bounds for parameter values

rth=0.05;                % step size parameter updates
```

```

% sample initial state vector x
%
x=ones(len,1);           % state 1
x(1)=unidrnd(2);
for it=2:len;
    if rand>p(1,2*x(it-1)-1)
        x(it)=2;
    end
end

% algorithm
%
th1=th(1,:);
for k=2:iter
    sp=zeros(len,2);      % allocate
    sf=zeros(len,2);
    sc=zeros(len,2);

    % sample transition probability matrix p
    %
    n12=sum(diff(x)==1);   % count transitions between states
    n21=sum(diff(x)==-1);
    n11=sum((x(1:end-1)==1)+(x(2:end)==1)==2);
    n22=sum((x(1:end-1)==2)+(x(2:end)==2)==2);

    % sample from Beta distribution
    p1=betarnd([g(1,1)+n11 g(1,3)+n21],[g(1,2)+n12 g(1,4)+n22]);
    p1=[p1(1) 1-p1(1) p1(2) 1-p1(2)];
    p(k,:)=p1;            % update

    % Metropolis-Hastings algorithm
    %
    idx=unidrnd(5);       % random parameter choice
    % generate

    th1(idx)=th(k-1,idx) + rth*randn;
    st=[x==1 x==2];

    % complete-data likelihood

```

```

                                % call mclh.m
llh=sum(log(mclh(y,st,th1,tT,dt,t,len))...
        -log(mclh(y,st,th(k-1,:),tT,dt,t,len)));

                                % prior densities
pr=unifpdf([th1(idx) th(k-1,idx)],rg(1,idx),rg(2,idx));

                                % acceptance probability
alpha=min(1,llh*pr(1)/pr(2));
if rand<=alpha,th(k,:)=th1;
else th(k,:)=th(k-1,:);
end

% filtering, FFBS algorithm
%
% forward filtering
%
sp(1,:)=0.5*(p1(1:2)+p1(3:end));
sf(1,:)=sp(1,:);
sf(2:end,1)=mclh(y(2:end),[ones(len-1,1) zeros(len-1,1)],...
                th(k,:),tT(2:end),dt,t(2:end),len-1);
sf(2:end,2)=mclh(y(2:end),[zeros(len-1,1) ones(len-1,1)],...
                th(k,:),tT(2:end),dt,t(2:end),len-1);

for it=2:len
    sf(it,:)=sf(it,:).*sp(it-1,:);
    sf(it,:)=sf(it,:)/(sf(it,1)+sf(it,2));
    sp(it,:)=sf(it,1)*p1(1:2)+sf(it,2)*p1(3:end);
end

% backward sampling
%
x=ones(len,1);          % state 1
if rand>sf(end,1)
    x(len)=2;
end

for it=len-1:-1:1
    sc(it,:)=p1([x(it+1) x(it+1)+2]).*sf(it,:);
    sc(it,:)=sc(it,:)/(sc(it,1)+sc(it,2));
end

```

```

        if rand>sc(it,1)
            x(it)=2;
        end
    end
    th1=th(k,:);          % update
end

% output
%
mu_th=mean(th(burnin+1:end,:))
mu_p=mean(p(burnin+1:end,:))

```

In the routine for the Gibbs sampler, the function **mclh.m** is called to evaluate the complete-data likelihood for various values of the parameters. This function is given below.

**function [lh,sig]=mclh(y,x,th,tT,dt,t,T), mclh.m**

Returns components of complete-data likelihood function

```

sig=zeros(T,2);
ct=.8+.08*(1+sin(.31+2*pi.*t));
sig(:,1)=sum(x.*th(ones(1,T),1:2),2).*(th(5)+
            exp(-sum(x.*th(ones(1,T),3:4),2).*tT)*(1-th(5))));
sig=ct(:,ones(2,1)).*sig;
mu=dt*sum(sig.*sig,2);

lh=exp(-0.5*((y+0.5*mu).^2)./mu);
lh=lh./sqrt(mu);

```

# Appendix C

## nbp\_2fac.m

### two-factor volatility model, nbp\_2fac.m

Generates forward dynamics of a two-factor regime switching model

```
% parameters
%
dt=1/365;           % daily
tT=(3:-dt:0)';    % time-to-maturity
t=flipud(tT);     % time

p0=[.75 .25 .85 .15]; % transition probability matrix
th0=[.5  1   1.5 1 .6 ...
      .75 1.3 2   3 .8]; % initial values for 2 factors &
                        % both states 1,2
                        % volatility parameters:
                        % \sigma_11, \sigma_12
                        % attenuation parameters:
                        % \alpha_11, \alpha_12
                        % long-term volatility parameter
                        % \sigma_l_1

                        % volatility parameters:
                        % \sigma_21, \sigma_22
                        % attenuation parameters:
                        % \alpha_21, \alpha_22
                        % long-term volatility parameter
                        % \sigma_l_2
```

```

c=.8; d=.08; f=.31;          % seasonality parameters
ct=c+d*(1+sin(f+2*pi.*t)); % seasonality function

% sample state vector x0
%
len=size(tT,1);
x0=ones(len,1);          % state 1
x0(1)=unidrnd(2);
for it=2:len;
    if rand>p0(1,2*x0(it-1)-1)
        x0(it)=2;
    end
end
x=x0;

% volatility function
%
x0=[x0==1 x0==2];
sig(:,1)=sum(x0.*th0(ones(1,len),1:2),2).*...
    (exp(-sum(x0.*th0(ones(1,len),3:4),2).*tT)*(1-th0(5))+th0(5)));
sig(:,2)=sum(x0.*th0(ones(1,len),6:7),2).*...
    (th0(10)-exp(-sum(x0.*th0(ones(1,len),8:9),2).*tT));
sig=ct(:,ones(2,1)).*sig;

% simulate path of forward returns
%
y=-0.5*dt*sum(sig.*sig,2)+sum(sig.*randn(len,2)*sqrt(dt),2);

% plot
%
figure, plot(y)
title('forward data sample, two-factor model')
xlabel('time t'), ylabel('y(t,T)')

```

# Appendix D

## mcmc.m

### Gibbs sampler, mcmc.m

Parameter estimation of Markov switching model with MCMC

```
clc, clear all

% initialize
%
iter=1e5;
burnin=iter/5;

p=zeros(iter,4);           % allocate
th=zeros(iter,10);

% generate simulated data
%
nbp_2fac;                 % call nbp_2fac.m for data

% priors
%
p(1,:)=[.64 .36 .16 .84]; % guess transition probabilities
g=p(1, :)*15;            % prior expectation of
                        % probability matrix

th(1,:)=[.25 .75 2 2 .5 ...
        .5 1 2 2 1];      % initial parameter values
rg=[zeros(1,10);1.5 1.5 5 5 1 ...
    1.5 1.5 5 5 1];      % bounds for parameter values
```



```

rth=0.025; % step size parameter updates

% sample initial state vector x
%
x=ones(len,1); % state 1
x(1)=unidrnd(2);
for it=2:len;
    if rand>p(1,2*x(it-1)-1)
        x(it)=2;
    end
end

% algorithm
%
th1=th(1,:);
for k=2:iter
    sp=zeros(len,2); % allocate
    sf=zeros(len,2);
    sc=zeros(len,2);

    % sample transition probability matrix p
    %
    n12=sum(diff(x)==1); % count transitions between states
    n21=sum(diff(x)==-1);
    n11=sum((x(1:end-1)==1)+(x(2:end)==1)==2);
    n22=sum((x(1:end-1)==2)+(x(2:end)==2)==2);

    % sample from Beta distribution
    p1=betarnd([g(1,1)+n11 g(1,3)+n21],[g(1,2)+n12 g(1,4)+n22]);
    p1=[p1(1) 1-p1(1) p1(2) 1-p1(2)];
    p(k,:)=p1; % update

% Metropolis-Hastings algorithm
%
    idx=unidrnd(10); % random parameter choice
    % generate

    th1(idx)=th(k-1,idx) + rth*randn;
    st=[x==1 x==2];

```

```

                                % complete-data likelihood
                                % call mclh.m
llh=sum(log(mclh(y,st,th1,tT,dt,t,len))...
        -log(mclh(y,st,th(k-1,:),tT,dt,t,len)));

                                % prior densities
pr=unifpdf([th1(idx) th(k-1,idx)],rg(1,idx),rg(2,idx));

                                % acceptance probability
alpha=min(1,llh*pr(1)/pr(2));
if rand<=alpha,th(k,:)=th1;
else th(k,:)=th(k-1,:);
end

% filtering, FFBS algorithm
%
% forward filtering
%
sp(1,:)=0.5*(p1(1:2)+p1(3:end));
sf(1,:)=sp(1,:);
sf(2:end,1)=mclh(y(2:end),[ones(len-1,1) zeros(len-1,1)],...
                th(k,:),tT(2:end),dt,t(2:end),len-1);
sf(2:end,2)=mclh(y(2:end),[zeros(len-1,1) ones(len-1,1)],...
                th(k,:),tT(2:end),dt,t(2:end),len-1);

for it=2:len
    sf(it,:)=sf(it,:).*sp(it-1,:);
    sf(it,:)=sf(it,:)/(sf(it,1)+sf(it,2));
    sp(it,:)=sf(it,1)*p1(1:2)+sf(it,2)*p1(3:end);
end

% backward sampling
%
x=ones(len,1);           % state 1
if rand>sf(end,1)
    x(len)=2;
end

for it=len-1:-1:1

```

```

        sc(it,:)=p1([x(it+1) x(it+1)+2]).*sf(it,:);
        sc(it,:)=sc(it,:)/(sc(it,1)+sc(it,2));
        if rand>sc(it,1)
            x(it)=2;
        end
    end
    th1=th(k,:);          % update
end

% output
%
mu_th=mean(th(burnin+1:end,:))
mu_p=mean(p(burnin+1:end,:))

```

In the routine for the Gibbs sampler, the function **mclh.m** is called to evaluate the complete-data likelihood for various values of the parameters. This function is given below.

**function [lh,sig]=mclh(y,x,th,tT,dt,t,T), mclh.m**

Returns components of complete-data likelihood function

```

sig=zeros(T,2);
ct=.8+.08*(1+sin(.31+2*pi.*t));
sig(:,1)=sum(x.*th(ones(1,T),1:2),2).*(th(5)+
    exp(-sum(x.*th(ones(1,T),3:4),2).*tT)*(1-th(5)));
sig(:,2)=sum(x.*th(ones(1,T),6:7),2).*(th(10)-
    exp(-sum(x.*th(ones(1,T),8:9),2).*tT));
sig=ct(:,ones(2,1)).*sig;
mu=dt*sum(sig.*sig,2);

lh=exp(-0.5*((y+0.5*mu).^2)./mu);
lh=lh./sqrt(mu);

```

# Appendix E

## mcmc.m

### Gibbs sampler, mcmc.m

Parameter estimation of Markov switching model with MCMC

```
clc, clear all

% initialize
%
iter=2*1e2;
burnin=iter/5;

p=zeros(iter,4);           % allocate
th=zeros(iter,13);

% generate simulated data
%
load f, load tT
y=log(f(2:end,1:12))-log(f(1:end-1,1:12));
tT=tT(1:end-1,1:12);

dt=1/252;% daily

% priors
%
p(1,:)=[.64 .36 .16 .84]; % guess transition probabilities
g=p(1, :)*15;             % prior expectation of
                           % probability matrix
```

```

th(1,:)=[.3 .8 2 2 .5 ...
        .5 1 1 3 .7 .8 .08 .3];          % initial parameter values

rg=[zeros(1,13);1.5 1.5 5 5 1 1.5 1.5 5 5 1 1 1 1]; % bounds for parameter v

rth=0.03;                                % step size parameter updates

% implement 'randomized' algorithm
%

dim=size(y);
lam=unidrnd(dim(2),dim(1),1);
sample = sub2ind(dim, 1:dim(1), lam');% use linear indexing

% construct return & maturity vectors with
% corresponding data points
y=y(sample)';
tT=tT(sample)';
t=flipud(tT);% time

% draw x
len=size(tT,1);
x=ones(len,1);
x(1)=unidrnd(2);
for it=2:len;
    if rand>p(1,2*x(it-1)-1),x(it)=2;end
end

% algorithm
%
th1=th(1,:);
for k=2:iter
    sp=zeros(len,2);          % allocate
    sf=zeros(len,2);
    sc=zeros(len,2);

    % sample transition probability matrix p
    %

```

```

n12=sum(diff(x)==1);      % count transitions between states
n21=sum(diff(x)==-1);
n11=sum((x(1:end-1)==1)+(x(2:end)==1)==2);
n22=sum((x(1:end-1)==2)+(x(2:end)==2)==2);

                                % sample from Beta distribution
p1=betarnd([g(1,1)+n11 g(1,3)+n21],[g(1,2)+n12 g(1,4)+n22]);
p1=[p1(1) 1-p1(1) p1(2) 1-p1(2)];
p(k,:)=p1;                    % update

% Metropolis-Hastings algorithm
%
    idx=unidrnd(13);           % random parameter choice
                                % generate

    th1(idx)=th(k-1,idx) + rth*randn;
    st=[x==1 x==2];

                                % complete-data likelihood
                                % call mclh.m
    llh=sum(log(mclh2(y,st,th1,tT,dt,t,len))...
            -log(mclh2(y,st,th(k-1,:),tT,dt,t,len)));

                                % prior densities
    pr=unifpdf([th1(idx) th(k-1,idx)],rg(1,idx),rg(2,idx));

                                % acceptance probability
    alpha=min(1,llh*pr(1)/pr(2));
    if rand<=alpha,th(k,:)=th1;
    else th(k,:)=th(k-1,:);
    end

% filtering, FFBS algorithm
%
    % forward filtering
    %
    sp(1,:)=0.5*(p1(1:2)+p1(3:end));
    sf(1,:)=sp(1,:);
    sf(2:end,1)=mclh2(y(2:end),[ones(len-1,1) zeros(len-1,1)],...
        th(k,:),tT(2:end),dt,t(2:end),len-1);

```

```

sf(2:end,2)=mclh2(y(2:end),[zeros(len-1,1) ones(len-1,1)],...
    th(k,:),tT(2:end),dt,t(2:end),len-1);

for it=2:len
    sf(it,:)=sf(it,:).*sp(it-1,:);
    sf(it,:)=sf(it,:)/(sf(it,1)+sf(it,2));
    sp(it,:)=sf(it,1)*p1(1:2)+sf(it,2)*p1(3:end);
end

% backward sampling
%
x=ones(len,1);           % state 1
if rand>sf(end,1)
    x(len)=2;
end

for it=len-1:-1:1
    sc(it,:)=p1([x(it+1) x(it+1)+2]).*sf(it,:);
    sc(it,:)=sc(it,:)/(sc(it,1)+sc(it,2));
    if rand>sc(it,1)
        x(it)=2;
    end
end
th1=th(k,:);           % update

% debug
%th(k,6)
end

% output
%
mu_th=mean(th(burnin+1:end,:))
mu_p=mean(p(burnin+1:end,:))

% plot
%

% volatility parameters
subplot(2,2,1), plot(th(:,1))

```

```

subplot(2,2,2), plot(th(:,2))
subplot(2,2,3), plot(th(:,6))
subplot(2,2,4), plot(th(:,7))

% attenuation parameters
figure
subplot(2,2,1), plot(th(:,3))
subplot(2,2,2), plot(th(:,4))
subplot(2,2,3), plot(th(:,8))
subplot(2,2,4), plot(th(:,9))

```

In the routine for the Gibbs sampler, the function **mclh2.m** is called to evaluate the complete-data likelihood for various values of the parameters. This function is given below.

```
function [lh,sig]=mclh2(y,x,th,tT,dt,t,T)
```

Returns components of complete-data likelihood function

```

sig=zeros(T,2);
%ct=.8+.08*(1+sin(.31+2*pi.*t));

ct=th(11)+th(12)*(1+sin(th(13)+2*pi.*t));

sig(:,1)=sum(x.*th(ones(1,T),1:2),2).*(th(5)+exp(-sum(x.*th(ones(1,T),3:4),2).*t));
sig(:,2)=sum(x.*th(ones(1,T),6:7),2).*(th(10)-exp(-sum(x.*th(ones(1,T),8:9),2).*t));
sig=ct(:,ones(2,1)).*sig;
mu=dt*sum(sig.*sig,2);

lh=exp(-0.5*((y+0.5*mu).^2)./mu);
lh=lh./sqrt(mu);

```



# Bibliography

- [1] Benth, F. & Koekebakker, S. (2008), “Stochastic modeling of financial electricity contracts”, *Energy Economics* 30, 1116-1157.
- [2] Breslin, J., Clewlow, L., Kwok, C. & Strickland, C. (2008), “Gaining from complexity: MFMC models”, *Energy Risk*, April, 60-64.
- [3] Carlin, B. & Polson, N. (1992), “Monte Carlo Bayesian methods for discrete regression models and categorical time series”, *Bayesian Statistics* 4, 577-586.
- [4] Chiarella, C., Clewlow, L. & Kang, B. (2009) “Modelling and estimating the forward price curve in the energy market”, *Working Paper*.
- [5] Chib, S. (1995), “Marginal likelihood from the Gibbs output”, *Journal of the American Statistical Association* 90, 1313-1321.
- [6] Chib, S. (1996), “Calculating posterior distributions and modal estimates in Markov Mixture models”, *Journal of Econometrics* 75, 79-97.
- [7] Clewlow, L. & Strickland, C. (1999), “A multi-factor model for energy derivatives”, *QFRC Research Paper Series 28, University of Technology, Sydney*.
- [8] Davis, M. (1984), “Piecewise-deterministic Markov processes: a general class of non-diffusion stochastic models”, *Journal of Royal Statistical Society B*, 46, 353–88.
- [9] de Jong, C. & R. Huisman (2002), “Option formulas for mean-reverting power prices with spikes.”, *Working Paper, Rotterdam School of Management at Erasmus University*.
- [10] Deng, S. (2000), “Pricing electricity derivatives under alternative stochastic spot price models.”, *Proceedings of the 33rd Hawaii International Conference on System Sciences*.

- [11] Di Masi, G.B., Kabanov, Yu. M. & Runggaldier W.J. (1994), “Mean-variance hedging of options on stocks with Markov volatilities”, *Theory of Probability and its Applications* 39, 172-182.
- [12] Elliott, R. Aggoun, L. & Moore, J. (1995), “Hidden Markov Models: estimation and control, applications of mathematics”, Springer-Verlag, New York.
- [13] Elliott, R. & Wilson, C. (2001), “The term structure of interest rates when a Markov chain is driving drift and volatility parameters of the short rate diffusion process”, *Working Paper*.
- [14] Frühwirth-Schnatter, S. (2006), “Finite Mixture and Markov Switching Models”, Springer, New York.
- [15] Gibson, R. & Schwartz, E.S. (1990), “Stochastic convenience yield and the pricing of oil contingent claims, ” *The Journal of Finance* 45, 959-976.
- [16] Gilks, W.R., Richardson, S. & Spiegelhalter D.J. (1998), “Markov Chain Monte Carlo in Practice”, Chapman & Hall, London.
- [17] Goldfeld, S.M. & Quandt, R.E. (1973), “A Markov model for switching regressions”, *Journal of Econometrics* 1, 3-15.
- [18] Hahn, M., Frühwirth-Schnatter, S. & Sass, J. (2007), “Markov Chain Monte Carlo methods for parameter estimation in multidimensional continuous time Markov switching models”, *RICAM, Linz* 2007-09.
- [19] Hamilton, J.D. (1989), “A New approach to the economic analysis of nonstationary time series and the business cycle”, *Econometrica* 57, 357-384.
- [20] Hamilton, J.D. (1990), “Analysis of time series subject to changes in regime”, *Journal of Econometrics* 45, 39-70.
- [21] Hansen, A.T. & Poulsen, R. (2000), “A simple regime switching term structure model”, *Finance and Stochastics* 4, 409-29.
- [22] Heath, D., Jarrow, R. & Morton, A. (1992), “Bond pricing and the term structure of interest rates: a new methodology”, *Econometrica* 60, 77–105.
- [23] Johannes, M. & Polson, N. (2002). “MCMC methods for continuous-time financial econometrics”, *Handbook of Financial Econometrics*.
- [24] Naik, V. (1993), “Option valuation and hedging strategies with jumps in the volatility of asset returns”, *Journal of Finance* 48, 1969-1984.

- [25] Pilipovic, D. (2007), “Energy risk: valuing and managing energy derivatives”, Second edition, McGraw-Hill.
- [26] Rambharat, B.R., Brockwel A.E. & Seppi, D.J. (2005), “A threshold autoregressive model for wholesale electricity prices”, *Journal of the Royal Statistical Society C*, 54, 287-299.
- [27] Scott, S. (2002), “Bayesian methods for Hidden Markov models”, *The Journal of the American Statistical Association* 97, 337-351.
- [28] Seifert, J. & Uhrig-Homburg, M. (2007), “Modelling jumps in electricity prices: theory and empirical evidence”, *Review of Derivatives Research* 10, 59-85.
- [29] Valchev, S. (2004), “Stochastic volatility Gaussian Heath-Jarrow-Morton models”, *Applied Mathematical Finance* 11, 347-368.