The Journal of Energy Markets (1–23) Volume 3/Number 3, Fall 2010

Modeling electricity forward prices using the multivariate normal inverse Gaussian distribution

Arne Andresen
Norwegian University of Science and Technology,
Department of Industrial Economics and Technology Management,
NO-7491 Trondheim, Norway; email: arne.andresen@iot.ntnu.no

Steen Koekebakker
University of Agder, Faculty of Economics and Social Sciences,
Service Box 422, NO-4604 Kristiansand, Norway;
email: steen.koekebakker@uia.no

Sjur Westgaard
Norwegian University of Science and Technology,
Department of Industrial Economics and Technology Management,
NO-7491 Trondheim, Norway; email: sjur.westgaard@iot.ntnu.no

This paper presents a discrete random-field model for forward prices driven by the multivariate normal inverse Gaussian distribution. The model captures the idiosyncratic risk and adequately addresses the heavy tails characterizing electricity forward prices. We fit the model to forward prices from the Nordic power exchange using a Markov chain Monte Carlo algorithm. This is then compared with Gaussian-based multifactor models in terms of goodness of fit to historical log returns. Our finding is that the proposed model offers a superior fit to the empirical distributions.

1 INTRODUCTION

In volatile markets like the electricity market, the return distribution of forward prices is leptokurtic, ie, it has a high center peak and heavy tails (see, for example, Frestad *et al* (2010) or Benth and Koekebakker (2008) for discussions). This suggests that

We thank Stein-Erik Fleten, Ane Marte Heggedal, Alexander Walnum, the guest editor (Fred Espen Benth) and an anonymous referee for helpful comments. Arne Andresen and Sjur Westgaard also acknowledge financial support from the Research Council of Norway, Agder Energi and Troms Kraft.

Gaussian-based models will have shortcomings in representing the risk related to the substantial market movements observed in these markets. Since risk measures and option prices depend on the distribution used in a model, it is important to develop models that are able to handle the non-Gaussian nature of electricity markets.

There are two main approaches to modeling forward prices: one is to specify a stochastic model for the spot price, and from this model derive the dynamics of forward prices based on no-arbitrage principles. The alternative is to follow the Heath–Jarrow–Morton approach and specify the dynamics of the forward contracts directly.

Most of the existing literature focuses on developing models for the spot price. Some examples are Lucia and Schwartz (2002), Cartea and Figueroa (2005) and Benth *et al* (2007a). The Heath–Jarrow–Morton approach to modeling forward prices in the electricity market has been proposed by several authors: Benth and Koekebakker (2008), Bjerksund *et al* (2000), Clewlow and Strickland (2000) and Kiesel *et al* (2009) all model the forward prices using multifactor models driven by Brownian motion.

Empirical findings in Koekebakker and Ollmar (2005) raise doubts about the validity of low-dimensional multifactor models in electricity markets, since a substantial amount of variation in forward prices cannot be explained by a few common factors. This observation is pursued in Frestad (2008) using the framework of Ross (1976), allowing common and unique factors to influence forward prices. The idea of letting each forward contract have some unique risk was first proposed by Audet *et al* (2004). They model weekly contracts at Nord Pool (the Nordic power exchange) using a Gaussian random-field model with an exponential correlation function. However, the models that directly specify the dynamics of the forward contracts ignore the fact that the return of forward prices in electricity markets is far from Gaussian distributed.

The benefit of modeling the forward prices directly is that, in contrast to spot models, there is no problem fitting the model to the current forward prices. In addition, since electricity is largely non-storable, there is no cost-of-carry relationship linking spot and forward prices. For example, Quinn *et al* (2005) examine spot and forward data from the PJM[1] market and find that electricity forward prices are largely disconnected from current spot prices. Benth *et al* (2008) draw similar conclusions when examining spot and forward prices at Nord Pool. They find that only the contracts in the very short end of the forward structure are highly correlated with the spot price. Their conclusion is that different dynamics drive the spot and forward prices. Inferring the dynamics of forward contracts from spot price models is therefore problematic.

Given the well-known shortcomings of the Gaussian distribution with regard to modeling the return distribution of financial assets, it is natural to look for other dis-

[1] For the period under examination the PJM electricity market covered Pennsylvania, New Jersey, Maryland, Delaware and the District of Columbia.

tributions that are better capable of capturing heavy tails and skewness. Frestad *et al* (2010) analyze the distribution of daily log returns of individual forward contracts at Nord Pool and find that the univariate normal inverse Gaussian (NIG) distribution captures the stylized facts of the returns. In this paper we develop a model for the joint dynamics of forward prices based on the semi-heavy-tailed multivariate normal inverse Gaussian (MNIG) distribution. The modeling framework that we use is a discrete random-field model, implying a blending of the market model setup of Jamshidian (1997), Brace *et al* (1997) and others, and the random-field modeling approach introduced by Kennedy (1994). In this approach, each forward contract is a distinct random variable that may be correlated with the other contracts. In general, each forward contract is therefore exposed to idiosyncratic risk, and cannot be perfectly hedged by a portfolio of forward contracts with different maturities.

This paper is organized as follows. In the next section we present our model and review the definition of the MNIG distribution and some of its properties. In Section 3 we investigate the empirical characteristics of forward prices at Nord Pool, choose suitable functional forms of the correlation, skewness and volatility structure in our model, and fit the model to forward prices. We investigate the quality of our model with emphasis on the fit to observed log returns and covariance structure in Section 4. In Section 5 we conclude.

2 A DISCRETE RANDOM-FIELD MODEL DRIVEN BY THE MULTIVARIATE NORMAL INVERSE GAUSSIAN DISTRIBUTION

In this section we describe the MNIG distribution and some of its properties before presenting the model. We will also make use of the NIG distribution in this paper. For the definition and properties of this distribution we refer the reader to Barndorff-Nielsen (1998).

2.1 The MNIG distribution

A d-dimensional vector $\boldsymbol{X}$ is MNIG distributed if its probability density function reads:

$$f(\boldsymbol{X}) = \frac{\delta}{2^{(d-1)/2}} \left[\frac{\alpha}{\pi q(\boldsymbol{x})} \right]^{(d+1)/2} \times \exp[p(\boldsymbol{x})] K_{(d+1)/2}[\alpha q(\boldsymbol{x})] \tag{1}$$

where:

$$q(\boldsymbol{x}) = \sqrt{\delta^2 + (\boldsymbol{x} - \boldsymbol{\mu})' \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\mu})} \quad \text{and} \quad p(\boldsymbol{x}) = \delta \sqrt{\alpha^2 - \boldsymbol{\beta}' \Sigma \boldsymbol{\beta}} + \boldsymbol{\beta}' (\boldsymbol{x} - \boldsymbol{\mu})$$

$K_d(x)$ is the modified Bessel function of the second kind with index $d, \delta > 0$, $\alpha^2 > \boldsymbol{\beta}' \Sigma \boldsymbol{\beta}$, $\boldsymbol{\beta} \in \mathbb{R}^d$, $\boldsymbol{\mu} \in \mathbb{R}^d$ and $\Sigma \in \mathbb{R}^{d \times d}$ is a symmetric positive semidefinite matrix with determinant 1.

The parameters in the MNIG distribution have interpretations relating to the overall shape of the density as follows. The parameters α and $\boldsymbol{\beta}$ control the shape of the density, δ is a scale parameter and $\boldsymbol{\mu}$ is the location parameter as such when $\boldsymbol{\beta} = \mathbf{0}$, $\boldsymbol{\mu}$ denotes the mean of the distribution. Finally, the structure matrix Σ determines the overall shape of the covariance structure between the components of $\boldsymbol{X}$. We will refer to $\boldsymbol{\beta}$ as the skewness parameter even though the skewness of the marginal components also depends on α, δ and Σ.

The cumulant generating function of a d-dimensional MNIG-distributed variable $\boldsymbol{X}$ is given by:

$$\Psi(\boldsymbol{\omega}) = \mathrm{i}\boldsymbol{\omega}'\boldsymbol{\mu} + \delta\left(\sqrt{\alpha^2 - \boldsymbol{\beta}'\Sigma\boldsymbol{\beta}} - \sqrt{\alpha^2 - (\boldsymbol{\beta} + \mathrm{i}\boldsymbol{\omega})'\Sigma(\boldsymbol{\beta} + \mathrm{i}\boldsymbol{\omega})}\right)$$

where $\mathrm{i} = \sqrt{-1}$. From the cumulant generating function we can easily derive the mean vector of an MNIG-distributed variable $\boldsymbol{X}$:

$$E[\boldsymbol{X}] = \boldsymbol{\mu} + \delta\Sigma\boldsymbol{\beta}(\alpha^2 - \boldsymbol{\beta}'\Sigma\boldsymbol{\beta})^{-1/2}$$

and covariance matrix:

$$\mathrm{var}[\boldsymbol{X}] = \delta(\alpha^2 - \boldsymbol{\beta}'\Sigma\boldsymbol{\beta})^{-1/2}[\Sigma + (\alpha^2 - \boldsymbol{\beta}'\Sigma\boldsymbol{\beta})^{-1}\Sigma\boldsymbol{\beta}\boldsymbol{\beta}'\Sigma] \tag{2}$$

If the MNIG distribution is symmetric, ie, the skewness parameter $\boldsymbol{\beta} = \mathbf{0}$, we see from Equation (2) that the correlation matrix is solely determined by the matrix Σ. For a non-symmetric MNIG distribution, however, the correlation structure depends on the three parameters Σ, $\boldsymbol{\beta}$ and α.

The MNIG distribution is fairly complicated, but the distribution has a simple characterization as a variance–mean mixture of a d-dimensional Gaussian random variable $\boldsymbol{Y}$ with a univariate inverse Gaussian-distributed mixing variable Z (Barndorff-Nielsen (1997)). Hence, an MNIG-distributed random variable $\boldsymbol{X}$ can be constructed from:

$$\boldsymbol{X} = \boldsymbol{\mu} + Z\Sigma\boldsymbol{\beta} + \sqrt{Z}\Sigma^{1/2}\boldsymbol{Y} \tag{3}$$

where $\boldsymbol{Y} \sim N_d(\mathbf{0}, I)$ and $Z \sim \mathrm{IG}[\delta^2, \alpha^2 - \boldsymbol{\beta}'\Sigma\boldsymbol{\beta}]$. Here $\mathrm{IG}[\chi, \psi]$, $\chi, \psi > 0$, denotes the inverse Gaussian distribution. This observation is important because it provides an easy way to simulate MNIG variables when pricing derivatives using the Monte Carlo method.

From Equation (3) we see that $\boldsymbol{X}$ can be interpreted as a stochastic variable where the inverse Gaussian variable Z represents both stochastic volatility and stochastic mean. The term $\sqrt{Z}$ is responsible for altering the tail thickness relative to a Gaussian distribution, while the term Z adds asymmetry to X ie, when modeling return series, $\boldsymbol{X}$ is able to capture the fact that negative returns often have heavier tails than positive returns.

The multivariate Gaussian distribution is a limiting distribution for the MNIG distribution in the limit $\delta \to \infty$ and $\alpha \to \infty$ but such that $\delta/\alpha = \sigma^2$ (Øigård *et al* (2005)). Another important special case for the MNIG distribution is the multivariate t-distribution with one degree of freedom. This occurs when $\Sigma = I$ and $\alpha \to 0$ (Øigård *et al* (2005)).

2.2 A discrete random-field model based on the multivariate normal inverse Gaussian distribution

We consider a discrete-time forward market where trades occur on discrete dates indexed $\mathbb{T} = \{t \mid t = 0, 1, \dots, n\}$. In the market, d different electricity forward contracts are traded. Electricity forward contracts deliver a constant flow of electricity over a specified time interval. Typical delivery periods are daily, weekly, monthly, quarterly or yearly. It is common to encounter contracts having overlapping delivery periods. For instance, one may buy three monthly forwards with delivery in the first three months of a year, or a quarterly contract with delivery over the first quarter. In order to avoid arbitrage, one needs to have certain relations between the prices of these contracts. This means, for example, that the cost of buying a portfolio of four consecutive quarterly forward contracts must equal the cost of buying the yearly forward contract with the same delivery period. We restrict our attention to non-overlapping contracts, thereby avoiding the complications that overlapping contracts lead to. More specifically, we will consider non-overlapping synthetic forward contracts obtained from a smoothed forward curve (see Fleten and Lemming (2003) or Benth *et al* (2007b) for details on how to construct synthetic forward curves from traded swaps). The forward contracts that we will work with are thus obtained by first constructing a smooth forward curve from traded forward contracts, and then computing synthetic forward contracts from the smooth forward curve. In this way we can construct datasets of forward contracts with the preferred delivery structure.

Assume that we have d different forward contracts with disjoint delivery periods $[T_1^{\mathrm{s}}, T_1^{\mathrm{e}}], [T_2^{\mathrm{s}}, T_2^{\mathrm{e}}], \dots, [T_d^{\mathrm{s}}, T_d^{\mathrm{e}}]$, where $[T_{\mathrm{c}}^{\mathrm{s}}, T_{\mathrm{c}}^{\mathrm{e}}]$, $T_{\mathrm{c}}^{\mathrm{s}} < T_{\mathrm{c}}^{\mathrm{e}}$, denotes the delivery period for contract c, $c = 1, \dots, d$. Let $F_{\mathrm{c}}(t)$ denote the price of a forward contract with delivery period $[T_{\mathrm{c}}^{\mathrm{s}}, T_{\mathrm{c}}^{\mathrm{e}}]$ at time t, $0 \leqslant t \leqslant T_{\mathrm{c}}^{\mathrm{s}}$. Under the real-world probability measure P, the one-period return of forward contracts is assumed to have the following dynamics:

$$\ln\left(\frac{\boldsymbol{F}(t+1)}{\boldsymbol{F}(t)}\right) = \boldsymbol{v}(t, \boldsymbol{T}^{\mathrm{s}}) + \Lambda(t, \boldsymbol{T}^{\mathrm{s}})\boldsymbol{L}(t, \boldsymbol{T}^{\mathrm{s}}) \tag{4}$$

Here:

$$0 \leqslant t \leqslant \min\{T_1^{\mathrm{s}}, \dots, T_d^{\mathrm{s}}\}$$
$$\boldsymbol{F}(t) = [F_1(t), F_2(t), \dots, F_d(t)]'$$
$$\boldsymbol{\upsilon}(t, \boldsymbol{T}^{\mathrm{s}}) = [\upsilon(t, T_1^{\mathrm{s}}), \upsilon(t, T_2^{\mathrm{s}}), \dots, \upsilon(t, T_d^{\mathrm{s}})]'$$

where $\upsilon(t, T_{\mathrm{c}}^{\mathrm{s}})$, $c = 1, \dots, d$, are deterministic scalar-valued functions, $\Lambda(t, \boldsymbol{T}^{\mathrm{s}})$ is a diagonal matrix with $\lambda_{\mathrm{c}}(t, T_{\mathrm{c}}^{\mathrm{s}})$ on the main diagonal, where $\lambda_{\mathrm{c}}(t, T_{\mathrm{c}}^{\mathrm{s}})$, $c = 1, \dots, d$, are positive deterministic functions scaling the random price fluctuations as a function of time and start of delivery. $\boldsymbol{L}(t, \boldsymbol{T}^{\mathrm{s}})$ are independent d-dimensional MNIG-distributed column vectors with location parameter equal to zero:

$$\boldsymbol{L}(t, \boldsymbol{T}^{\mathrm{s}}) \sim \mathrm{MNIG}(\alpha, \delta, \boldsymbol{\mu} = \boldsymbol{0}, \boldsymbol{\beta}(t, \boldsymbol{T}^{\mathrm{s}}), \Sigma(t, \boldsymbol{T}^{\mathrm{s}}))$$

Note that the product of an MNIG-distributed vector with a non-stochastic matrix is MNIG distributed with parameters given by Property 3 in Appendix A. Moreover, the sum of a non-stochastic vector and an MNIG-distributed variable is also MNIG distributed, implying that log return is MNIG distributed. Consequently, we could embed both the drift and scaling terms in the MNIG-distributed vector $\boldsymbol{L}(t, \boldsymbol{T}^{\mathrm{s}})$, but we have chosen to extract them in order to present the model in a form similar to the LIBOR market model.

The MNIG-distributed variables $\boldsymbol{L}(t, \boldsymbol{T}^{\mathrm{s}})$ are characterized by the parameters $\alpha, \delta, \boldsymbol{\mu} = \boldsymbol{0}$, $\boldsymbol{\beta}(t, \boldsymbol{T}^{\mathrm{s}})$ and $\Sigma(t, \boldsymbol{T}^{\mathrm{s}})$. α and δ are scalars, whereas $\boldsymbol{\beta}(t, \boldsymbol{T}^{\mathrm{s}})$ is a d-dimensional vector and is assumed to be determined by the continuous scalar-valued function $\beta(\cdot)$:

$$\boldsymbol{\beta}(t, \boldsymbol{T}^{\mathrm{s}}) = [\beta(t, T_1^{\mathrm{s}}), \beta(t, T_2^{\mathrm{s}}), \dots, \beta(t, T_d^{\mathrm{s}})]'$$

The matrix $\Sigma(t, \boldsymbol{T}^{\mathrm{s}})$, which determines the overall shape of the correlation structure, is assumed to be determined by a positive definite scalar-valued function $R(t, T_j^{\mathrm{s}}, T_i^{\mathrm{s}})$ taking time and start of delivery as arguments:

$$\Sigma(t, \boldsymbol{T}^{\mathrm{s}})_{j,i} = R(t, T_j^{\mathrm{s}}, T_i^{\mathrm{s}}), \quad i, j \in 1, 2, \dots, d$$

The assumptions made above correspond to assuming an exponential model for the forward price dynamics:

$$\boldsymbol{F}(t) = \boldsymbol{F}(0) \exp\left(\sum_{i=0}^{t-1} \boldsymbol{\upsilon}(i, \boldsymbol{T}^{\mathrm{s}}) + \sum_{i=0}^{t-1} \Lambda(i, \boldsymbol{T}^{\mathrm{s}}) \boldsymbol{L}(i, \boldsymbol{T}^{\mathrm{s}}) \right) \tag{5}$$

The model states that the one-period return of forward contracts is MNIG distributed. By Property 2 in Appendix A we also know that the marginal one-period return distribution of each forward contract is NIG distributed. However, since a sum of MNIG-distributed variables with different α, $\boldsymbol{\beta}$ or Σ parameters is not MNIG distributed, the log returns over longer time periods do not, in general, have a known distribution.

In order to avoid arbitrage opportunities, forward contracts with overlapping delivery periods must have prices that are consistent with each other. Consider, for instance, that the prices $F_{Q_1}(t), \dots, F_{Q_4}(t)$ of four quarterly forward contracts with delivery periods each quarter over the next year are modeled by the dynamics given in Equation (5). The following relation between the quarterly contracts and the price $F_Y(t)$ of a forward with delivery period over the whole next year must hold:

$$F_Y(t) = \sum_{i=1}^{4} w_i F_{Q_i}(t)$$

Here the weight function is given by:

$$w_i = \frac{\int_{Q_i} w(u)\,\mathrm{d}u}{\int_Y w(u)\,\mathrm{d}u}$$

and $w(u) \equiv 1$ if we assume that the settlement of the forward takes place at the end of the delivery period (Benth and Koekebakker (2008)). Hence, if we fit the model to the quarterly contracts, we cannot derive the distribution of the one-period log returns of the yearly contract using no-arbitrage relationships. In order to state anything about the distribution of the yearly contract we have to rely on Monte Carlo analysis. In addition, it is not possible to derive any dynamics for the three monthly contracts making up the quarterly one. That is, the MNIG dynamics of the one-period return distribution stated by the model only applies to the contracts explicitly modeled.

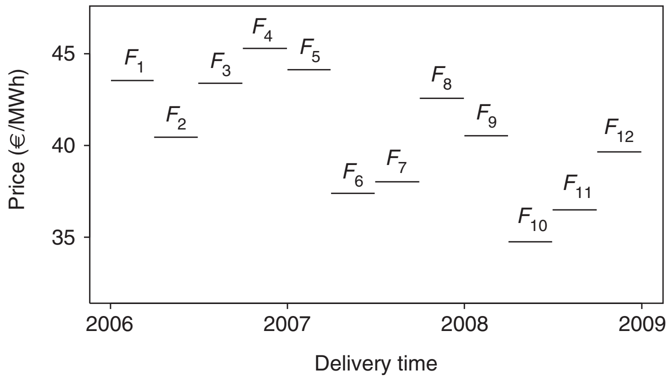
3 FITTING THE MODEL TO FORWARD PRICES AT NORD POOL

The distribution of the log return of forward contracts is determined by the parameters $\boldsymbol{\upsilon}(t, \boldsymbol{T}^{\mathrm{s}})$, $\Lambda(t, \boldsymbol{T}^{\mathrm{s}})$, α, δ, $\boldsymbol{\beta}(t, \boldsymbol{T}^{\mathrm{s}})$ and $\Sigma(t, \boldsymbol{T}^{\mathrm{s}})$. To simplify the estimation procedure we assume that the vector and matrix parameters are determined by time-homogenous functions: that is, the functions only depend on the time to start of delivery. It is important that these functions are chosen such that they are in conformity with the observed data. The aim of this section, therefore, is to investigate the empirical characteristics of forward prices at Nord Pool, and to choose the functional forms of these functions according to the empirical findings.

3.1 The dataset

The dataset that we consider in this section is made up of synthetic forward contracts with a delivery period of one quarter. Coverage spans from January 2001 to December

FIGURE 1 Term structure of synthetic forward prices on January 2, 2006.



Each contract has a delivery period of three months.

2006, with a total of 1,478 observations. For each trading day in this period, 12 forward contracts, each with a delivery period of one quarter, are constructed from a smoothed forward curve as described in Section 8.2 of Benth *et al* (2008). For each trading day, the first contract has start-settlement time zero and end-settlement time three months ahead. The second contract has start settlement three months ahead and end settlement six months ahead and so on (see Figure 1 for an illustration). We label the prices of these forward contracts $F_c(t)$, where t denotes time and $c = 1, \dots, 12$ denotes the number of quarters until end settlement.

Let $\boldsymbol{r}_t$ denote the 12-dimensional vector containing the log returns of the 12 forward contracts at time t as follows:

$$\boldsymbol{r}_t = \left[\ln\left(\frac{F_1(t+1)}{F_1(t)}\right) \ln\left(\frac{F_2(t+1)}{F_2(t)}\right) \cdots \ln\left(\frac{F_{12}(t+1)}{F_{12}(t)}\right) \right]'$$

According to Equation (4) and the assumptions that the parameters $\boldsymbol{\upsilon}(t, \boldsymbol{T}^{\mathrm{s}})$, $\Lambda(t, \boldsymbol{T}^{\mathrm{s}})$, $\boldsymbol{\beta}(t, \boldsymbol{T}^{\mathrm{s}})$ and $\Sigma(t, \boldsymbol{T}^{\mathrm{s}})$ are determined by time-stationary functions, the return vectors $\boldsymbol{r}_t$ are independent and identical MNIG-distributed variables. Table 1 on the facing page shows descriptive statistics of the individual forward contracts based on the dataset $\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_{1477}$.

Table 1 on the facing page shows that the empirical mean is slightly positive and rather constant. Regarding the empirical skewness, the numbers indicate the presence, to some extent, of negative skewness in the data. The kurtosis fluctuates a lot, but there seems to be a trend of decreasing kurtosis as the time to delivery increases. However, as demonstrated by Kim and White (2004), the conventional estimators of skewness

TABLE 1 Summary statistics of log return of the individual forward contracts.

Quarters to end delivery	Mean (%)	Annualized volatility (%)	Skewness	Excess kurtosis
1	0.06	48.90	−0.76	19.29
2	0.07	45.12	−1.06	12.10
3	0.07	49.72	−0.41	36.37
4	0.06	36.40	−0.22	17.19
5	0.07	30.54	−0.57	9.91
6	0.07	29.25	−0.83	11.19
7	0.07	29.39	−1.18	14.93
8	0.06	24.96	−0.32	6.23
9	0.06	26.56	0.07	6.21
10	0.06	30.40	−0.08	5.50
11	0.06	26.97	−0.25	7.89
12	0.05	24.79	−0.43	10.36

Volatility is annualized using 250 trading days a year.

and kurtosis are extremely sensitive to single outliers or small groups of outliers. These estimates should therefore be interpreted with caution.

3.2 Modeling $\boldsymbol{\upsilon}(t, \boldsymbol{T}^{\mathrm{s}})$, $\boldsymbol{\Lambda}(t, \boldsymbol{T}^{\mathrm{s}})$, $\boldsymbol{\beta}(t, \boldsymbol{T}^{\mathrm{s}})$ and $\boldsymbol{\Sigma}(t, \boldsymbol{T}^{\mathrm{s}})$

In our model we have assumed that the drift parameter $\boldsymbol{\upsilon}(t, \boldsymbol{T}^{\mathrm{s}})$ and the skewness parameter $\boldsymbol{\beta}(t, \boldsymbol{T}^{\mathrm{s}})$ are determined by time-stationary continuous functions of time t and time to start of delivery $\boldsymbol{T}^{\mathrm{s}}$; that is, $\beta(t, T^{\mathrm{s}})$ and $\upsilon(t, T^{\mathrm{s}})$ depend only on $T^{\mathrm{s}} - t$. For simplicity and ease of calibration, we choose a linear function for both $\beta(T^{\mathrm{s}} - t)$ and $\upsilon(T^{\mathrm{s}} - t)$. That is, $\beta(T^{\mathrm{s}} - t) = b_0 + b_1(T^{\mathrm{s}} - t)$ and $\upsilon(T^{\mathrm{s}} - t) = a_0 + a_1(T^{\mathrm{s}} - t)$, where b_0, b_1, a_0 and a_1 are parameters that must be estimated. Even though $\boldsymbol{\upsilon}(t, \boldsymbol{T}^{\mathrm{s}})$ and $\boldsymbol{\beta}(t, \boldsymbol{T}^{\mathrm{s}})$ are determined by linear functions, the expectation of an MNIG distribution, or the skewness of the marginal distributions of an MNIG distribution, is not necessarily a linear function of time until start of settlement.

The matrix $\Lambda(t, \boldsymbol{T}^{\mathrm{s}})$ determines the overall shape of the volatility structure, whereas the matrix $\Sigma(t, \boldsymbol{T}^{\mathrm{s}})$ determines the overall shape of the correlation structure. From Table 1 we observe that volatility increases as time to maturity decreases. We therefore model the volatility as a negative exponential function of time until start of settlement.[2]

We assume that all forward contracts share the same time-stationary scaling function $\lambda(t, T_{\mathrm{c}}^{\mathrm{s}})$, $c = 1, \dots, 12$. This seems reasonable as long as all contracts have the same length of delivery. Motivated by the empirical findings, we want the scaling

[2] Note, however, that a monotone decreasing volatility function does not fit the data perfectly.

function to decay slowly for the first contracts and to level out relatively quickly. We therefore choose an exponential scaling function where the time argument is squared:

$$\lambda(T^{\mathrm{s}} - t) = \exp(-\gamma_1 (T^{\mathrm{s}} - t)^2) + \gamma_2$$

Here $\gamma_1, \gamma_2 > 0$ are parameters to be estimated. Note that it is not necessary to include an additional parameter γ_3, yielding a scaling function of the form $\gamma_3 \exp(-\gamma_1 (T^{\mathrm{s}} - t)^2) + \gamma_2$. To see this, consider an MNIG-distributed vector $\boldsymbol{X}$. If we multiply this vector by the diagonal matrix Γ with γ on the main diagonal, it is straightforward to see, using Property 3 of Appendix A, that the structure matrix of the resulting MNIG-distributed vector remains unchanged. Therefore, it is possible to make the distribution of $\boldsymbol{X}$ equal to the distribution of $\Gamma \boldsymbol{X}$ by appropriately changing the vector and scalar parameters of the distribution. Consequently, multiplying the scaling function by a constant γ_3 would not improve the fit to historical log returns.

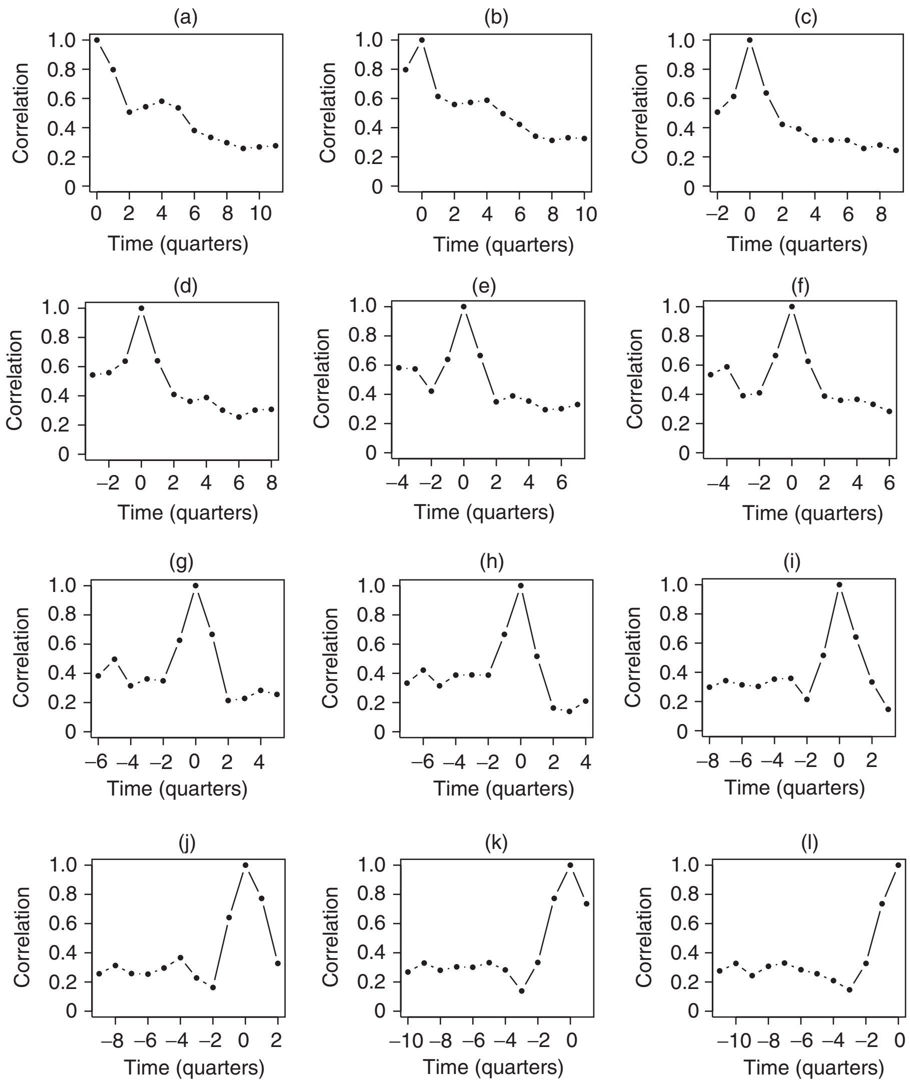
Next we investigate the correlation structure between forward contracts in order to assign a suitable functional form to the time-stationary positive definite function $R(\cdot)$ determining $\Sigma(t, \boldsymbol{T}^{\mathrm{s}})$. To estimate the correlation structure we use Spearman's correlation coefficient because, in contrast to the classical Pearson product–moment correlation estimator, it does not require the assumption that the returns have an elliptical distribution.

Figure 2 on the facing page shows the empirical correlation between the contracts. What is common across all plots is that the correlation decreases rapidly for the nearest two or three quarters. For contracts in the short end of the forward curve there seems to be a seasonal pattern in the correlation: the correlation with the other contracts decreases steadily before it bottoms out two quarters ahead, then stays rather constant for two quarters before again starting to decrease. In the long end of the forward curve, the plot indicates that the correlation is lowest with the forward contracts with time of delivery three months ahead or three months earlier. For the rest of the contracts there is a mixed pattern in the correlation structure. After the initial rapid fall in correlation in the two to three nearest quarters, it stays relatively constant or decays slowly. For a more thorough empirical examination of the swap correlation structure at Nord Pool we refer the reader to Frestad (2007).

In order to partly accommodate the remarks made above, we construct a function which is capable of capturing a seasonal pattern in the correlation structure. Because $\Sigma(t, \boldsymbol{T}^{\mathrm{s}})$ must be positive definite, we must restrict our attention to positive definite functions when assigning a functional form to $R(\cdot)$. Since both the product and the sum of two positive definite functions is a positive definite function, we see that the following function is positive definite:

$$R(x) = (1-k)\exp(-\theta_1 x^q) + k \frac{\sin(\theta_2 x)}{\theta_2 x} \exp(-\theta_3 x^p)$$

FIGURE 2 Empirical correlation between the contract in the title and the other contracts.



(a) $F_1(t)$, (b) $F_2(t)$, (c) $F_3(t)$, (d) $F_4(t)$, (e) $F_5(t)$, (f) $F_6(t)$, (g) $F_7(t)$, (h) $F_8(t)$, (i) $F_9(t)$, (j) $F_{10}(t)$, (k) $F_{11}(t)$ and (l) $F_{12}(t)$. The horizontal axis shows the lag in quarters between the contract in the title and the other contracts.

Here $x = |(t - T_i^s) - (t - T_j^s)|$, $i, j \in \{1, 2, \dots, 12\}$, whereas $k \in [0, 1]$, $p, q \in (0, 2)$ and $\theta_1, \theta_2, \theta_3 > 0$ are parameters that must be estimated from the data. The functions $\exp(-\theta x^p)$ and $\sin(\theta x)/\theta x$ are known to be positive definite (see, for example, Banerjee *et al* (2004)). In the proposed correlation function, the term $\sin(\theta x)/\theta x$ is responsible for capturing the seasonal pattern, whereas the exponential functions capture the fact that correlation decreases as time between contracts increases. The parameter k determines how prominent the seasonal pattern is, and the parameter θ_3 determines how persistent the seasonal pattern is.

3.3 Estimated parameter values

Estimating the parameters of the MNIG distribution is fairly complicated, because a direct maximization of the likelihood proves to be difficult (Øigård *et al* (2005)). Generally, the parameters of the MNIG distribution are most efficiently estimated by the expectation–maximization algorithm (see McNeil *et al* (2005) or Øigård *et al* (2005) for details on how to implement this algorithm). However, because the elements $\boldsymbol{\upsilon}(t, \boldsymbol{T}^{\mathrm{s}})$, $\Lambda(t, \boldsymbol{T}^{\mathrm{s}})$, $\boldsymbol{\beta}(t, \boldsymbol{T}^{\mathrm{s}})$ and $\Sigma(t, \boldsymbol{T}^{\mathrm{s}})$ are determined in our model by functions, explicit formulas for updating these parameters in the expectation–maximization scheme are generally not available. In this paper we therefore adopt a Bayesian approach to parameter estimation, using the Markov chain Monte Carlo method outlined in Appendix B (for background on Markov chain Monte Carlo algorithms we refer the reader to Robert and Casella (2004) or Gilks *et al* (1996)).

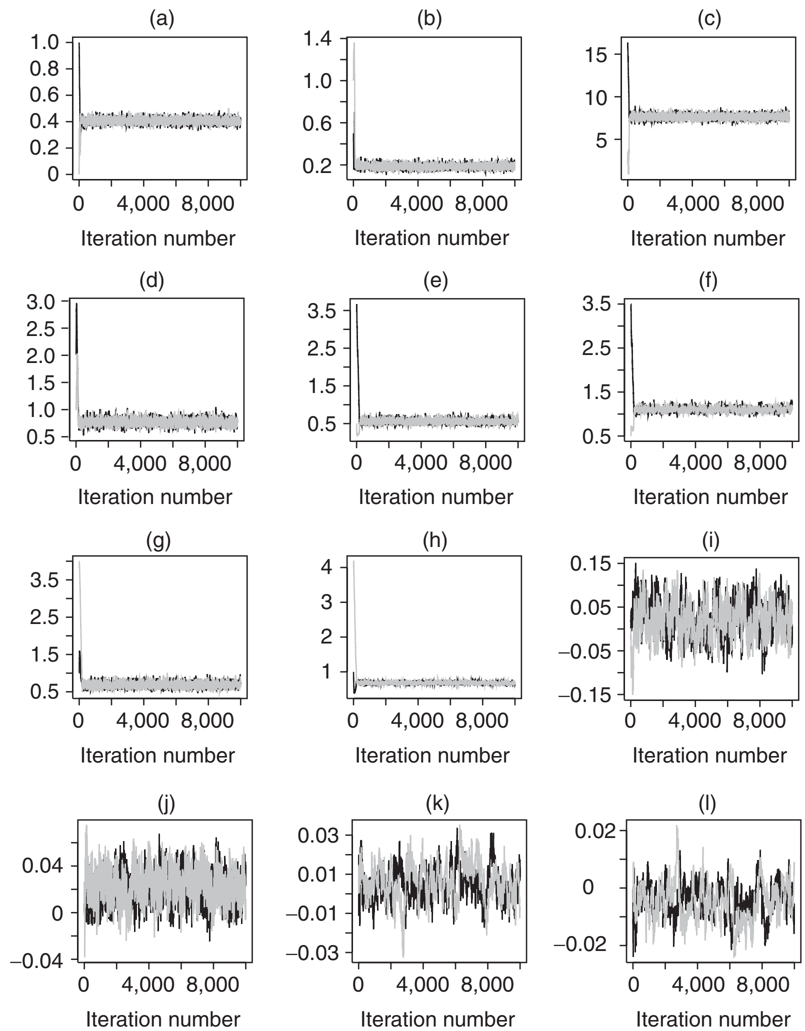
To reduce the number of parameters that must be estimated, we set the values of the parameters p and q in the correlation function equal to 1. Figure 3 on the facing page shows the simulated Markov chains from two runs of the algorithm with different initial values. Starting the algorithm with different initial values provides valuable information for monitoring the convergence of the algorithm, because chains started from different values can make lack of convergence apparent (see, for example, Cowles and Carlin (1996) for a review on monitoring the convergence Markov chain Monte Carlo methods).

Figure 3 on the facing page indicates that convergence to the stationary distributions is fast and that the mixing in most cases is good. The exceptions are the chains for the parameters b_0 and b_1, where the chains are highly autocorrelated. Nevertheless, both chains seem to have converged to the stationary distributions, and the only problem is that the exploration of the stationary distribution is slower than desired.

In both runs of the algorithm we set the burn-in length to 2,000 iterations and obtain the pooled estimates,[3] with corresponding standard deviations, as shown in Table 2 on page 14. Using a significance level of 5%, the parameters a_0, b_0 and b_1 turn out not to be significantly different from zero.

[3] We have also fitted the model to monthly forward contracts. The shape of the covariance structure is qualitatively the same, but the parameter estimates are slightly different. The interested reader can contact the authors for additional details.

FIGURE 3 Output from the Gibbs sampler plotted against iteration number.



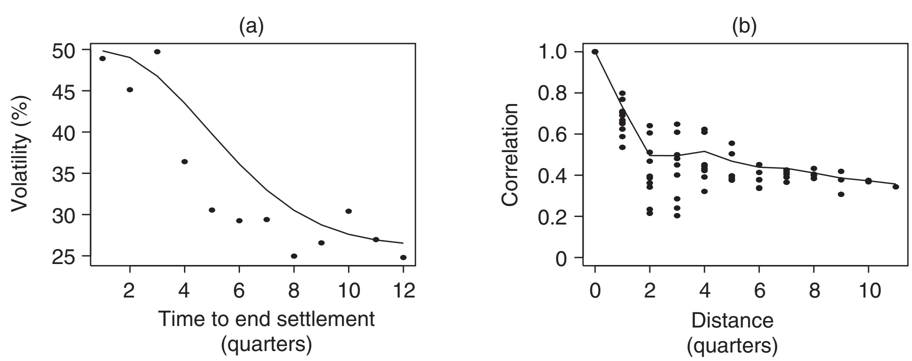
(a) k, (b) θ_1, (c) θ_2, (d) θ_3, (e) γ_1, (f) γ_2, (g) α, (h) δ, (i) a_0, (j) a_1, (k) b_0 and (l) b_1. The black line shows output from the first run, while the gray line shows output from the second run.

The solid lines in Figure 4 on the next page represent the implied correlation and volatility structure from our model using the parameter estimates from Table 2 on the next page. Historical correlation and volatility are represented by the dots.

TABLE 2 Estimated parameter values with corresponding standard deviation.

Parameter	Estimated value	Standard deviation
k	0.407	0.013
θ_1	0.186	0.017
θ_2	7.631	0.079
θ_3	0.792	0.032
γ_1	0.554	0.058
γ_2	1.106	0.078
α	0.690	0.052
δ	0.680	0.035
a_0	0.0239	0.0278
a_1	0.0217	0.0110
b_0	0.0063	0.0086
b_1	−0.0050	0.0061

FIGURE 4 Observed and modeled volatility and correlation.



Volatility is annualized using 250 trading days.

4 ASSESSING THE FIT TO FORWARD RETURNS

In assessing the quality of the fit to forward returns we focus on two issues:

1) how well the MNIG model fits the univariate empirical distributions;

2) how well the MNIG model predicts the correlation and variance structure when compared with data that was not used in the calibrating procedure.

Issue 1) corresponds to an investigation of the in-sample fit to forward returns, while issue 2) provides an out-of-sample examination of the fit to the correlation and variance structure.

In order to have some relative measure of quality, we compare the MNIG model with a market model where the log return of forward contracts is Gaussian distributed. More specifically, under the objective probability measure P, the dynamics of forward contracts is assumed[4] to have the following form:

$$\ln \frac{F_{\mathrm{c}}(t+1)}{F_{\mathrm{c}}(t)} = \mu(t - T_{\mathrm{c}}^{\mathrm{s}}) + \sum_{i=1}^{3} \sigma_i (t - T_{\mathrm{c}}^{\mathrm{s}}) W_i, \quad c = 1, 2, \ldots, d \tag{6}$$

Here $F_{\mathrm{c}}(t)$ denotes the price of a forward contract with delivery period $[T_{\mathrm{c}}^{\mathrm{s}}, T_{\mathrm{c}}^{\mathrm{e}}]$ at time t, $0 \leqslant t \leqslant T_{\mathrm{c}}^{\mathrm{s}}$, $\mu(t - T_{\mathrm{c}}^{\mathrm{s}})$ is a drift term, $\sigma_i(t - T_{\mathrm{c}}^{\mathrm{s}})$ and W_i, $i = 1, 2, 3$, are volatility functions and independent standard Gaussian-distributed variables, respectively.

Let S denote the empirical covariance matrix of the dataset consisting of log returns of the 12 forward contracts described in Section 3. By using principal components analysis, S can be decomposed as $S = VDV'$, where D is a diagonal matrix whose diagonal elements $d_{1,1}, d_{2,2}, \ldots, d_{12,12}$ are the eigenvalues of S, and V is an orthogonal matrix of order 12 whose ith column $\boldsymbol{v}_i$ is the eigenvector corresponding to $d_{i,i}$. We estimate the volatility functions as:

$$\sigma_i(t - T_{\mathrm{c}}^{\mathrm{s}}) = \sqrt{d_{i,i}}\, v_{c,i}$$

where $i = 1, 2, 3$, $c = 1, \ldots, 12$ and $v_{c,i}$ is the cth element of the ith eigenvector. $\mu(t - T_{\mathrm{c}}^{\mathrm{s}})$ is simply estimated as the empirical mean of the log return of the contract with time to start of delivery $t - T_{\mathrm{c}}^{\mathrm{s}}$.

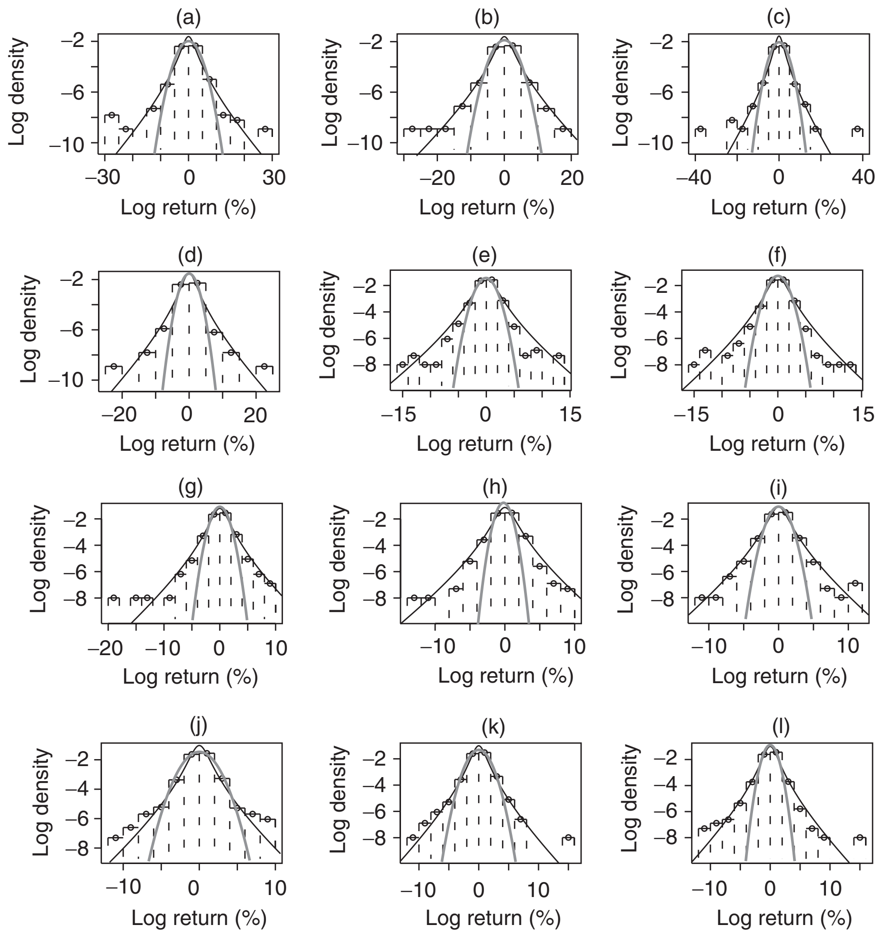
Although our proposed model is capable of modeling the joint dynamics of forward prices, for illustrative purposes we find it more appealing to investigate how well the model fits the univariate distributions. From Property 2 of Appendix A we can easily deduce the marginal distribution of individual forward contracts in our MNIG model. We then examine how well our model fits the empirical univariate distribution of the forward contracts compared with the 3-factor model.

From Figure 5 on the next page we see that the MNIG model offers a better fit to the observed return data when compared with the 3-factor model. The improvement is especially visible in the tails.

From Figure 4 on the facing page we see that the model's fit to historical variance and correlation is not perfect. We therefore want to investigate the relative quality of this fit

[4] A 3-factor model usually captures most of the term structure variation (see, for example, Litterman and Scheinkman (1991)). For a more dedicated discussion of the term structure of electricity forwards see Section 7.3 of Benth *et al* (2008).

FIGURE 5 Empirical log-density histograms of log-return distributions of individual forward contracts.



(a) $F_1(t)$, (b) $F_2(t)$, (c) $F_3(t)$, (d) $F_4(t)$, (e) $F_5(t)$, (f) $F_6(t)$, (g) $F_7(t)$, (h) $F_8(t)$, (i) $F_9(t)$, (j) $F_{10}(t)$, (k) $F_{11}(t)$ and (l) $F_{12}(t)$. The solid black line shows the fit from the MNIG model while the gray line shows the fit from the 3-factor model.

using cross-validation. We exclude some of the data when calibrating the model, then compare the covariance structure implied by the model with the covariance structure of the dataset left out when calibrating. In order to have some relative measure of the quality, we compare our model with the 3-factor Gaussian market model. As a second benchmark, we include a Gaussian market model in the same form as indicated by Equation (6), but with 12 factors instead of 3. The 12 volatility factors of this model

TABLE 3 Cross-validation variance prediction error for the proposed model and the factor models with 3 and 12 factors.

Year left out	Mean-squared difference between observed and model-implied variance structure: Proposed model	3-factor model	12-factor model
2001	22.55	8.03	13.24
2002	23.35	36.50	28.05
2003	31.51	47.16	36.89
2004	12.91	8.56	11.13
2005	11.69	2.68	5.32
2006	2.77	5.34	2.54
Sum mean-squared prediction error	104.78	108.27	97.17

TABLE 4 Cross-validation correlation prediction error for the proposed model and the factor models with 3 and 12 factors.

Year left out	Mean-squared difference between observed and model-implied correlation structure: Proposed model	3-factor model	12-factor model
2001	0.0528	0.0476	0.1079
2002	0.0139	0.0882	0.0096
2003	0.0822	0.2693	0.0390
2004	0.0977	0.2766	0.0475
2005	0.0126	0.0680	0.0450
2006	0.0149	0.0996	0.0121
Sum mean-squared prediction error	0.2741	0.8493	0.2611

are estimated using principal components analysis and, accordingly, this model is capable of reproducing the historical covariance structure perfectly.

Table 3 shows that the overall variance prediction errors do not differ much, with about 11% difference in sum mean-squared prediction error between the largest and smallest sum mean-squared prediction error. However, when it comes to predicting the correlation structure of data that was not used to calibrate the model, there are significant differences between the models. The total mean-squared correlation prediction error of the 3-factor model is about 200% higher than it is for the other two models. We also note that the cross-validating analysis of the 12-factor model indicates that the covariance structure is not time stationary. Audet *et al* (2004) argue that the volatility of forward contracts traded at Nord Pool is usually high in winter and low in summer, and they use a deterministic seasonal volatility function to capture this effect. They also note that, due to different annual hydro inflow, the volatility is different in different years. For example, autumn 2002 was unusually dry, leading to record high volatility levels in the winter.

5 CONCLUDING REMARKS

We have proposed the use of an MNIG distribution to model electricity forward prices. We have shown that the MNIG model offers a superior fit to the empirical distribution of log returns when compared with a Gaussian multifactor model. This comes as no surprise, since the MNIG distribution nests the multivariate Gaussian distribution. We have also demonstrated that our model, which has separate functions for determining the volatility and correlation structure, compares favorably with a 3-factor model when it comes to prediction of out-of-sample correlation structure.

Since the distribution of log return over each time step has a known distribution, which can be characterized as a variance–mean mixture of a d-dimensional Gaussian random variable with a univariate inverse Gaussian-distributed mixing variable, it is easy to simulate from the model.

APPENDIX A: SOME PROPERTIES OF THE MULTIVARIATE NORMAL INVERSE GAUSSIAN DISTRIBUTION

Property 1

Let $\boldsymbol{X}_1, \dots, \boldsymbol{X}_M$ be M independent MNIG variables with common shape parameters α, $\boldsymbol{\beta}$ and Σ but potential different location parameters $\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_M$ and scale parameters $\delta_1, \dots, \delta_M$. Then the sum $\boldsymbol{Y} = \boldsymbol{X}_1 + \cdots + \boldsymbol{X}_M$ is also MNIG distributed (see Øigård *et al* (2005)):

$$\boldsymbol{Y} \sim \text{MNIG}\left(\alpha, \boldsymbol{\beta}, \sum_{i=1}^{M} \boldsymbol{\mu}_i, \sum_{i=1}^{M} \delta_i, \Sigma\right)$$

Property 2

The marginal distributions of the MNIG distribution are univariate NIG distributions (Lillestøl (2000)). If we denote the parameters of the marginal distribution of the ith component X_i of $\boldsymbol{X}$ by α_i, β_i, δ_i and μ_i, then they are related to the parameters α, $\boldsymbol{\beta}$, δ, Σ and $\boldsymbol{\mu}$ in the following way:

$$\mu_i = \boldsymbol{\mu}_i$$
$$\delta_i = \sqrt{\Sigma_{ii}}\delta$$
$$\beta_i = \frac{1}{\Sigma_{ii}} \sum_{k=1}^{d} \Sigma_{ik} \boldsymbol{\beta}_k$$
$$\alpha_i = \sqrt{\frac{1}{\Sigma_{ii}}(\alpha^2 - \boldsymbol{\beta}'\Sigma\boldsymbol{\beta}) + \beta_i^2}$$

Property 3

Assume that A is a real-valued $d \times d$ coefficient matrix and that $\boldsymbol{b}$ is a $d \times 1$-dimensional real vector. Then the linear transformation $\boldsymbol{Y} = \boldsymbol{b} + A\boldsymbol{X}$ of an MNIG-distributed variable $\boldsymbol{X}$ with parameters $\alpha, \boldsymbol{\beta}, \delta, \Sigma$ and $\boldsymbol{\mu}$ is also MNIG distributed with parameters given by (Øigård *et al* (2005)):

$$\tilde{\alpha} = \alpha |\det A|^{-1/d}$$
$$\tilde{\boldsymbol{\beta}} = (A^{-1})'\boldsymbol{\beta}$$
$$\tilde{\delta} = \delta |\det A|^{1/d}$$
$$\tilde{\boldsymbol{\mu}} = \boldsymbol{b} + A\boldsymbol{\mu}$$
$$\tilde{\Sigma} = A\Sigma A' |\det A|^{-2/d}$$

APPENDIX B: SAMPLING SCHEME FOR ESTIMATING THE PARAMETERS OF THE MULTIVARIATE NORMAL INVERSE GAUSSIAN MODEL

In our forward-market model the parameters $\boldsymbol{\upsilon}(t, \boldsymbol{T}^{\mathrm{s}})$, $\boldsymbol{\beta}(t, \boldsymbol{T}^{\mathrm{s}})$, $\Lambda(t, \boldsymbol{T}^{\mathrm{s}})$ and $\Sigma(t, \boldsymbol{T}^{\mathrm{s}})$ are determined by the functions $\upsilon(\cdot)$, $\beta(\cdot)$, $\lambda(\cdot)$ and $R(\cdot)$, respectively. These functions depend on the parameters a_0, a_1, b_0, b_1, γ_1, γ_2, k, p, q, θ_1, θ_2 and θ_3, as described in Section 3. To reduce the dimension of the parameter space and, therefore, to ease the estimation procedure, we have decided *a priori* to set the values of the parameters p and q equal to 1. Together with the parameters α and δ, we stack the remaining parameters in a vector $\boldsymbol{\eta} = (\alpha, \delta, \gamma_1, \gamma_2, \theta_1, \theta_2, \theta_3, a_0, a_1, b_0, b_1, k)$. Our goal is then to obtain the posterior distribution of $\boldsymbol{\eta}$ given the observed log returns $\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_{1477}$. We now show how this can be accomplished using the Gibbs sampler and the Metropolis–Hastings algorithm.

Because the log returns $\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_{1477}$ are assumed to be independent and identical MNIG distributed, the joint likelihood factorizes into a product of marginal likelihoods:

$$L(\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_{1477} \mid \boldsymbol{\eta}) = \prod_{i=1}^{1477} L(\boldsymbol{r}_i \mid \boldsymbol{\eta})$$

Here $L(\boldsymbol{r}_i \mid \boldsymbol{\eta})$ denotes the likelihood function of an MNIG distribution. Let $\pi(\boldsymbol{\eta})$ denote the prior density of $\boldsymbol{\eta}$. Bayes's theorem relates the posterior distribution $\pi(\boldsymbol{\eta} \mid \boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_{1477})$ to the likelihood function and the prior distribution via the following formula:

$$\pi(\boldsymbol{\eta} \mid \boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_{1477}) = \frac{\pi(\boldsymbol{\eta}) \prod_{i=1}^{1477} L(\boldsymbol{r}_i \mid \boldsymbol{\eta})}{\int \pi(\boldsymbol{\eta}) \prod_{i=1}^{1477} L(\boldsymbol{r}_i \mid \boldsymbol{\eta}) \, \mathrm{d}\boldsymbol{\eta}}$$

We are interested in obtaining an estimate of $\boldsymbol{\eta}$ conditional on the observations $\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_{1477}$. This can be obtained by finding the expectation of $\boldsymbol{\eta}$:

$$E[\boldsymbol{\eta} \mid \boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_{1477}] = \int \boldsymbol{\eta} \pi(\boldsymbol{\eta} \mid \boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_{1477}) \, \mathrm{d}\boldsymbol{\eta}$$

Explicit evaluations of these integrals are not possible. However, the Markov chain Monte Carlo method provides an alternative, where we draw a sample from the posterior distribution and obtain sample estimates of the quantities of interest, thereby performing the integration implicitly.

The distributions involved above are too complex to generate samples directly. We therefore rely on a sampling scheme like the Gibbs sampler. The Gibbs sampler proceeds by splitting the parameter vector into groups and updating each group in turn by a series of Gibbs transitions.

We assume that the components of parameter vector $\boldsymbol{\eta}$ are independent *a priori*: that is, the prior distribution is of the form:

$$\pi(\boldsymbol{\eta}) \propto \prod_{i=1}^{12} \pi(\eta_i)$$

In the absence of good prior information, a convenient strategy is to use diffuse proper priors. The prior distributions adopted are as follows: $\pi(\eta_i) \sim N_0^{\infty}(1, 10^3)$, $i = 1, \dots, 7$, $\pi(\eta_i) \sim N(0, 10^3)$, $i = 8, \dots, 11$, $\pi(\eta_{12}) \sim U(0, 1)$. Here $N_0^{\infty}(1, 10^3)$ is a truncated Gaussian distribution with mean 1 and standard deviation 10^3. The lower truncation limit is 0 and the upper truncation limit is ∞. $N(0, 10^3)$ is a Gaussian distribution with mean 0 and standard deviation 10^3, $U(0, 1)$ is a uniform distribution on the interval $[0, 1]$. With these prior specifications, parameter inferences are virtually unaffected by the prior distributions.

Let the observed log returns $\boldsymbol{r}_1, \boldsymbol{r}_2, \dots, \boldsymbol{r}_{1477}$ be denoted by $\boldsymbol{r}$. The posterior distribution is then given by $\pi(\boldsymbol{\eta} \mid \boldsymbol{r}) \propto \pi(\boldsymbol{\eta}) L(\boldsymbol{r} \mid \boldsymbol{\eta})$, where $L(\boldsymbol{r} \mid \boldsymbol{\eta})$ is the likelihood function of the log returns. The full conditional distribution of the parameter η_i,

$p_i(\eta_i \mid \boldsymbol{r}, \boldsymbol{\eta}_{j \neq i})$, is easily found by considering all the other parameters $\boldsymbol{\eta}_{j \neq i}$, as given constants in the posterior distribution. Unfortunately, it seems to be the case that the full conditional distributions are not standard distributions that we can easily generate samples from directly. Therefore, when using the Gibbs sampler to estimate the parameters, we use a method like the Metropolis–Hastings algorithm within the Gibbs sampler.

The sampling scheme is as follows. Give the parameter vector:

$$\boldsymbol{\eta} = (\alpha, \delta, \gamma_1, \gamma_2, \theta_1, \theta_2, \theta_3, a_0, a_1, b_0, b_1, k)$$

suitable starting values such as $\boldsymbol{\eta}^{(0)} = (\alpha^{(0)}, \delta^{(0)}, \dots, k^{(0)})$. Repeat for $n = 1, 2, \dots, N$. For $i = 1, 2, \dots, 12$, do the following. Draw η_i^* from the proposal distribution $q_i(\eta_i^* \mid \eta_i^{(n-1)})$ and set:

$$\boldsymbol{\eta}^* = (\eta_1^{(n)}, \dots, \eta_{i-1}^{(n)}, \eta_i^*, \eta_{i+1}^{(n-1)}, \dots, \eta_{12}^{(n-1)})$$
$$\boldsymbol{\eta}^{(n)} = (\eta_1^{(n)}, \dots, \eta_{i-1}^{(n)}, \eta_i^{(n-1)}, \eta_{i+1}^{(n-1)}, \dots, \eta_{12}^{(n-1)})$$

Compute:

$$r = \frac{q_i(\eta_i^{(n-1)} \mid \eta_i^*)}{q_i(\eta_i^* \mid \eta_i^{(n-1)})} \exp[\log(\pi(\boldsymbol{\eta}^* \mid \boldsymbol{r})) - \log(\pi(\boldsymbol{\eta}^{(n)} \mid \boldsymbol{r}))]$$

If:

$$r \geqslant 1, \quad \text{set } \eta_i^{(n)} = \eta_i^*$$

If:

$$r < 1, \quad \text{set } \eta_i^{(n)} = \begin{cases} \eta_i^* & \text{with probability } r, \\ \eta_i^{(n-1)} & \text{with probability } 1 - r \end{cases}$$

We use the following proposal distributions: for $i = 1, \dots, 5$ and 7, $q_i(\eta_i^* \mid \eta_i^{(n-1)})$ is equal to $N_0^\infty(\eta_i^{(n-1)}, 0.05)$; $q_6(\eta_6^* \mid \eta_6^{(n-1)})$ is equal to $N_0^\infty(\eta_6^{(n-1)}, 0.25)$ for $i = 8, \dots, 11$; $q_i(\eta_i^* \mid \eta_i^{(n-1)})$ is equal to $N(\eta_i^{(n-1)}, 0.01)$ and $q_{12}(\eta_{12}^* \mid \eta_{12}^{(n-1)})$ is equal to $N_0^1(\eta_{12}^{(n-1)}, 0.05)$. Here $N(\cdot, \cdot)$ denotes a Gaussian distribution while $N_l^u(\cdot, \cdot)$ denotes a truncated Gaussian distribution with lower truncation limit l and upper truncation limit u.

When obtained at iteration n, $\boldsymbol{\eta}^{(n)}$ converges in distribution to a draw from the true joint posterior distribution $\pi(\boldsymbol{\eta} \mid \boldsymbol{r})$. This means that, for sufficiently large n, bigger than n_0, say, $\{\boldsymbol{\eta}^n \mid n = n_0 + 1, \dots, N\}$ is a sample from the true posterior, from which any posterior quantities of interest may be estimated. To estimate the ith component of $\boldsymbol{\eta}$ we simply use:

$$\hat{\eta}_i = (N - n_0)^{-1} \sum_{n=n_0+1}^{N} \eta_i^n$$

REFERENCES

Audet, N., Heiskanen, P., Keppo, J., and Vehvilainen, I. (2004). Modeling electricity forward curve dynamics in the Nordic market. In *Modeling Prices in Competitive Electricity Markets*, Bunn, D. (ed), pp. 251–265. John Wiley & Sons.

Banerjee, S., Carlin, B. P., and Gelfand, A. E. (2004). *Hierarchical Modeling and Analysis For Spatial Data*. Chapman & Hall/CRC, London.

Barndorff-Nielsen, O. E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics* **24**, 1–13.

Barndorff-Nielsen, O. E. (1998). Processes of normal inverse Gaussian type. *Finance and Stochastics* **2**, 41–68.

Benth, F. E., and Koekebakker, S. (2008). Stochastic modeling of financial electricity contracts. *Energy Economics* **30**, 1116–1157.

Benth, F. E., Kallsen J., and Meyer-Brandis, T. (2007a). A non-Gaussian Ornstein–Uhlenbeck process for electricity spot price modeling and derivatives pricing. *Applied Mathematical Finance* **14**, 153–169.

Benth, F. E., Koekkebakker, S., and Ollmar, F. (2007b). Extracting and applying smooth forward curves from average-based commodity contracts with seasonal variation. *Journal of Derivatives* **15**, 52–66.

Benth, F. E., Benth, J. S., and Koekebakker, S. (2008). *Stochastic Modeling of Electricity and Related Markets*. World Scientific.

Bjerksund, P., Rasmussen, H., and Stensland, G. (2000). Valuation and risk management in the Nordic electricity market. Working Paper, Institute of Finance and Management Sciences, Norwegian School of Economics and Business Administration.

Brace, A., Gatarek, D., and Musiela, M. (1997). The market model of interest rate dynamics. *Mathematical Finance* **7**, 127–155.

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

Clewlow, L., and Strickland, C. (2000). *Energy Derivatives: Pricing and Risk Management*. Lacima, Houston, TX.

Cowles, M. K., and Carlin, B. P. (1996). Markov chain Monte Carlo convergence diagnostics: a comparative review. *Journal of the American Statistical Association* **91**, 883–904.

Fleten, S.-E., and Lemming, J. (2003). Constructing forward price curves in electricity markets. *Energy Economics* **25**, 409–424.

Frestad, D. (2007). Electricity forward return correlations. Working Paper, University of Agder, Norway.

Frestad, D. (2008). Common and unique factors influencing daily swap returns in the Nordic electricity market, 1997–2005. *Energy Economics* **30**, 1081–1097.

Frestad, D., Benth, F. E., and Koekebakker, S. (2010). Modeling term structure dynamics in the Nordic electricity swap market. *Energy Journal* **31**, 53–86.

Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (1996). *Markov Chain Monte Carlo in Practice*. Chapman & Hall/CRC, London.

Jamshidian, F. (1997). LIBOR and swap market models and measures. *Finance and Stochastics* **1**, 293–330.

Kennedy, D. P. (1994). The term structure of interest rates as a Gaussian random field. *Mathematical Finance* **4**, 247–258.

Kiesel, R., Schindlmayr, G., and Boerger, R. H. (2009). A two-factor model for the electricity forward market. *Quantitative Finance* **9**, 279–287.

Kim, T. H., and H. White (2004). On more robust estimation of skewness and kurtosis. *Finance Research Letters* **1**, 56–73.

Koekebakker, S., and Ollmar, F. (2005). Forward curve dynamics in the Nordic electricity market. *Managerial Finance* **31**, 73–94.

Lillestøl, J. (2000). Risk analysis and the NIG distribution. *The Journal of Risk* **2**, 41–56.

Litterman, R. B., and Scheinkman, J. (1991). Common factors affecting bond returns. *Journal of Fixed Income* **1**, 54–61.

Lucia, J. J., and Schwartz, E. S. (2002). Electricity prices and power derivatives: evidence from the Nordic power exchange. *Review of Derivatives Research* **5**, 5–50.

McNeil, A. J., Frey, R., and Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press.

Øigård, T. A., Hanssen, A., Hansen, R. E., and Godtliebsen, F. (2005). EM-estimation and modeling of heavy-tailed processes with the multivariate normal inverse Gaussian distribution. *Signal Processing* **85**, 1655–1673.

Quinn, J. A., Reitzes, J. D., and Schumacher, A. C. (2005). Forward and spot prices in electricity and gas markets. In *Obtaining the Best from Regulation and Competition*, Crew, M. A. and Spiegel, M. (eds), pp. 109–134. Springer.

Robert, C. P., and Casella, G. (2004). *Monte Carlo Statistical Methods*. Springer.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory* **13**, 341–360.