Calibration of a multifactor model for the forward markets of several commodities

Enrico Edoli[*1,2], Davide Tasinato[1], and Tiziano Vargiolu[†1]

[1]Department of Mathematics, via Trieste 63, University of Padova, I-35121 Padova, Italy
[2]Finalyst SAS, via Trentino 28, I-30010 Campolongo Maggiore, Italy

September 28, 2012

Abstract

We propose a model for the evolution of forward prices of several commodities, which is an extension of the two-factor forward model by Kiesel *et al.* (2009), originally conceived for the electricity forward market, to a market where multiple commodities are traded. We then show how to calibrate this model in a market where few or no derivative assets on forward contracts are present. We thus perform a calibration based on historical forward prices. First we calibrate separately the four coefficients of every single commodities, using an approach based on quadratic variation. Then we pass to estimate the mutual correlation among the Brownian motions driving the different commodities, the estimates being based now on the quadratic covariation between forward prices of different commodities. This calibration is compared to a calibration method used by practitioners, which uses rolling time series, which however requires a modification of the model.

Keywords: two-factor model for forward prices, historical calibration, quadratic variation, quadratic covariation.

1 Introduction

When dealing with forward prices of a single commodity having different maturities, the two-factor model proposed by Kiesel *et al.* [6] is quite simple to understand, analytically tractable and gives a good fit of several stylized fact. The first is the so-called Samuelson effect, i.e. the local volatility of a short-term forward contract is greater than the local volatility of a long-term contract, and in particular an exponential decay is observed as the time to maturity of the contract grows. The second stylized fact

[*]`enricoedoli@gmail.com`
[†]Corresponding Author, +39 049 8271383 `vargiolu@math.unipd.it`.

is that this volatility does not go to zero, but rather to a fixed value, called long-term volatility, due to long term uncertainty factors like technological innovations, change in geo-political equilibria, structural modifications to commodity prices, and so on. Moreover, the model is consistent with market data and with the Schwarz-Smith model for the spot price [8], (see [1] for details), which exibits mean reversion, another stylized fact which is observed in the markets.

We extend this model by assuming to have $K \geqslant 2$ commodities in our market, and that, for each one of the commodity, their forward prices follow the following two-dimensional model: by denoting with $F^k(t,T)$ the price at time t of a forward contract on the commodity $k = 1, \dots, K$ with maturity T, we assume that under a forward-neutral probability measure $\mathbb{Q}_T$ its dynamics are

$$dF^k(t,T) = F^k(t,T)(e^{-\lambda^k(T-t)}\sigma_1^k \, dW_1^k(t) + \sigma_2^k \, dW_2^k(t))$$

where W_1^k and W_2^k are two correlated Brownian motions with correlation ρ^k and the other parameters represent, respectively:

- σ_1^k - spot volatility, i.e. how much the forward price is influenced by short period shocks;
- σ_2^k - long term volatility, i.e. how much the forward price is influenced by long period uncertainty;
- λ^k - mean-reversion speed, or speed of decaying of the spot volatility.

Thus, when fitting this model to the market data of the k-th commodity, we have to calibrate the four parameters $p^k = (\sigma_1^k, \sigma_2^k, \lambda^k, \rho^k)$. Moreover, we assume that the Brownian motions of the commodities also have an inter-commodity correlation, given by the correlation matrix

$$\rho_{a,b}^{k,m} := corr(W_a^k(t), W_b^m(t)) = \text{Cov}\,(W_a^k(t), W_b^m(t))/t, \qquad \text{i.e.} \qquad \rho_{a,b}^{k,m} := \text{Cov}\,(W_a^k(1), W_b^m(1))$$

for all $a, b = 1, 2$ and $k, m = 1, \dots, K$: of course

$$\rho_{1,2}^{k,k} = \rho_{2,1}^{k,k} = \rho^k$$

Thus, the $2K$-dimensional Brownian motion $(W_1^1, W_2^1, \dots, W_1^K, W_2^K)$ has correlation matrix

$$\boldsymbol{\rho} = (\boldsymbol{\rho}^{k,m})_{1 \leqslant k,m \leqslant K} := \begin{pmatrix} \boldsymbol{\rho}^{1,1} & \cdots & \boldsymbol{\rho}^{1,K} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\rho}^{K,1} & \cdots & \boldsymbol{\rho}^{K,K} \end{pmatrix} \tag{1}$$

where

$$\boldsymbol{\rho}^{k,m} = (\rho_{a,b}^{k,m})_{1 \leqslant a,b \leqslant 2} := \begin{pmatrix} \rho_{1,1}^{k,m} & \rho_{1,2}^{k,m} \\ \rho_{2,1}^{k,m} & \rho_{2,2}^{k,m} \end{pmatrix}$$

Recall that, being $\boldsymbol{\rho}$ a correlation matrix, it is symmetric, semi-positive definite, with $\rho^{k,k}_{a,a} = 1$ for all $k = 1, \ldots, K$ and $a = 1, 2$ and $\rho^{k,m}_{1,1} \in [-1, 1]$ for all $k, m = 1, \ldots, K$ and $a, b = 1, 2$.

This model is analytically tractable because, under the forward measure $\mathbb{Q}_T$, each $F^k(\cdot, T)$ has a lognormal evolution, given by

$$F^k(t,T) = F^k(t_0,T) \exp\left(\int_{t_0}^t e^{-\lambda^k(T-s)} \sigma^k_1 \, dW^k_1(s) + \int_{t_0}^t \sigma^k_2 \, dW^k_2(s) - \frac{1}{2} \int_{t_0}^t \Sigma^k(s,T)^2 \, ds \right)$$

where $\Sigma_k(s,T)$ is a sort of local volatility at time s, given by

$$\Sigma^k(s,T) := \sqrt{e^{-2\lambda^k(T-s)}(\sigma^k_1)^2 + 2\rho^k e^{-\lambda^k(T-s)} \sigma^k_1 \sigma^k_2 + (\sigma^k_2)^2}$$

Thus, $\log F^k(t,T)$ has a Gaussian distribution, with mean

$$\mathbb{E}_{\mathbb{Q}_T}[\log F^k(t,T)] = \log F^k(t_0,T) - \frac{1}{2} \int_{t_0}^t \Sigma^2_k(s,T) \, ds$$

and variance

$$\mathrm{Var}_{\mathbb{Q}_T}[\log F^k(t,T)] = \int_{t_0}^t \Sigma^2_k(s,T) \, ds$$

In this paper we want to calibrate this model in a situation where, for each commodity $k = 1, \ldots, K$, forward contracts with (a finite number of) different maturities $T^k_1, \ldots, T^k_{N_k}$ are present, and few or no derivatives on these forward contracts are traded, as can be the case of some markets and/or some commodities. We thus perform a calibration based on historical forward prices. The strategy is first to calibrate separately the four coefficients of every single commodities, as we want them to have priority and greater precision than the correlations among different commodities: in fact, the main aim of our calibration is that it should reproduce well first of all the price behaviour of single-commody products. Secondly, we estimate the correlation matrix also in the inter-commodity correlations.

More in detail, Section 2 shows the calibration procedure of the four parameters of a single commodity, with an approach based on quadratic variation-covariation. Section 3 shows the calibration procedure of the residual parameters, i.e. the inter-commodity correlations, again with an approach based on quadratic covariation. Section 4 present an alternative calibration method which is mostly used by practitioners and uses rolling time series: this method is simpler but, to be made rigorous, it requires to work with a modified model. In Section 5 we show how to perform the intercommodity calibration of the global correlation matrix $\boldsymbol{\rho}$ in a way which is numerically efficient, based on the Cholesky decomposition. In Section 6 we test the two methods against simulated data at two different time scales, namely with daily data and with high-frequency data (200 per day). Section 7 concludes.

2 Single commodity calibration

We now fix the commodity $k = 1, \ldots, K$ and assume that, as already mentioned in the Introduction, we have a market where forward contracts with maturities $T_1, \ldots, T_N$ are traded (in this section we omit the dependences on k of the maturities). Then, by denoting $X_i^k(t) := \log F^k(t, T_i)$, we have that

$$dX_i^k(t) = e^{-\lambda^k(T_i - t)}\sigma_1^k \, dW_1^k(t) + \sigma_2^k \, dW_2^k(t) + \text{ drift}$$

under the forward-neutral probability $\mathbb{Q}_T$. Since we want to perform an historical calibration, we need dynamics under the real world probability $\mathbb{P}$. By the Girsanov theorem, the dynamics of X_i^k under $\mathbb{P}$ is given by

$$dX_i^k(t) = e^{-\lambda^k(T_i - t)}\sigma_1^k \, d\tilde{W}_1^k(t) + \sigma_2^k \, d\tilde{W}_2^k(t) + \text{ drift}$$

where $\tilde{W}_1^k$ and $\tilde{W}_2^k$ are Brownian motions under $\mathbb{P}$, still with mutual correlation ρ^k, but the drift in the two dynamics are different, as in the second drift also the market price of risk is present. We notice that the coefficients $p^k = (\sigma_1^k, \sigma_2^k, \lambda^k, \rho^k)$ can be estimated directly under $\mathbb{P}$. A more direct writing of the dynamics of X_i^k under $\mathbb{P}$ is

$$dX_i^k(t) = \Sigma_i^k(t) \, d\overline{W}^k(t) + \text{ drift}$$

where

$$\Sigma_i^k(t) := \Sigma^k(s, T_i) = \sqrt{e^{-2\lambda^k(T_i - t)}(\sigma_1^k)^2 + 2\rho^k e^{-\lambda^k(T_i - t)}\sigma_1^k\sigma_2^k + (\sigma_2^k)^2}$$

and $\overline{W}^k$ is a suitable 1-dimensional Brownian motion under $\mathbb{P}$.

The fact that the diffusion coefficient of the X_i^k, $i = 1, \ldots, N$, under $\mathbb{P}$ is deterministic gives us a easy way to estimate the parameters. In fact, the quadratic variation of X_i^k under $\mathbb{P}$ is given by

$$\langle X_i^k \rangle_{t_0}^t := \lim_{n \to \infty} \sum_{l=1}^{n} (X_i^k(t_{l+1}) - X_i^k(t_l))^2 = \int_{t_0}^{t} (\Sigma_i^k(u))^2 \, du \tag{2}$$

where $t_0 < t_1 < \ldots < t_n = t$ are suitable sequences, and the quadratic covariation of X_i^k, X_j^k, always under $\mathbb{P}$, is given by

$$\langle X_i^k, X_j^k \rangle_{t_0}^t := \lim_{n \to \infty} \sum_{l=1}^{n} (X_i^k(t_{l+1}) - X_i^k(t_l))(X_j^k(t_{l+1}) - X_j^k(t_l)) = \int_{t_0}^{t} \Sigma_{i,j}^k(u) \, du \tag{3}$$

(for more details, see [7]). Now, the last term of these equalities is explicitly computable ($\Sigma_{i,j}^k(u)$ will be specified later in the next Lemma 2.2), while the middle term can be approximated with historical observations. This gives us an idea to calibrate the model: given the historical quadratic covariations, our aim is to find coefficients p^k such that the theoretical quadratic covariations of all forward contracts match as close as possible the historical quadratic covariations.

In order to do this, we must calculate analytically the integrals in Equations (2–3).

Lemma 2.1 *The quadratic variation of the process X_i^k is given by*

$$\begin{aligned}\langle X_i^k\rangle_{T_0^i}^{T_i^1} &= \int_{T_0^i}^{T_i^1} (\Sigma_i^k(u))^2\, du = \\ &= \frac{\left(\sigma_1^k\right)^2}{2\lambda^k}\left(e^{-2\lambda^k\left(T_i-T_i^1\right)} - e^{-2\lambda^k\left(T_i-T_i^0\right)}\right) + \left(\sigma_2^k\right)^2 (T_i^1 - T_i^0) + \\ &\quad + \frac{2\sigma_1^k\sigma_2^k\rho_{1,2}^{k,k}}{\lambda^k}\left(e^{-\lambda^k\left(T_i-T_i^1\right)} - e^{-\lambda^k\left(T_i-T_i^0\right)}\right)\end{aligned}$$

Proof. We have

$$\begin{aligned}\int_{T_i^0}^{T_i^1}\left(\Sigma_i^k(t)\right)^2 dt &= \int_{T_i^0}^{T_i^1}\left(e^{-2\lambda^k(T_i-t)}\left(\sigma_1^k\right)^2 + \left(\sigma_2^k\right)^2 + 2e^{-\lambda^k(T_i-t)}\sigma_1^k\sigma_2^k\rho_{1,2}^{k,k}\right) dt \\ &= \left(\sigma_1^k\right)^2\left[e^{-2\lambda^k(T_i-t)}\right]_{T_i^0}^{T_i^1} + \left(\sigma_2^k\right)^2 (T_i^1 - T_i^0) + 2\sigma_1^k\sigma_2^k\rho_{1,2}^{k,k}\left[e^{-\lambda^k(T_i-t)}\right]_{T_i^0}^{T_i^1} \\ &= \frac{\left(\sigma_1^k\right)^2}{2\lambda^k}\left(e^{-2\lambda^k\left(T_i-T_i^1\right)} - e^{-2\lambda^k\left(T_i-T_i^0\right)}\right) + \left(\sigma_2^k\right)^2 (T_i^1 - T_i^0) + \\ &\quad + \frac{2\sigma_1^k\sigma_2^k\rho_{1,2}^{k,k}}{\lambda^k}\left(e^{-\lambda^k\left(T_i-T_i^1\right)} - e^{-\lambda^k\left(T_i-T_i^0\right)}\right)\end{aligned}$$

□

Lemma 2.2 *The quadratic covariation of the processes X_i^k, X_j^k is given by*

$$\begin{aligned}\langle X_i^k, X_j^k\rangle_{T_0^i}^{T_i^1} &= \left(\sigma_2^k\right)^2 (T_{i,j}^1 - T_{i,j}^0) - \frac{e^{-\lambda^k(T_i+T_j)}\left(\sigma_1^k\right)^2}{2\lambda^k}\left(e^{2\lambda^k T_{i,j}^1} - e^{2\lambda^k T_{i,j}^0}\right) + \\ &\quad + \frac{\sigma_1^k\sigma_2^k\rho_{1,2}^{k,k}\left(e^{-\lambda^k T_i} + e^{-\lambda^k T_j}\right)}{\lambda^k}\left(e^{\lambda^k T_{i,j}^1} - e^{\lambda^k T_{i,j}^0}\right)\end{aligned}$$

Proof. The best way to proceed is to use the so-called polarization identity

$$2\left\langle X_i^k, X_j^m\right\rangle_{T_{i,j}^0}^{T_{i,j}^1} = \left(\left\langle X_i^k + X_j^m\right\rangle_{T_{i,j}^0}^{T_{i,j}^1} - \left\langle X_i^k\right\rangle_{T_{i,j}^0}^{T_{i,j}^1} - \langle X_j^m\rangle_{T_{i,j}^0}^{T_{i,j}^1}\right) \tag{4}$$

(which will valid also for inter-commodity covariations), where the only missing thing here is $\left\langle X_i^k + X_j^k\right\rangle_{T_{i,j}^0}^{T_{i,j}^1}$: in order to calculate this, first we obtain the stochastic differential of $X_i^k + X_j^k$ as

$$d\left(X_i^k + X_j^k\right) = \left(e^{-\lambda^k T_i} + e^{-\lambda^k T_j}\right) e^{\lambda^k t}\sigma_1^k dW_1^k(t) + 2\sigma_2^k dW_2^k(t) + \text{drift}$$

The variation $\left\langle X_i^k + X_j^k \right\rangle_{T_{i,j}^0}^{T_{i,j}^1}$ is then equal to

$$\begin{aligned}
\left\langle X_i^k + X_j^k \right\rangle_{T_{i,j}^0}^{T_{i,j}^1} &= \int_{T_{i,j}^0}^{T_{i,j}^1} \left(e^{-\lambda^k T_i} + e^{-\lambda^k T_j}\right)^2 e^{2\lambda^k t} \left(\sigma_1^k\right)^2 + 4\left(\sigma_2^k\right)^2 + 4\sigma_1^k \sigma_2^k \rho_{1,2}^{k,k} \left(e^{-\lambda^k T_i} + e^{-\lambda^k T_j}\right) e^{\lambda^k t}\, dt \\
&= \frac{\left(e^{-\lambda^k T_i} + e^{-\lambda^k T_j}\right)^2 \left(\sigma_1^k\right)^2}{2\lambda^k} \left(e^{2\lambda^k T_{i,j}^1} - e^{2\lambda^k T_{i,j}^0}\right) + 4\left(\sigma_2^k\right)^2 \left(T_{i,j}^1 - T_{i,j}^0\right) + \\
&\quad + \frac{4\sigma_1^k \sigma_2^k \rho_{1,2}^{k,k} \left(e^{-\lambda^k T_i} + e^{-\lambda^k T_j}\right)}{\lambda^k} \left(e^{\lambda^k T_{i,j}^1} - e^{\lambda^k T_{i,j}^0}\right)
\end{aligned}$$

By putting all together, the result is

$$\begin{aligned}
2\left\langle X_i^k, X_j^k \right\rangle_{T_{i,j}^0}^{T_{i,j}^1} &= 2\left(\sigma_2^k\right)^2 \left(T_{i,j}^1 - T_{i,j}^0\right) - \frac{e^{-\lambda^k (T_i + T_j)} \left(\sigma_1^k\right)^2}{\lambda^k} \left(e^{2\lambda^k T_{i,j}^1} - e^{2\lambda^k T_{i,j}^0}\right) + \\
&\quad + \frac{2\sigma_1^k \sigma_2^k \rho_{1,2}^{k,k} \left(e^{-\lambda^k T_i} + e^{-\lambda^k T_j}\right)}{\lambda^k} \left(e^{\lambda^k T_{i,j}^1} - e^{\lambda^k T_{i,j}^0}\right)
\end{aligned}$$

which gives the desired result □

As already pointed out, our strategy is to have the model quadratic covariations $\left\langle X_i^k, X_j^k \right\rangle_{T_{i,j}^0}^{T_{i,j}^1}$ as close as possible to the market quadratic covariations, which are estimated using the **realized variation estimators**

$$\overline{\langle X_i^k \rangle_{T_i^0}^{T_i^1}} := \sum_{j=1}^{n} \left(X_i^k(t_{j+1}) - X_i^k(t_j) \right)^2 \tag{5}$$

and the **realized covariation estimators**

$$\overline{\left\langle X_i^k, X_j^m \right\rangle_{T_{i,j}^0}^{T_{i,j}^1}} := \sum_{l=1}^{n} \left(X_i^k(t_{l+1}) - X_i^k(t_l) \right) \left(X_j^m(t_{l+1}) - X_j^m(t_l) \right) \tag{6}$$

(which in this section we will use only with $k = m$). Ideally, we would impose that

$$\overline{\left\langle X_i^k, X_j^k \right\rangle_{T_{i,j}^0}^{T_{i,j}^1}} = \left\langle X_i^k, X_j^k \right\rangle_{T_{i,j}^0}^{T_{i,j}^1} \qquad \text{for all } i, j = 1, \ldots, N_k$$

However, the second terms of this system depend only on the four parameters $p^k = (\sigma_1^k, \sigma_2^k, \lambda^k, \rho^k)$, so the system is likely to be overdetermined for $N_k > 2$. For this reason, we estimate the four parameters with a mean-square estimation, i.e. define $\hat{p}^k$ as the 4-dimensional vector which solves

$$\min_{p^k} \sum_{i,j=1}^{N_k} \left(\overline{\left\langle X_i^k, X_j^k \right\rangle_{T_{i,j}^0}^{T_{i,j}^1}} - \left\langle X_i^k, X_j^k \right\rangle_{T_{i,j}^0}^{T_{i,j}^1} \right)^2$$

In this way we obtain all the parameters p^k for all the single commodities, while the inter-commodity correlations $(\rho_{a,b}^{k,m})_{a,b=1,2,k\neq m}$ still remain to be estimated.

3 Calibration of the intercommodity correlations

In order to calibrate for the intercommodity correlations, we continue to use the idea of using the quadratic covariations among the log-forward prices X_k^i for $k = 1, \dots, K$ and $i = 1, \dots, N_k$. In fact, for all suitable i, j, k, m, the quadratic covariations of X_k^i, X_m^j is given by

$$\langle X_i^k, X_j^m \rangle_{t_0}^t := \lim_{n\to\infty} \sum_{l=1}^{n} (X_i^k(t_{l+1}) - X_i^k(t_l))(X_j^m(t_{l+1}) - X_j^m(t_l)) = \int_{t_0}^{t} \Sigma_{i,j}^{k,m}(u)\, du$$

As before, the middle term of these equalities can be estimated with historical observations, while the last term is explicitly computable, in a slightly more complex way than the previous case. In fact, as done in the single commodity case, the best way to calculate it is via the polarization inequality (4), which leads us to calculate first $\left\langle X_i^k + X_j^m \right\rangle_t$.

Lemma 3.1 *We have*

$$\left\langle X_i^k + X_j^m \right\rangle_{T_{i,j}^0}^{T_{i,j}^1} = \int_{T_{i,j}^0}^{T_{i,j}^1} \left(\Sigma_{i,j}^{k,m}(t)\right)^2 dt = \int_{T_{i,j}^0}^{T_{i,j}^1} \Theta_{i,j}^{k,m} R^{k,m} \left(\Theta_{i,j}^{k,m}\right)^T dt$$

where

$$\Theta_{i,j}^{k,m} = \left(\begin{array}{cccc} e^{-\lambda^k (T_i - t)} \sigma_1^k, & \sigma_2^k, & e^{-\lambda^m (T_j - t)} \sigma_1^m, & \sigma_2^m \end{array} \right)$$

and

$$R^{k,m} = \left(\begin{array}{cc} \boldsymbol{\rho}^{k,k} & \boldsymbol{\rho}^{k,m} \\ \boldsymbol{\rho}^{m,k} & \boldsymbol{\rho}^{m,m} \end{array} \right) = \left(\begin{array}{cccc} 1 & \rho_{1,2}^{k,k} & \rho_{1,1}^{k,m} & \rho_{1,2}^{k,m} \\ \rho_{1,2}^{k,k} & 1 & \rho_{2,1}^{k,m} & \rho_{2,2}^{k,m} \\ \rho_{1,1}^{k,m} & \rho_{2,1}^{k,m} & 1 & \rho_{1,2}^{m,m} \\ \rho_{1,2}^{k,m} & \rho_{2,2}^{k,m} & \rho_{1,2}^{m,m} & 1 \end{array} \right)$$

Proof.

$$d\left(X_i^k + X_j^m\right) = \Theta_{i,j}^{k,m}\, dW^{k,m}(t) + \text{ drift} \tag{7}$$

where $W^{k,m}(t) := \left(W_1^k(t), W_2^k(t), W_1^m(t), W_2^m(t)\right)^T$ results in a Gaussian process with independent stationary increments, zero mean and self-correlation matrix given by $R^{k,m}$.

In order to calculate the quadratic variation of $X_i^k + X_j^m$, we now want to represent $W^{k,m}$ as a linear function of a 4-dimensional Brownian motion $\bar{W}^{k,m}$, i.e. $W^{k,m} = \Lambda^{k,m}\bar{W}^{k,m}$ (where the components of $\bar{W}^{k,m}$ are independent 1-dimensional Brownian motions), then we have $R^{k,m} = \Lambda^{k,m}(\Lambda^{k,m})^T$. We can choose to perform a Cholesky decomposition, so that $\Lambda^{k,m}$ can be taken as a lower triangular matrix: in fact, since $R^{k,m}$ is semipositive definite, it can be written as $R^{k,m} = L^{k,m}D^{k,m}\left(L^{k,m}\right)^T$, with $L^{k,m}$ unitary and lower triangular and $D^{k,m}$ diagonal; we can then let $\widetilde{\Lambda}^{k,m} := L^{k,m}\left(D^{k,m}\right)^{\frac{1}{2}}$, with $\left(D^{k,m}\right)^{\frac{1}{2}}$ the matrix having the diagonal elements which are square roots of those of $D^{k,m}$, we have that

$$\widetilde{\Lambda}^{k,m}\left(\widetilde{\Lambda}^{k,m}\right)^T = L^{k,m}\left(D^{k,m}\right)^{\frac{1}{2}}\left(L^{k,m}\left(D^{k,m}\right)^{\frac{1}{2}}\right)^T = L^{k,m}D^{k,m}\left(L^{k,m}\right)^T = R^{k,m}$$

Then,

$$d\left(X_i^k + X_j^m\right) = \Theta_{i,j}^{k,m}\tilde{\Lambda}^{k,m}d\bar{W}^{k,m} + \text{ drift}$$

so that

$$\left\langle X_i^k + X_j^m\right\rangle_{T_{i,j}^0}^{T_{i,j}^1} = \int_{T_{i,j}^0}^{T_{i,j}^1} \Theta_{i,j}^{k,m}R^{k,m}\left(\Theta_{i,j}^{k,m}\right)^T dt$$

□

The integrand, in extended form, is given by

$$\begin{aligned}
\Theta_{i,j}^{k,m}R^{k,m}\left(\Theta_{i,j}^{k,m}\right)^T &= \left(\sigma_1^k\right)^2 e^{-2\lambda^k(T_i-t)} + 2\left(\sigma_2^k\rho_{1,2}^{k,k} + \sigma_2^m\rho_{1,2}^{k,m}\right)\sigma_1^k e^{-\lambda^k(T_i-t)} + \\
&+ \left(\sigma_1^m\right)^2 e^{-2\lambda^m(T_j-t)} + 2\left(\sigma_2^k\rho_{2,1}^{k,m} + \sigma_2^m\rho_{1,2}^{m,m}\right)\sigma_1^m e^{-\lambda^m(T_j-t)} + \\
&+ 2\sigma_1^m\sigma_1^k\rho_{1,1}^{k,m}e^{-\lambda^kT_i-\lambda^mT_j}e^{\left(\lambda^k+\lambda^m\right)t} + 2\sigma_2^m\sigma_2^k\rho_{2,2}^{k,m} + \sigma_2^k + \sigma_2^m
\end{aligned}$$

This results in

$$\begin{aligned}
\left\langle X_i^k + X_j^m\right\rangle_{T_{i,j}^0}^{T_{i,j}^1} &= \frac{\left(\sigma_1^k\right)^2\left(e^{-2\lambda^k\left(T_i-T_{i,j}^1\right)} - e^{-2\lambda^k\left(T_i-T_{i,j}^0\right)}\right)}{2\lambda^k} + \\
&+ \frac{2\left(\sigma_2^k\rho_{1,2}^{k,k} + \sigma_2^m\rho_{1,2}^{k,m}\right)\sigma_1^k\left(e^{-\lambda^k\left(T_i-T_{i,j}^1\right)} - e^{-\lambda^k\left(T_i-T_{i,j}^0\right)}\right)}{\lambda^k} + \\
&+ \frac{\left(\sigma_1^m\right)^2\left(e^{-2\lambda^m\left(T_j-T_{i,j}^1\right)} - e^{-2\lambda^m\left(T_j-T_{i,j}^0\right)}\right)}{2\lambda^m} + \\
&+ \frac{2\left(\sigma_2^k\rho_{2,1}^{k,m} + \sigma_2^m\rho_{1,2}^{m,m}\right)\sigma_1^m\left(e^{-\lambda^m\left(T_j-T_{i,j}^1\right)} - e^{-\lambda^m\left(T_j-T_{i,j}^0\right)}\right)}{\lambda^m} + \\
&+ \frac{2\sigma_1^m\sigma_1^k\rho_{1,1}^{k,m}e^{-\lambda^kT_i-\lambda^mT_j}\left(e^{\left(\lambda^k+\lambda^m\right)T_{i,j}^1} - e^{\left(\lambda^k+\lambda^m\right)T_{i,j}^0}\right)}{\lambda^k+\lambda^m} + \\
&+ \left(2\sigma_2^m\sigma_2^k\rho_{2,2}^{k,m} + \sigma_2^k + \sigma_2^m\right)\left(T_{i,j}^1 - T_{i,j}^0\right)
\end{aligned}$$

Plugging this into the polarization identity (4), we obtain

$$\left\langle X_i^k, X_j^m \right\rangle_{T_{i,j}^0}^{T_{i,j}^1} = \rho_{1,2}^{k,m} A_{i,j}^{k,m} + \rho_{2,1}^{k,m} B_{i,j}^{k,m} + \rho_{1,1}^{k,m} C_{i,j}^{k,m} + \rho_{2,2}^{k,m} D^{k,m}$$

where

$$\begin{aligned}
A_{i,j}^{k,m} &:= \frac{\sigma_2^m \sigma_1^k \left(e^{-\lambda^k \left(T_i - T_{i,j}^1\right)} - e^{-\lambda^k \left(T_i - T_{i,j}^0\right)} \right)}{\lambda^k} \\
B_{i,j}^{k,m} &:= \frac{\sigma_2^k \sigma_1^m \left(e^{-\lambda^m \left(T_j - T_{i,j}^1\right)} - e^{-\lambda^m \left(T_j - T_{i,j}^0\right)} \right)}{\lambda^m} \\
C_{i,j}^{k,m} &:= \frac{\sigma_1^m \sigma_1^k \left(e^{-\lambda^k \left(T_i - T_{i,j}^1\right) - \lambda^m \left(T_j - T_{i,j}^1\right)} - e^{-\lambda^k \left(T_i - T_{i,j}^0\right) - \lambda^m \left(T_j - T_{i,j}^0\right)} \right)}{\lambda^k + \lambda^m} \\
D_{i,j}^{k,m} &:= \sigma_2^m \sigma_2^k \left(T_{i,j}^1 - T_{i,j}^0 \right)
\end{aligned}$$

are known from the calibration of the previous section, and the $\rho_{a,b}^{k,m}$ are still to be estimated. As before, one should aim to solve the linear system

$$\begin{aligned}
\overline{\left\langle X_i^k, X_j^m \right\rangle_{T_{i,j}^0}^{T_{i,j}^1}} &= \rho_{1,2}^{k,m} A_{i,j}^{k,m} + \rho_{2,1}^{k,m} B_{i,j}^{k,m} + \rho_{1,1}^{k,m} C_{i,j}^{k,m} + \rho_{2,2}^{k,m} D_{i,j}^{k,m} \\
& \qquad \forall k, m \in \{1, \ldots, K\} \qquad \forall i \in N_k \qquad \forall j \in N_m
\end{aligned}$$

which, as before, is overdetermined as soon as $|N_k| \times |N_m| > 4$. Thus, again we estimate the $\rho_{a,b}^{k,m}$ with a mean-square estimation, i.e. define the $\rho_{a,b}^{k,m}$ as the minimizers of the problem

$$\min_{\rho_{a,b}^{k,m}} \sum_{i,j,(k \neq m)} \left(\rho_{1,2}^{k,m} A_{i,j}^{k,m} + \rho_{2,1}^{k,m} B_{i,j}^{k,m} + \rho_{1,1}^{k,m} C_{i,j}^{k,m} + \rho_{2,2}^{k,m} D_{i,j}^{k,m} - \overline{\left\langle X_i^k, X_j^m \right\rangle_{T_{i,j}^0}^{T_{i,j}^1}} \right)^2 \quad (8)$$

Remark 3.1 If one minimizes over the original correlations $\rho_{a,b}^{k,m}$, then one must impose, besides $\rho_{a,b}^{k,m} \in [-1, 1]$, that the global correlation matrix $\boldsymbol{\rho}$ is semipositive definite, which is computationally very demanding. An alternative way is to make, similarly to what done in Lemma 3.1, a Cholesky decomposition of $\boldsymbol{\rho}$: this allows to not impose the positive semidefiniteness of the global correlation matrix $\boldsymbol{\rho}$. This will be done more in details in Section 5.

4 An alternative calibration

Now we present an alternative calibration method, which is used among practitioners, but has the fault that, to be rigorous, works on an approximation of the original model. This method is based on the use of the so-called **rolling time series**. Assume from now on, as is quite realistic for those commodities which do not have forward contracts with long deliveries traded in the market, that the maturities $T_1,\ldots,T_N$ are consecutive ends of months. Then the method of rolling time series consists in taking the forward contract with maturity month T_i and treating it, in the current month, as if its volatility were constant (and thus approximately equal to $\Sigma_k(s,T_i)$ with s a suitable point in the current month). When the current month ends and the next begins, take these observations and paste it to the observations of the forward with maturity month T_{i+1}: in this way, we obtain a time series of a forward contract with more or less the same relative maturity.

This method can be made rigorous by redefining the model as

$$\frac{dF^k(t,T)}{F^k(t,T)} = e^{-\frac{\lambda^k}{12}\lceil 12(T-t)\rceil}\sigma_1^k dW_1^k(t) + \sigma_2^k dW_2^k(t) \qquad \forall k=1,\ldots,K \tag{9}$$

If we, as before, denote $X_i^k(t) := \log F^k(t,T_i)$, then we have that

$$\bar{X}_i^k(t_1,t_2) := X_i^k(t_2) - X_i^k(t_1) = \int_{t_1}^{t_2} \sigma_1^k e^{-\frac{\lambda^k}{12}\lceil 12(T-s)\rceil} dW_1^k(s) + \int_{t_1}^{t_2} \sigma_2^k dW_2^k(s) + \text{ drift} \tag{10}$$

where "drift" denotes a quantity which is deterministic both under the risk-neutral probability as well as the real world probability (but of course possibly different). Thus, if we have an equispaced grid $t_1 < \ldots < t_\ell$, with $t_{l+1} - t_l \equiv \Delta$ in a given month, then $(\bar{X}_i^k(t_l,t_{l+1}))_{l=1,\ldots,\ell-1}$ are i.i.d. Gaussian random variables with variance

$$\Sigma_{i,i}^{k,k} = \left(\sigma_1^k\right)^2 e^{-2\lambda^k T_i}\Delta + 2\rho_{1,2}^{k,k}\sigma_1^k\sigma_2^k e^{-\lambda^k T_i}\Delta + \left(\sigma_2^k\right)^2\Delta \tag{11}$$

(recall that, being the T_i ends of months, one has $\frac{1}{12}\lceil 12T_i\rceil = T_i$) and the same applies when we extend this to the rolling time series in the following months. Moreover, if we take two different maturities T_i, T_j, then the two sequences of Gaussian random variables $(\bar{X}_i^k(t_l,t_{l+1}))_{l=1,\ldots,\ell-1}$ and $(\bar{X}_j^k(t_l,t_{l+1}))_{l=1,\ldots,\ell-1}$ have covariance given by

$$\begin{aligned} &\text{Cov}\,(\bar{X}_i^k(t_l,t_{l+1}), \bar{X}_j^m(t_l,t_{l+1})) = \Sigma_{i,j}^{k,k} := \\ &:= \left(\sigma_1^k\right)^2 e^{-\lambda^k(T_i+T_j)}\Delta + \sigma_1^k\sigma_2^k\rho_{1,2}^{k,k}\left(e^{-\lambda^k T_i} + e^{-\lambda^k T_j}\right)\Delta + \left(\sigma_2^k\right)^2\Delta \end{aligned} \tag{12}$$

Finally, if we take two different maturities T_i, T_j, then the two sequences of Gaussian random variables $(\bar{X}_i^k(t_l,t_{l+1}))_{l=1,\ldots,\ell-1}$ and $(\bar{X}_j^m(t_l,t_{l+1}))_{l=1,\ldots,\ell-1}$ have covariance given by

$$\begin{aligned} &\text{Cov}\,(\bar{X}_i^k(t_l,t_{l+1}), \bar{X}_j^m(t_l,t_{l+1})) = \Sigma_{i,j}^{k,m} := \Delta\times \\ &\times \left[\rho_{1,1}^{k,m}\sigma_1^k\sigma_1^m(e^{-\lambda^k T_i} + e^{-\lambda^m T_j}) + \rho_{1,2}^{k,m}\sigma_1^k\sigma_2^m e^{-\lambda^k T_i} + \rho_{2,1}^{k,m}\sigma_2^k\sigma_1^m e^{-\lambda^m T_j} + \rho_{2,2}^{k,m}\sigma_2^k\sigma_2^m\right] \end{aligned} \tag{13}$$

These model variances and covariances can be estimated using the standard estimators

$$\bar{\Sigma}_{i,j}^{k,m} := s_{\bar{X}_i^k,\bar{X}_j^m} = \frac{\sum_l \bar{X}_i^k(t_l,t_{l+1})\bar{X}_j^m(t_l,t_{l+1})}{n} - \frac{\sum_l \bar{X}_i^k(t_l,t_{l+1})}{n}\frac{\sum_l \bar{X}_j^m(t_l,t_{l+1})}{n} \tag{14}$$

where n is the number of contemporary realizations of the time series $(\bar{X}_i^k(t_l, t_{l+1}))_l$ and $(\bar{X}_j^m(t_l, t_{l+1}))_l$. Define then $\bar{\boldsymbol{\Sigma}}^{k,m}$ as

$$\bar{\boldsymbol{\Sigma}}^{k,m} := \left(\bar{\Sigma}_{i,j}^{k,m}\right)_{i\leqslant N_k, j\leqslant N_m} = \begin{pmatrix} \bar{\Sigma}_{1,1}^{k,m} & \cdots & \bar{\Sigma}_{1,N_m}^{k,m} \\ \vdots & \ddots & \vdots \\ \bar{\Sigma}_{N_k,1}^{k,m} & \cdots & \bar{\Sigma}_{N_k,N_m}^{k,m} \end{pmatrix}$$

and $\bar{\boldsymbol{\Sigma}}$, which will be our realized covariance matrix, as

$$\bar{\boldsymbol{\Sigma}} \left(\ \bar{\boldsymbol{\Sigma}}^{k,m} \ \right)_{k,m\leqslant K} = \begin{pmatrix} \bar{\boldsymbol{\Sigma}}^{1,1} & \cdots & \bar{\boldsymbol{\Sigma}}^{K,1} \\ \vdots & \ddots & \vdots \\ \bar{\boldsymbol{\Sigma}}^{1,K} & \cdots & \bar{\boldsymbol{\Sigma}}^{K,K} \end{pmatrix}$$

This has to be compared to the model covariance matrix $\boldsymbol{\Sigma}$, defined as

$$\boldsymbol{\Sigma} \ := \ \left(\ \boldsymbol{\Sigma}^{k,m} \ \right)_{k,m\leqslant K} = \begin{pmatrix} \boldsymbol{\Sigma}^{1,1} & \cdots & \boldsymbol{\Sigma}^{K,1} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}^{1,K} & \cdots & \boldsymbol{\Sigma}^{K,K} \end{pmatrix}$$

where

$$\boldsymbol{\Sigma}^{k,m} := \left(\Sigma_{i,j}^{k,m}\left(p^{k,m}\right)\right)_{i\leqslant N_k, j\leqslant N_m} \quad = \quad \begin{pmatrix} \Sigma_{1,1}^{k,m} & \cdots & \Sigma_{1,N_m}^{k,m} \\ \vdots & \ddots & \vdots \\ \Sigma_{N_k,1}^{k,m} & \cdots & \Sigma_{N_k,N_m}^{k,m} \end{pmatrix}$$

As in the previous sections, one is tempted to let

$$\boldsymbol{\Sigma}(\boldsymbol{p}) = \bar{\boldsymbol{\Sigma}}$$

which is, as usual, overdetermined. We thus proceed as in the previous calibrations: first of all we estimate all the parameters for each commodity $k = 1, \ldots, K$ separately, by making a least-square estimation in the usual way:

$$\min_{p^k} \sum_{i,j=1}^{N_k} \left(\Sigma_{i,j}^{k,k}\left(p^k\right) - \overline{\Sigma}_{i,j}^{k,k}\right)^2$$

Once that the $p^k = (\sigma_1^k, \sigma_2^k, \lambda^k, \rho^k)$ have been estimated, they are kept fixed and the second calibration is performed, again by least-squares, as

$$\min_{\rho_{a,b}^{k,m}} \sum_{k\neq m} \sum_{i=1}^{N_k} \sum_{j=1}^{N_m} \left(\Sigma_{i,j}^{k,m} - \overline{\Sigma}_{i,j}^{k,m}\right)^2 \tag{15}$$

which gives the intercommodity correlations $\rho_{a,b}^{k,m}$, $a, b = 1, 2$, $k \neq m$.

Remark 4.1 As in Section 3, here too it is convenient to work with the Cholesky decomposition of the matrix $\mathbf{\Sigma}$: in this way, analogously with what happens in Remark 3.1, one has the same number of coefficients (in fact, $\mathbf{\Sigma}$ is symmetric and its Cholesky square root is lower triangular with the same dimension), but one has the constraint of $\mathbf{\Sigma}$ being positive semidefinite which is automatically satisfied. As in the previous method, this is done with more details in the following Section 5.

5 Calibration of the intercommodity correlation matrix

Both the calibrations based on the quadratic variation-covariation approach (Sections 2 and 3) as well as on the variance-covariance of rolling time series (Section 4) are based on the two-steps procedure: first calibrate the 4 parameters $p^k = (\sigma_1^k, \sigma_2^k, \lambda^k, \rho^k)$ for each commodity $k = 1, \ldots, K$, and then calibrate the intercommodity correlations $\rho_{a,b}^{k,m}$ for $a, b = 1, 2$ and $k \neq m$. This second step must be self-consistent, in the sense that the resulting global correlation matrix $\boldsymbol{\rho}$, defined in Equation (1), must be nonnegative definite, being the correlation matrix of a $2K$-dimensional Brownian motion.

As reported in Remarks 3.1 and 4.1, if one imposes this constraint naively on the functions to be minimized in Equations (8) and (15), the resulting problem is highly nonlinear in its constraints, thus very time consuming as soon as $K > 2$.

A more clever way to formulate the semi-definiteness constraint is based, as anticipated in both 3.3 and 4.1, on the Cholesky decomposition of $\boldsymbol{\rho}$. Recall that the Cholesky decomposition states that, being $\boldsymbol{\rho}$ semipositive definite, it can be written as $\boldsymbol{\rho} = WW^T$ for a suitable square lower-triangular matrix W.

The advantage of the Cholesky decomposition is that, by calling W_i the i-th row of W for $i = 1, \ldots, 2K$, we can express the constraints on $\boldsymbol{\rho}$ via bilinear constraints on the W_i. More in details, the fact that the principal diagonal of $\boldsymbol{\rho}$ has unitary elements is translated into the condition

$$\|W_i\|^2 = 1 \qquad \forall i = 1, \ldots, 2K, \tag{16}$$

where $\|\cdot\|$ denotes the Euclidean norm.

The fact that the elements of $\boldsymbol{\rho}$ which correspond to a ρ^k which was already calibrated (i.e. $\rho_{1,2}^{k,k} = \rho_{2,1}^{k,k} = \rho^k$) must be taken as already assigned is translated into

$$W_{2k-1} W_{2k}^T = \rho^k \qquad \forall i = 1, \ldots, K, \tag{17}$$

Finally, the fact that $|\rho_{1,2}^{k,m}| \leqslant 1$ follows from the Cauchy-Schwarz inequality for Euclidean norm and scalar product in ^{2K} and from Equation (16): in fact,

$$|\rho_{1,2}^{k,m}| = |W_{2k-1} W_{2m}^T| \leqslant \sqrt{\|W_{2k-1}\| \cdot \|W_{2m}^T|} = 1$$

and the heaviest constraint, i.e. the semipositive definiteness of $\boldsymbol{\rho}$, is automatically satisfied by the very definition of W.

Now, the functions to be minimized in Equations (8) and (15) can be written, in a more abstract form, as

$$G(\boldsymbol{\rho}) := \|\boldsymbol{\Xi}\boldsymbol{\rho} - \mathbf{E}\|_2^2$$

with $\boldsymbol{\Xi}$ and $\mathbf{E}$ suitably defined in the two problems. This minimization problem translates into

$$\min_W \|\boldsymbol{\Xi} W W^T - \mathbf{E}\|_2^2$$

where W varies over the space of all lower-triangular matrices with the constraints in Equations (16-17). This is a quadratic problem with quadratic constraints, which is numerically time-efficient and quite stable.

6 Empirical findings

In order to test the two methods, we simulate daily prices of 36 futures of a single commodity, where maturities are equispaced with a 1-month interval. The parameters that we impose are $\sigma_1 = \sigma_2 = 0.02$, $\lambda = 0.04$, $\rho = 0.3$.

The green circles represent the square of the local volatility structure calculated with the true parameters, while the red curve represents the estimate with the covariation method (Sections 2-3) and the blue curve represents the estimate with the method of rolling time series (Section 4). The results can be seen in Figure 1.

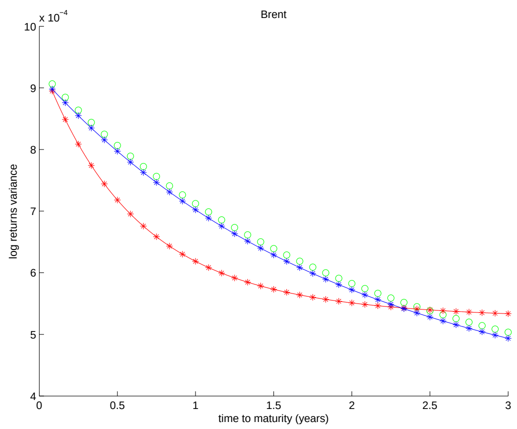

Figure 1: Log-return variance, i.e. squared local volatility, with respect to time to maturity, daily simulation. The green circles represent the square of the local volatility structure calculated with the true parameters, while the red curve represents the estimate with the covariation method (Sections 2-3) and the blue curve represents the estimate with the method of rolling time series (Section 4).

We can see that the rolling time series method gives quite a good fit, while the covariation method gives a fit which is quite far from the real volatility shape. The reason for this misbehaviour could be that the quadratic covariation needs a limit to be performed, while we only have a finite number of observation.

Of course, the more the interval between observations becomes thinner, the more the estimators that we use come near to the theoretical quadratic covariations. For this reason, we do another simulation, with the same parameters, but now with 200 observations per day. This results in a much better fit for the covariation method, but it is also evident that the method based on rolling time series gives now a perfect fit.

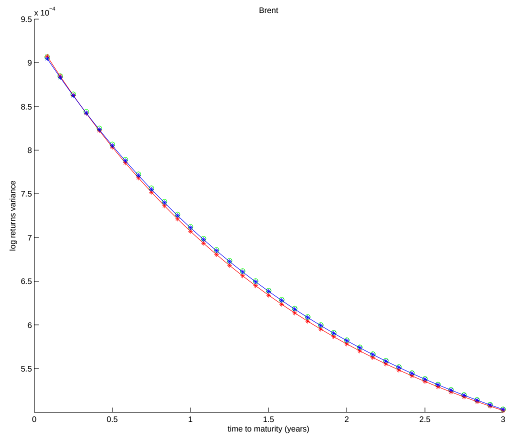

Figure 2: Log-return variance, i.e. squared local volatility, with respect to time to maturity, 200 simulations per day. The colors are as in Figure 1.

This is actually bad news, at least for the quadratic covariation approach. In fact, it is true that some commodities (e.g. Brent) have a number of transactions on some forward contracts which allow this daily number of observations to be performed. However it is also true that, for maturities from 9 months on, exchanges of forward contracts are less frequent. This would result in covariation estimators performed on real data to give many zeroes in the estimators of Equations (5–6), resulting in a bias towards zero: this is known as the Epps effect [3]. One way to circumvent this would be to use estimators which prevent the Epps effects in asynchronous observations, such as for example the Fourier estimators studied in [5].

7 Conclusions

We present a multicommodity model for forward prices which extends the single-commodity model presented by Kiesel *et al.* [6]. We show two calibration methods based on time series of past forward prices, which can be used when liquid derivatives are not traded in the market. The first calibration method, presented in Sections 2 and 3, is based on the quadratic variations and covariations of log-prices, which are analytically computable, while the second method, presented in Section 4, uses the idea of rolling time series, but requires a modification of the model to be used exactly. Both the methods require to estimate first four parameters per commodity, and then a global intercommodity correlation matrix. For this last step, it is numerically convenient to express the global correlation matrix via its Cholesky decomposition: this results in a quadratic minimization problem with quadratic constraints, which is numerically tractable, as detailed in Section 5. Finally in Section 6 we test the two methods against simulated data, and find out that the first method performs poorly when dealing with daily data, while the second method gives already a good fit at this time scale. We then test the two methods with high-frequency simulated data (200 observations per day): the first method now performs much better, but still the second method is the best. The conclusion is that the rolling series method seems to perform well at very different time scales, while the first one needs high-frequency data to produce reliable results. This can be perhaps improved with more sophisticated estimators, like those presented in [5], but we leave this improvement to a future work.

References

[1] F. E. Benth, N. Lange, T. A. Myklebust. *Pricing and hedging quanto options in energy markets*. EnergyFinance2012 Conference, 2012

[2] T. Bjork. Arbitrage Theory in Continuous Time. Oxford University Press, 3^{rd} edition, 2009.

[3] T. Epps. *Comovements in stock prices in the very short run*. Journal of the American Statistical Association **74**, 291–298, 1979.

[4] H. Geman. Commodities and commodity derivatives: modelling and pricing for agriculturals, metals, and energy. Wiley finance series. John Wiley & Sons , 2005.

[5] M. E. Mancino, S. Sanfelici. *Estimating covariance via Fourier method in the presence of asynchronous trading and microstructure noise*. Journal of Financial Econometrics **9 (2)**, 367–408, 2011.

[6] R. Kiesel, G. Schindlmayr, R. H. Börger. A two-factor model for the electricity forward market. Quantitative Finance **9 (3)**, 2009.

[7] B. Øksendal, *Stochastic differential equations*, 5^{th} ed., Springer, 2000.

[8] E. Schwartz, J. E. Smith, *Short-Term Variations and Long-Term Dynamics in Commodity Prices*, Management Science **46 (7)**, 2000