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 3, 2013

Abstract

We propose a model for the evolution of forward prices of several commodities, which is an extension of the factor forward model in [3, 10], to a market where multiple commodities are traded. We calibrate this model in a market where forward contracts on multiple commodities are present, using historical forward prices. First we calibrate separately the four coefficients of every single commodity, using an approach based on quadratic variation/covariation of forward prices. Then, with the same technique, we pass to estimate the mutual correlation among the Brownian motions driving the different commodities. This calibration is compared to a calibration method used by practitioners, which uses rolling time series and requires a modification of the model, but turns out to be more accurate in practice, especially with a low frequency of observed transaction. We present efficient methods to perform the calibration with both methods, as well as the calibration of the intercommodity correlation matrix. Then we calibrate our model to WTI, ICE Brent and ICE Gasoil forward prices. Finally we present how to estimate spot volatility from forward parameters, with an application to the WTI spot volatility.

Keywords: two-factor model for commodity forward prices, historical calibration, quadratic variation/covariation, rolling time series, non-convex optimization, semidefinite programming.
2010 MSC Classification: 60H35, 62H20, 62M05, 90C22, 90C26, 91B25

1 Introduction

When dealing with forward prices of a single commodity having different maturities, the factor model proposed in [3, 10] is quite simple to understand, analytically tractable

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

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 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 [12], (see [4] 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

$$dF^k(t,T) = F^k(t,T)(e^{-\lambda^k(T-t)}\sigma_1^k \, dW_1^k(t) + \sigma_2^k \, dW_2^k(t)) \quad 0 \leqslant t < T$$

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

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

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

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

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

This model is an extension of the model in [10] to a multicommodity framework, and also an extension of the model in [3], where only two commodities are taken into account. The model is analytically tractable because, under the forward measure $\mathbb{Q}_T$, each $F^k(\cdot, T)$ has a lognormal evolution, given by

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

where $\Sigma^k(s, T)$ is a sort of local volatility at time s, defined as

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

Thus, conditional to the information up to time t_0, $\log F^k(t, T)$ has a Gaussian distribution, with mean

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

and variance

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

In this paper we want to calibrate this model in a situation where, for each commodity $k = 1, \dots, K$, forward contracts with (a finite number of) different maturities $T^k_1, \dots, T^k_{N_k}$ 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-commodity products. Secondly, we estimate the correlation matrix also in the inter-commodity correlations.

More in details, 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 efficiently calibrate the single-commodity parameters in an efficient way in both procedures: in both cases, this results in a non-convex problem in 4 variables, which can be brought to a non-convex problem in 1 variable. In Section 6 we show how to perform the inter-commodity calibration of the global correlation matrix $\boldsymbol{\rho}$ in a way which is numerically efficient, based on the Cholesky decomposition: in doing this, we solve a non-standard semidefinite programming problem. In Section 7 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). In Section 8 we calibrate the model to WTI, ICE Brent and ICE Gasoil, while in Section 9 we show how to estimate the spot volatility from forward parameters, presenting a numerical estimation on the WTI spot. Section 10 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. Being k fixed, 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

$$dX_i^k(t) = e^{-\lambda^k(T_i - t)} \sigma_1^k \, dW_1^k(t) + \sigma_2^k \, dW_2^k(t) + \text{ drift} \quad 0 \leqslant t < T_i$$

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 the same mutual correlation ρ^k, but the drift in the two dynamics are possibly 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

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

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, for $t_0 < t \leqslant T_i$, the quadratic variation of X_i^k under $\mathbb{P}$ is given by

$$\langle X_i^k \rangle_{t_0}^t := \lim_{n \to \infty} \sum_{l=1}^{n} (X_i^k(t_{l+1}) - X_i^k(t_l))^2 = \int_{t_0}^{t} (\Sigma_i^k(u))^2 \, du \tag{3}$$

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

$$\langle X_i^k, X_j^k \rangle_{t_0}^t := \lim_{n \to \infty} \sum_{l=1}^{n} (X_i^k(t_{l+1}) - X_i^k(t_l))(X_j^k(t_{l+1}) - X_j^k(t_l)) = \int_{t_0}^{t} \Sigma_{i,j}^k(u)\, du \qquad (4)$$

(for more details, see [11]). 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 (3–4).

Lemma 2.1 *For $t_0 < t \leqslant T_i$, the quadratic variation of the process X_i^k is given by*

$$\begin{aligned} \langle X_i^k \rangle_{t_0}^t &= \int_{t_0}^{t} (\Sigma_i^k(u))^2\, du = \\ &= \frac{\left(\sigma_1^k\right)^2}{2\lambda^k} \left(e^{-2\lambda^k(T_i - t)} - e^{-2\lambda^k(T_i - t_0)}\right) + \left(\sigma_2^k\right)^2 (t - t_0) + \\ &\quad + \frac{2\sigma_1^k \sigma_2^k \rho^k}{\lambda^k} \left(e^{-\lambda^k(T_i - t)} - e^{-\lambda^k(T_i - t_0)}\right) \end{aligned}$$

Proof. See Appendix. □

Lemma 2.2 *For $t_0 < t \leqslant \min(T_i, T_j)$, the quadratic covariation of the processes X_i^k, X_j^k is given by*

$$\begin{aligned} \langle X_i^k, X_j^k \rangle_{t_0}^t &= \left(\sigma_2^k\right)^2 (t - t_0) - \frac{e^{-\lambda^k(T_i + T_j)} \left(\sigma_1^k\right)^2}{2\lambda^k} \left(e^{2\lambda^k t} - e^{2\lambda^k t_0}\right) + \\ &\quad + \frac{\sigma_1^k \sigma_2^k \rho^k \left(e^{-\lambda^k T_i} + e^{-\lambda^k T_j}\right)}{\lambda^k} \left(e^{\lambda^k t} - e^{\lambda^k t_0}\right) \end{aligned}$$

Proof. See Appendix. □

As already pointed out, our strategy is to have the model quadratic covariations as close as possible to the market quadratic covariations. More in details, at a generic observation time t, for each pair of maturities T_i, T_j the maximal interval in $[t_0, t]$ where we can define the quadratic covariation between X_i^k and X_j^k is given by $[T_{i,j}^0, T_{i,j}^1]$, where

$$T_{i,j}^0 := \max(t_0, \tilde{T}_i, \tilde{T}_j) \quad \text{and} \quad T_{i,j}^1 := \min(t, T_i, T_j) \tag{5}$$

and, for all $i = 1, \ldots, N$, the time $\tilde{T}_i$ is the official time from when the forward $F(t, T_i)$ can be traded. Then the model quadratic covariation will be $\langle X_i^k, X_j^k \rangle_{T_{i,j}^0}^{T_{i,j}^1}$, to be compared with the market quadratic covariation.

In order to estimate for the latter ones, for each $i, j = 1, \ldots, N$, we use the **realized variation estimators**

$$\overline{\langle X_i^k \rangle_{T_{i,j}^0}^{T_{i,j}^1}} := \sum_{j=1}^{n} \Big(X_i^k\left(t_{j+1}\right) - X_i^k\left(t_j\right) \Big)^2 \tag{6}$$

and the **realized covariation estimators**

$$\overline{\langle X_i^k, X_j^m \rangle_{T_{i,j}^0}^{T_{i,j}^1}} := \sum_{l=1}^{n} \Big(X_i^k\left(t_{l+1}\right) - X_i^k\left(t_l\right) \Big) \left(X_j^m\left(t_{l+1}\right) - X_j^m\left(t_l\right) \right) \tag{7}$$

(which in this section we will use only with $k = m$). It is a standard result [2] that these estimators are unbiased and consistent provided that the drifts of X_i^k and X_j^m are zero. Here this is not the case as the drifts are not zero, but it is possible to prove, by adapting results in [2], that these estimators are biased (but with the bias depending on the drift only on third order), consistent and asymptotically Gaussian.

Ideally, we would impose that

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

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 as follows.

Definition 2.1 *We define the estimator $\hat{p}^k$ as the 4-dimensional vector which solves*

$$\min_{p^k} \sum_{i,j=1}^{N_k} \left(\overline{\langle X_i^k, X_j^k \rangle_{T_{i,j}^0}^{T_{i,j}^1}} - \langle X_i^k, X_j^k \rangle_{T_{i,j}^0}^{T_{i,j}^1} \right)^2 \tag{8}$$

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. Unfortunately, though mean-square estimation is a classical in statistics, with this definition for $\hat{p}^k$ is very hard to prove properties like unbiasedness and consistency, due to the non-linear dependence of $\langle X_i^k, X_j^k \rangle_{T_{i,j}^0}^{T_{i,j}^1}$ on p^k (especially on λ^k). It is also true that classical recipes usually employed to obtain good estimators (like maximum likelihood, for example) here result in computations which are impossible to carry out to obtain explicit estimators. Thus, the compromise here is to use a classical recipe, like mean-squares estimations, to obtain some estimators.

Remark 2.1 The problem in Equation (8) is in principle a non-convex optimization problem in the four parameters $p^k = (\sigma_1^k, \sigma_2^k, \lambda^k, \rho^k)$, which can be numerically instable and with many local minima. In Section 5 we show how it is possible to reduce it to a 2-step optimization problem, where the first step is a quadratic optimization and the second step is a non-convex problem in one variable.

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 $k = 1, \ldots, K$ and $i = 1, \ldots, N_k$. In fact, for all suitable i, j, k, m, for $t_0 < t \leqslant \min(T_i^k, T_j^m)$, 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, as shown in the following lemmas.

Lemma 3.1 *For $t_0 < t \leqslant \min(T_i^k, T_j^m)$, we have*

$$\left\langle X_i^k + X_j^m \right\rangle_{t_0}^{t} = \int_{t_0}^{t} \left(\Sigma_{i,j}^{k,m}(t)\right)^2 dt = \int_{t_0}^{t} \Theta_{i,j}^{k,m} R^{k,m} \left(\Theta_{i,j}^{k,m}\right)^T dt$$

where

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

and

$$R^{k,m} = \begin{pmatrix} \boldsymbol{\rho}^{k,k} & \boldsymbol{\rho}^{k,m} \\ \boldsymbol{\rho}^{m,k} & \boldsymbol{\rho}^{m,m} \end{pmatrix} = \begin{pmatrix} 1 & \rho_{1,2}^{k,k} & \rho_{1,1}^{k,m} & \rho_{1,2}^{k,m} \\ \rho_{1,2}^{k,k} & 1 & \rho_{2,1}^{k,m} & \rho_{2,2}^{k,m} \\ \rho_{1,1}^{k,m} & \rho_{2,1}^{k,m} & 1 & \rho_{1,2}^{m,m} \\ \rho_{1,2}^{k,m} & \rho_{2,2}^{k,m} & \rho_{1,2}^{m,m} & 1 \end{pmatrix}$$

Proof. We have that

$$d\left(X_i^k + X_j^m\right) = \Theta_{i,j}^{k,m}\, dW^{k,m}(t) + \text{ drift} \tag{9}$$

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 $\hat{W}^{k,m}$, i.e. $W^{k,m} = \Lambda^{k,m}\hat{W}^{k,m}$ (where the components of $\hat{W}^{k,m}$ 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

$$\left\langle X_i^k + X_j^m\right\rangle_{t_0}^t = \int_{t_0}^t \Theta_{i,j}^{k,m}R^{k,m}\left(\Theta_{i,j}^{k,m}\right)^T dt$$

□

Lemma 3.2 *For $t_0 < t \leqslant \min(T_i^k, T_j^m)$, the quadratic covariation of X_i^k, X_j^m is given by*

$$\left\langle X_i^k, X_j^m\right\rangle_{t_0}^t = \rho_{1,2}^{k,m}A_{i,j}^{k,m} + \rho_{2,1}^{k,m}B_{i,j}^{k,m} + \rho_{1,1}^{k,m}C_{i,j}^{k,m} + \rho_{2,2}^{k,m}D^{k,m}$$

where

$$\begin{aligned}
A_{i,j}^{k,m} &:= \frac{\sigma_2^m\sigma_1^k\left(e^{-\lambda^k\left(T_i^k-t\right)} - e^{-\lambda^k\left(T_i^k-t_0\right)}\right)}{\lambda^k} \\
B_{i,j}^{k,m} &:= \frac{\sigma_2^k\sigma_1^m\left(e^{-\lambda^m\left(T_j^m-t\right)} - e^{-\lambda^m\left(T_j^m-t_0\right)}\right)}{\lambda^m} \\
C_{i,j}^{k,m} &:= \frac{\sigma_1^m\sigma_1^k\left(e^{-\lambda^k\left(T_i^k-t\right)-\lambda^m\left(T_j^m-t\right)} - e^{-\lambda^k\left(T_i^k-t_0\right)-\lambda^m\left(T_j^m-t_0\right)}\right)}{\lambda^k+\lambda^m} \\
D_{i,j}^{k,m} &:= \sigma_2^m\sigma_2^k(t-t_0)
\end{aligned}$$

Proof. See Appendix. □

As in the previous section, our strategy is to have the model quadratic covariations as close as possible to the market quadratic covariations. More in details, at a generic observation time t, for each pair of maturities T_i^k, T_j^m the maximal interval in $[t_0, t]$ where we can define the quadratic covariation between X_i^k and X_j^m is now $[T_{i,j}^{0,k,m}, T_{i,j}^{1,k,m}]$, where

$$T_{i,j}^{0,k,m} := \max(t_0, \tilde{T}_i^k, \tilde{T}_j^m) \quad \text{and} \quad T_{i,j}^{1,k,m} := \min(t, T_i^k, T_j^m)$$

where, for all $k = 1, \ldots, K$ and $i = 1, \ldots, N_k$, the time $\tilde{T}_i^k$ is the official time from when the forward $F^k(t, T_i)$ can be traded. Then again we compare the market quadratic covariations, estimated as in Equation (7), with the model quadratic covariations $\langle X_i^k, X_j^k \rangle_{T_{i,j}^{0,k,m}}^{T_{i,j}^{1,k,m}}$, where the $A_{i,j}^{k,m}$, $B_{i,j}^{k,m}$, $C_{i,j}^{k,m}$ and $D_{i,j}^{k,m}$, for all the k, m, i, j used in the calibration, are known from the calibration of the previous section, while $\rho_{a,b}^{k,m}$ are still to be estimated. As before, one should aim to solve the linear system

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

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 as follows.

Definition 3.1 *Define the $\hat{\rho}_{a,b}^{k,m}$ as the minimizers of the problem*

$$\min_{\rho_{a,b}^{k,m}} \sum_{i,j,(k \neq m)} \left(\rho_{1,2}^{k,m} A_{i,j}^{k,m} + \rho_{2,1}^{k,m} B_{i,j}^{k,m} + \rho_{1,1}^{k,m} C_{i,j}^{k,m} + \rho_{2,2}^{k,m} D_{i,j}^{k,m} - \overline{\left\langle X_i^k, X_j^m \right\rangle_{T_{i,j}^{0,k,m}}^{T_{i,j}^{1,k,m}}} \right)^2 \tag{10}$$

subject to the constraint that the global correlation matrix $\boldsymbol{\rho}$, defined in Equation (1), must be symmetric and semipositive definite, and to the other natural constraints

$$\begin{cases} \rho_{a,b}^{k,m} \in [-1, 1] & \forall k, m = 1, \ldots, K, \forall a, b = 1, 2, \\ \rho_{1,2}^{k,k} = \rho^k & \forall k = 1, \ldots, K, \\ \rho_{1,1}^{k,k} = \rho_{2,2}^{k,k} = 1 & \forall k = 1, \ldots, K, \end{cases} \tag{11}$$

In this way we obtain all the inter-commodity correlations $(\hat{\rho}_{a,b}^{k,m})_{a,b=1,2,k \neq m}$, which still were to be estimated. Unfortunately we again find ourselves in the same situation as with $\hat{p}^k$, i.e. with this definition for $\hat{\rho}_{a,b}^{k,m}$ is very hard to prove properties like unbiasedness and consistency, due this time to the nonlinear constraints on $\hat{\rho}_{a,b}^{k,m}$, in particular the semipositive definiteness of the global matrix $\boldsymbol{\rho}$. Also here, classical recipes like maximum likelihood result in computations which are impossible to carry out to obtain explicit estimators. Thus, the compromise here has been again to obtain some estimators through a classical recipe.

Remark 3.1 If one minimizes over the original correlations $\rho_{a,b}^{k,m}$, then the constraint for the global correlation matrix $\boldsymbol{\rho}$ to be semipositive definite is computationally very demanding; in fact, the problem results in non-standard semidefinite programming problem. 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 6.

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 maturities traded in the market, that the maturities $T_1,\ldots,T_N$ are the same for all the commodities and 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. A refinement of this method can be found for example in [1], but the version presented here is the most used among practitioners.

This method can be made rigorous by redefining the model as

$$\frac{dF^k(t,T_i)}{F^k(t,T_i)} = e^{-\frac{\lambda^k}{12}\lceil 12(T_i-t)\rceil}\sigma_1^k dW_1^k(t) + \sigma_2^k dW_2^k(t) \quad 0 \leqslant t < T_i \tag{12}$$

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

$$\bar{X}_i^k(t_1,t_2) := X_i^k(t_2) - X_i^k(t_1) = \int_{t_1}^{t_2} \sigma_1^k e^{-\frac{\lambda^k}{12}\lceil 12(T_i-s)\rceil} dW_1^k(s) + \int_{t_1}^{t_2} \sigma_2^k dW_2^k(s) + \text{ drift} \tag{13}$$

where "drift" denotes a quantity which is deterministic under the risk-neutral probability and is possibly non-deterministic under the real world probability, containing in this latter case also the market price of risk.

We now assume that the market price of risk is deterministic and stationary in time. Then, 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

$$\Sigma_{i,i}^{k,k} = \left(\sigma_1^k\right)^2 e^{-2\lambda^k T_i}\Delta + 2\rho^k\sigma_1^k\sigma_2^k e^{-\lambda^k T_i}\Delta + \left(\sigma_2^k\right)^2\Delta \tag{14}$$

(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 $(\bar{X}_i^k(t_l,t_{l+1}))_{l=1,\dots,\ell-1}$ and $(\bar{X}_j^k(t_l,t_{l+1}))_{l=1,\dots,\ell-1}$ have covariance given by

$$\begin{aligned} \mathrm{Cov}\,(\bar{X}_i^k(t_l,t_{l+1}),\bar{X}_j^k(t_l,t_{l+1})) = \Sigma_{i,j}^{k,k} :=& \\ := \quad \left(\sigma_1^k\right)^2 e^{-\lambda^k(T_i+T_j)}\Delta + \sigma_1^k\sigma_2^k\rho^k\left(e^{-\lambda^k T_i}+e^{-\lambda^k T_j}\right)\Delta + \left(\sigma_2^k\right)^2\Delta \end{aligned} \tag{15}$$

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

$$\begin{aligned} &\mathrm{Cov}\,(\bar{X}_i^k(t_l,t_{l+1}),\bar{X}_j^m(t_l,t_{l+1})) = \Sigma_{i,j}^{k,m} := \\ &:= \quad \Delta\left[\rho_{1,1}^{k,m}\sigma_1^k\sigma_1^m e^{-\lambda^k T_i-\lambda^m T_j} + \rho_{1,2}^{k,m}\sigma_1^k\sigma_2^m e^{-\lambda^k T_i} + \rho_{2,1}^{k,m}\sigma_2^k\sigma_1^m e^{-\lambda^m T_j} + \rho_{2,2}^{k,m}\sigma_2^k\sigma_2^m\right] \end{aligned} \tag{16}$$

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

$$\bar{\Sigma}_{i,j}^{k,m} := s_{\bar{X}_i^k,\bar{X}_j^m} = \frac{\sum_l \bar{X}_i^k(t_l,t_{l+1})\bar{X}_j^m(t_l,t_{l+1})}{n} - \frac{\sum_l \bar{X}_i^k(t_l,t_{l+1})}{n}\frac{\sum_l \bar{X}_j^m(t_l,t_{l+1})}{n} \tag{17}$$

where n is the number[1] of contemporary realizations of the time series $(\bar{X}_i^k(t_l,t_{l+1}))_l$ and $(\bar{X}_j^m(t_l,t_{l+1}))_l$. It is a standard result that these estimators are unbiased and consistent.

Define then $\bar{\mathbf{\Sigma}}^{k,m}$ as

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

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

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

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

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

[1]in order to be rigorous, one should define $T_{i,j}^{0/1}$ as in the previous sections: we choose not to do so, as it would make the notation heavier and we believe that all is clear from the context.

where

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

As in the previous sections, one is tempted to let

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

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:

$$\min_{p^k} \sum_{i,j=1}^{N_k} \left(\Sigma_{i,j}^{k,k}\left(p^k\right) - \overline{\Sigma}_{i,j}^{k,k}\right)^2 \tag{18}$$

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

$$\min_{\rho_{a,b}^{k,m}} \sum_{k\neq m} \sum_{i=1}^{N_k} \sum_{j=1}^{N_m} \left(\Sigma_{i,j}^{k,m} - \overline{\Sigma}_{i,j}^{k,m}\right)^2 \tag{19}$$

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

Remark 4.1 As in Section 2, also the optimization problem in Equation (18) is a non-convex optimization problem in the four parameters $p^k = (\sigma_1^k, \sigma_2^k, \lambda^k, \rho^k)$, which can be numerically instable and with many local minima. In Section 5 we show how it is possible to reduce also this to a 2-step optimization problem in the same way as in Remark 2.1.

Remark 4.2 As in Section 3, here too the calibration problem results in non-standard semidefinite programming problem. Thus, it is convenient to work with the Cholesky decomposition of the matrix $\boldsymbol{\Sigma}$: in this way, analogously with what happens in Remark 3.1, one has the same number of coefficients (in fact, $\boldsymbol{\Sigma}$ is symmetric and its Cholesky square root is lower triangular with the same dimension), but one has the constraint of $\boldsymbol{\Sigma}$ being positive semidefinite which is automatically satisfied. As in the previous method, this is done with more details in the following Section 6.

5 Single commodity 2-step calibration

As already noticed in Remarks 2.1 and 4.1, the optimization problems (8) and (18), needed to perform the mean-square calibration in the single commodities, are non-convex optimization problem in the four parameters $p^k = (\sigma_1^k, \sigma_2^k, \lambda^k, \rho^k)$. Here we show how they can be reduced to a 2-step optimization problem which has the problems (8) and (18) as particular cases, and whose 2 steps are numerically easier and more stable to implement. More in details, the first step is a quadratic optimization and the second step is a non-convex problem in only one variable.

We now formulate the general optimization problem, of which (8) and (18) can be seen as two particular cases. As we are working only on a single commodity k, we will omit the dependence on k in the sequel and write N instead of N_k and $(\sigma_1, \sigma_2, \lambda, \rho)$ instead of $(\sigma_1^k, \sigma_2^k, \lambda^k, \rho_{1,2}^{k,k})$. The general problem can be formulated as

$$\min_{\sigma_1,\sigma_2,\lambda\in\mathbb{R}^+,\rho\in[-1,1]} \sum_{i=1}^{N}\sum_{j=1}^{N}(a_{i,j}\sigma_1^2 + b_{i,j}\sigma_2^2 + c_{i,j}\rho\sigma_1\sigma_2 - \bar{X}_{i,j})^2 \tag{20}$$

We can obtain problem (8) by letting

$$\begin{aligned}
a_{i,j} &= \frac{e^{-\lambda\left(T_i+T_j-2T_{i,j}^1\right)} - e^{-\lambda\left(T_i+T_j-2T_{i,j}^0\right)}}{2\lambda} \\
b_{i,j} &= T_{i,j}^1 - T_{i,j}^0 \\
c_{i,j} &= \frac{e^{-\lambda\left(T_i-T_{i,j}^1\right)} - e^{-\lambda\left(T_i-T_{i,j}^0\right)} + e^{-\lambda\left(T_j-T_{i,j}^1\right)} - e^{-\lambda\left(T_j-T_{i,j}^0\right)}}{\lambda}, \\
\bar{X}_{i,j} &= \overline{\left\langle X_i^k, X_j^k \right\rangle_{T_{i,j}^0}^{T_{i,j}^1}}
\end{aligned}$$

with $T_{i,j}^0$ and $T_{i,j}^1$ defined in Eq. (5). Problem (18) can be obtained by letting

$$\begin{aligned}
a_{i,j} &= e^{-\lambda(T_i+T_j)}\Delta \\
b_{i,j} &= \Delta, \\
c_{i,j} &= \left(e^{-\lambda T_i} + e^{-\lambda T_j}\right)\Delta, \\
\bar{X}_{i,j} &= \bar{\Sigma}_{i,j}
\end{aligned}$$

where $\bar{\Sigma}_{i,j}$ stands for $\bar{\Sigma}_{i,j}^{k,k}$ as defined in Eq. (17). In both specifications, as well as in the general formulation (20), we omit the dependence on λ of $a_{i,j}$, $b_{i,j}$ and $c_{i,j}$ for ease of notation.

Now we solve the general problem (20) in two steps. The first step consists in fixing a $\lambda > 0$ and making a change of variables $v_1 := \sigma_1^2$, $v_2 := \sigma_2^2$, $v_3 := \rho\sigma_1\sigma_2$. Then problem (20) can be written as

$$\min_{v_1,v_2\in\mathbb{R}^+,|v_3|\leqslant\sqrt{v_1v_2}} \sum_{i=1}^{N}\sum_{j=1}^{N}(a_{i,j}v_1 + b_{i,j}v_2 + c_{i,j}v_3 - \bar{X}_{i,j})^2 \tag{21}$$

which is a quadratic problem in v_1, v_2, v_3, to be solved for $v_1, v_2 \in \mathbb{R}^+$ (as they are variances) and $|v_3| \leqslant \sqrt{v_1 v_2}$ (as $|\rho| \leqslant 1$). The first order conditions turn out to be

$$\begin{cases} \sum_{i,j} \left(a_{i,j}v_1 + b_{i,j}v_2 + c_{i,j}v_3 - \bar{X}_{i,j}\right) a_{i,j} = 0 \\ \sum_{i,j} \left(a_{i,j}v_1 + b_{i,j}v_2 + c_{i,j}v_3 - \bar{X}_{i,j}\right) b_{i,j} = 0 \\ \sum_{i,j} \left(a_{i,j}v_1 + b_{i,j}v_2 + c_{i,j}v_3 - \bar{X}_{i,j}\right) c_{i,j} = 0 \end{cases}$$

where $\sum_{i,j}$ stands for $\sum_{i=1}^{N} \sum_{j=1}^{N}$. By collecting terms in the new variables $v := (v_1, v_2, v_3)^T$, these conditions can be rewritten as $\Pi v = p$, with

$$\begin{pmatrix} \sum_{i,j}\left(a_{i,j}^2\right) & \sum_{i,j}\left(a_{i,j}b_{i,j}\right) & \sum_{i,j}\left(a_{i,j}c_{i,j}\right) \\ \sum_{i,j}\left(a_{i,j}b_{i,j}\right) & \sum_{i,j}\left(b_{i,j}^2\right) & \sum_{i,j}\left(b_{i,j}c_{i,j}\right) \\ \sum_{i,j}\left(a_{i,j}c_{i,j}\right) & \sum_{i,j}\left(b_{i,j}c_{i,j}\right) & \sum_{i,j}\left(c_{i,j}^2\right) \end{pmatrix} = \Pi, \qquad \begin{pmatrix} \sum_{i,j} a_{i,j}\bar{X}_{i,j} \\ \sum_{i,j} b_{i,j}\bar{X}_{i,j} \\ \sum_{i,j} c_{i,j}\bar{X}_{i,j} \end{pmatrix} = p$$

For both problems (8) and (18), the matrix Π is invertible for each $\lambda > 0$, thus we obtain a unique solution $v = v^\lambda$ for the problem (21).

The second step is now to solve, in the single variable λ, the problem

$$\min_{\lambda \in \mathbb{R}^+} \sum_{i=1}^{N} \sum_{j=i}^{N} (a_{i,j}v_1^\lambda + b_{i,j}v_2^\lambda + c_{i,j}v_3^\lambda - \bar{X}_{i,j})^2$$

which is, in general, non-convex (recall that also $a_{i,j}$, $b_{i,j}$ and $c_{i,j}$ depend on λ).

This two-steps method works very well if v^λ, obtained as the solution of $\Pi v = p$, satisfies $v_1^\lambda, v_2^\lambda \in \mathbb{R}^+$, $|v_3^\lambda| \leqslant \sqrt{v_1^\lambda v_2^\lambda}$ for all $\lambda > 0$. If this is not the case, then v^λ must be found by applying for example the Kuhn-Tucker multipliers method (or other constrained optimization methods), which are slightly more time consuming. In our case, in the calibration example of Section 8, the two-steps method gives results in less than 10 seconds, while the original optimization problem is instead solved in times that are greater of a factor from 10 to 20.

6 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 $k = 1, \ldots, K$ (for which we presented an efficient procedure in Section 5), 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 consistent both with itself as with the first step. More in details, 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. Besides, $\boldsymbol{\rho}$ must have a diagonal with all entries equal to 1 and with some other fixed entries, which are the ones found during the first calibration.

As reported in Remarks 3.1 and 4.2, if one imposes this constraint naively on the functions to be minimized in Equations (10) and (19), one obtains a non-standard semidefinite programming problem, which 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 Remarks 3.1 and 4.2, 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

$$\|W_i\|^2 = 1 \qquad \forall i = 1, \ldots, 2K, \tag{22}$$

where $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^{2K}$.

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

$$W_{2k-1} W_{2k}^T = \rho^k \qquad \forall k = 1, \ldots, K, \tag{23}$$

Finally, the fact that $|\rho^{k,m}_{1,2}| \leqslant 1$ follows from the Cauchy-Schwarz inequality for Euclidean norm and scalar product in $\mathbb{R}^{2K}$ and from Equation (22): in fact,

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

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 (10) and (19) 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

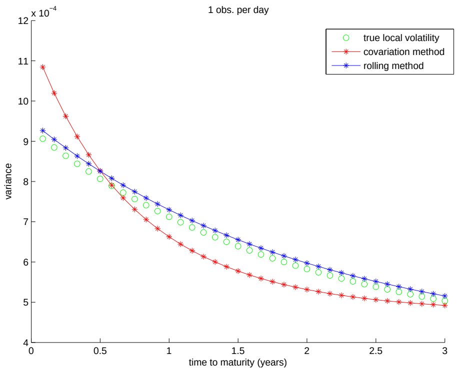
$$\min_W \left\|\boldsymbol{\Xi} W W^T - \mathbf{E}\right\|_2^2$$

where W varies over the space of all lower-triangular matrices with the constraints in Equations (22-23). This is a quadratic problem with quadratic constraints, which is numerically time-efficient and quite stable. In our case, in the calibration example of Section 8, this method gives results in less than 20 seconds.

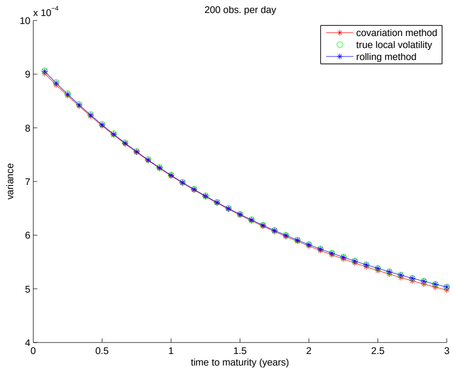
7 A numerical test for the two methods

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 $\sigma_1 = \sigma_2 = 0.02$, $\lambda = 0.04$, $\rho = 0.3$. We then calibrate the model, based on these simulated observations, by using first the method based on the quadratic variations and covariations of Sections 2, and then the method based on rolling time series of Section 4. In both methods, we used the procedures of Section 5 in order to obtain a more tractable optimization problem.

The results of the test can be seen in Figure 1(a), where the green circles represent the square of the local volatility structure $\Sigma_{i,i}^{k,k}$ computed by inserting in Equation (15) 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).



(a) daily observations



(b) 200 observations per day

Figure 1: Log-return variance, i.e. squared local volatility, with respect to time to maturity. 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 instead 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, and report the result in Figure 1(b). 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.

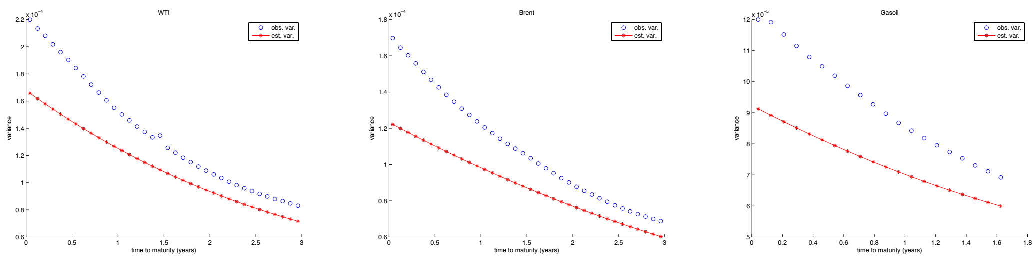
This is actually bad news, at least for the quadratic covariation approach. In fact, it is true that some commodities (e.g. ICE Brent) have a number of transactions on some forward contracts which allow this daily number of observations to be performed. However it is also true that, for maturities from 9 months on, exchanges of forward contracts are less frequent. Moreover, for other commodities (such as power, for instance), there are, in practice, no contracts that allow such high observation frequency.

As a consequence, this would result in covariation estimators performed on real data which give many zeroes in the estimators of Equations (6–7), resulting in a bias towards zero: this is known as the Epps effect [6]. One example of how this would underestimate the volatility is given at the beginning of Section 8. 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 [9]. This however is outside the scopes of this paper.

8 A calibration example

Given the results of the previous section, we decide to calibrate the multicommodity model using the method in Section 4, based on the variance-covariance structure of rolling time series. For the calibration, we selected three of the most liquid commodities in the international markets, namely the WTI (West Texas Intermediate) oil, also known as Texas Light Sweet, the ICE Brent oil, and the ICE Gasoil, with daily observations going from January 3, 2012 to June 3, 2013.

Before showing the real calibration, we present an illustration in Figure 2 on how the calibration with the first method, based on quadratic variations-covariations, effectively underestimates heavily the volatility surface. Our suspect is that this is due to the Epps effect mentioned at the end of Section 7. For this reason, from now on we concentrate ourselves on the calibration done with the rolling time series method.



(a) WTI (b) ICE Brent (c) ICE Gasoil

Figure 2: Log-return variance and covariance with respect to time to maturity (in years) for WTI, ICE Brent and ICE Gasoil, with the variation-covariation method. The red lines represent the estimates with the rolling series method of Section 4, while the blue circles are the empirical variances of the rolling series.

8.1 Single commodity calibrations

As seen in the previous sections, the first thing to do is to calibrate the four parameters p^k of commodity $k = 1, 2, 3$. We represent the results of the calibration in Figure 3 in two ways. The first one is a plot of the variance of rolling series with respect to their maturity, as shown in Figures 3(a), 3(c), 3(e), respectively. The second way is a plot of the cross-covariances of the forward contracts with respect to their maturities, as shown in Figures 3(b), 3(d), 3(f), respectively. This is possible because Equation (15) gives the model covariances $\Sigma_{i,j}^{k,k}$ of the forward log-returns, which can then be compared with the corresponding empirical estimates $\bar{\Sigma}_{i,j}^{k,k}$, with $k = 1, 2, 3$ and with i, j covering maturity spans which are commonly found in the market (3 year for WTI and ICE Brent, and 1 years and 8 months for ICE Gasoil).

More in details, for the WTI in Figures 3(a) and 3(b) we see a very good fit of the estimated model to the empirical variance-covariance structure, apart for the 1.5-years maturity. For the ICE Brent, in Figures 3(c) and 3(d) we notice a fit which is even better in terms of goodness. Finally, for the ICE Gasoil in Figures 3(e) and 3(f) we notice a satisfying fit, apart from the first maturities, probably due to anomalous price movements. All the parameters found have been reported in Table 1.

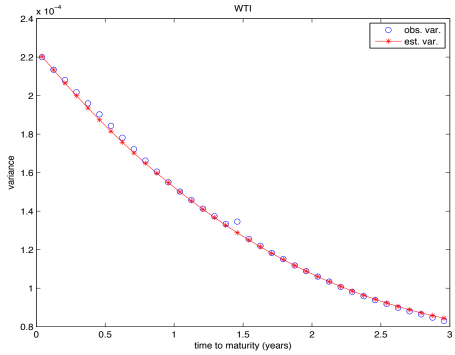
	WTI	ICE Brent	ICE Gasoil
σ_1	0.6892	5.4527	4.1725
σ_2	0.5162	5.2915	4.0473
λ	0.0826	0.0078	0.0105
$\rho_{1,2}$	-0.9628	-0.9997	-0.9996

Table 1: Single commodity calibration results

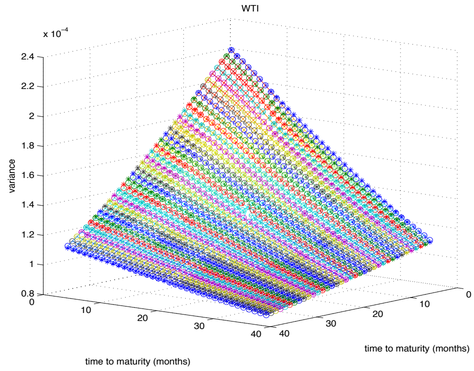
The most relevant remark that can be made here is that the $\bar{\rho}_{1,2}^{k,k}$, for $k = 1, 2, 3$, are all very near to -1: this means that the two factors driving these commodities are almost perfectly correlated, which is consistent with the common empirical findings about crude oil markets, where usually one single factor explains more than 95% of the total variance (see e.g. [7]). One could then argue that a more parsimonious model would have only one Brownian factor per commodity, with the global volatility $e^{-\lambda^k (T_i - t)}\sigma_1^k - \sigma_2^k$. This however could be misleading and not suited for all the possible scenarios.

To see this, we perform a rolling estimation, i.e. take a 60-days estimation horizon and obtain a time series of estimated parameters. A sample of this is reported in Table 2 for the WTI.

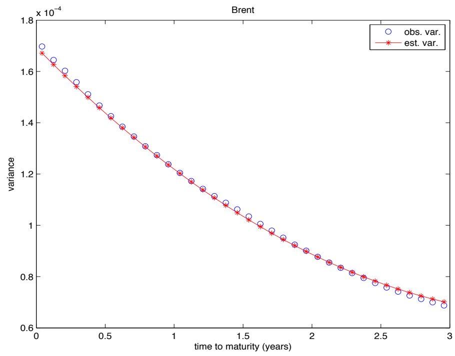
One can see that in some cases ρ can be far from -1, and in that case using a single factor is clearly insufficient. Besides, one can notice that this happens exactly when λ is far from 0, while for values of λ near 0, ρ is near -1. This behaviour could in principle be proved with an asymptotic expansion for $\lambda \to 0$, which is however beyond the scopes of this article. We instead show that this intuition is confirmed numerically when one looks at the joint rolling estimation graphics of λ and ρ, as shown in Figure 4.



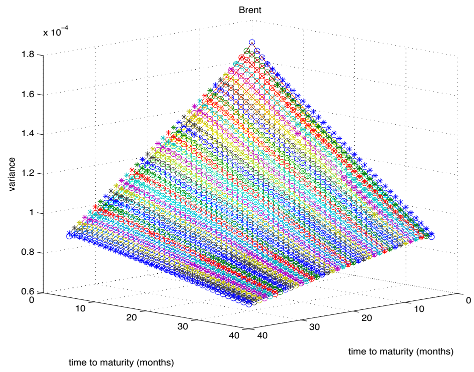
(a) Variance versus maturity, WTI



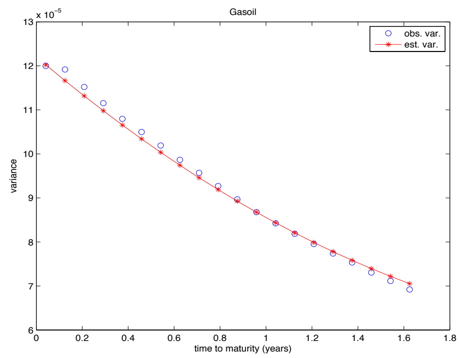
(b) Covariance versus maturity, WTI



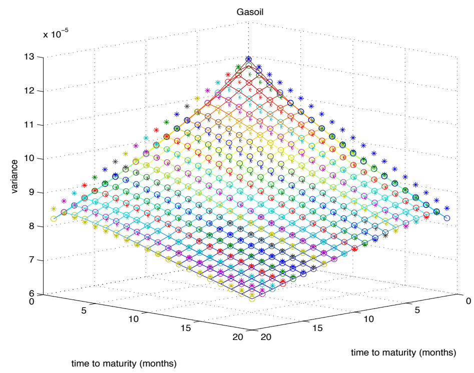
(c) Variance versus maturity, ICE Brent



(d) Covariance versus maturity, ICE Brent



(e) Variance versus maturity, ICE Gasoil



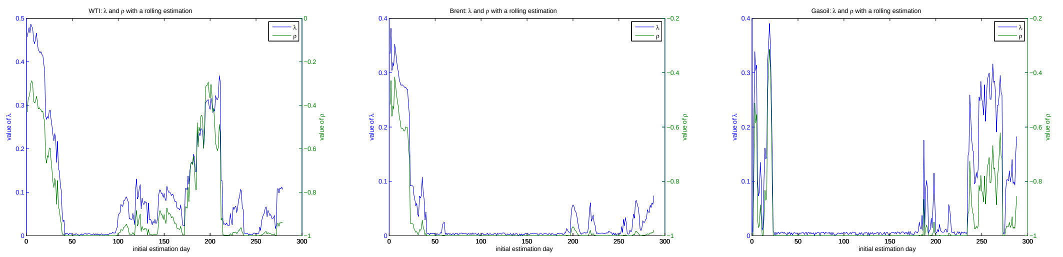
(f) Covariance versus maturity, ICE Gasoil

Figure 3: Log-return variance and covariance with respect to time to maturity (in years) for WTI, ICE Brent and ICE Gasoil. In Figures (a), (c) and (e), the red line represent the estimate with the rolling series method of Section 4, while the blue circles are the empirical variances of the rolling series. In Figures (b), (d) and (f), the lined surface represent the estimate with the rolling series method of Section 4, while the circles are the empirical covariances of the rolling series.

	σ_1	σ_2	λ	ρ
1–60	0.2148	0.1369	0.4943	-0.4151
61–120	6.0590	5.7911	0.0109	-0.9996
121–180	0.6071	0.4228	0.0948	-0.9492
181–240	0.3416	0.2045	0.1662	-0.6728
241–300	5.5947	5.4752	0.0070	-0.9999

Table 2: Rolling estimations on the WTI for some 60-days frames (from day 1 to 60, from day 61 to 120 and so on).

Each point of these graphics has been estimated by taking a 60-days estimation horizon starting from the day in the x-axis on. It is quite evident, for all the three commodities, that when $\lambda \simeq 0$, $\rho \simeq -1$: in these market scenarios one can effectively think that only one factor moves the entire term structure of volatilities. Conversely, when λ is far from 0, also ρ is far from -1, this corresponding to market scenarios where two factors are needed to move the term structure.



(a) WTI (b) ICE Brent (c) ICE Gasoil

Figure 4: Rolling estimations for λ and ρ in the three commodities, in two different scales: in all the three figures, the values for λ are in the left vertical scale, while those for ρ are in the right one.

8.2 Estimation of the global correlation matrix

The second step is to estimate the intercommodity correlation matrix as seen in Section 6. The result is

$$\bar{\boldsymbol{\rho}} = \left(\begin{array}{cc|cc|cc} 1 & -0.9628 & 0.7203 & -0.6734 & 0.4390 & -0.4135 \\ -0.9628 & 1 & -0.7167 & 0.6753 & -0.4306 & 0.4097 \\ \hline 0.7203 & -0.7167 & 1 & -0.9997 & 0.4339 & -0.4214 \\ -0.6734 & 0.6753 & -0.9997 & 1 & -0.4301 & 0.4179 \\ \hline 0.4390 & -0.4306 & 0.4339 & -0.4301 & 1 & -0.9996 \\ -0.4135 & 0.4097 & -0.4214 & 0.4179 & -0.9996 & 1 \end{array}\right)$$

Now some remarks can be formulated after these numerical results.

The first remark is mathematical and is that the $\bar{\rho}^{k,k}_{1,2}$, for $k = 1, 2, 3$, are exactly equal to those found in the single-commodity estimations, consistently with the procedure of Section 6.

A more empirical remark is that the six Brownian motions appear to be correlated among each other through the correlation matrix $\bar{\boldsymbol{\rho}}$, but less than one expects when observing market price comovements of the commodities. For example, it is a well known fact that WTI and ICE Brent prices have a very high correlation, being the same type of crude oil, while this is not evident at all from the corresponding correlations in $\bar{\boldsymbol{\rho}}$, all around ± 0.5. This is because the correlations in $\bar{\boldsymbol{\rho}}$ are those among the driving Brownian motions, and not those among the commodities prices.

In order to obtain those latter correlations, one has to use Equations (14) and (16) to recover the correlations between the (log) prices of two given market instruments: in fact, if we are interested in the correlation between the T_i-month forward of commodity k and the T_j-month forward of commodity m, then this can be easily computed as

$$R^{k,m}_{i,j} := \frac{\Sigma^{k,m}_{i,j}}{\sqrt{\Sigma^{k,k}_{i,i}\Sigma^{m,m}_{j,j}}}$$

We show the result of this calculation for the values $R^{k,m}_{i,i}$ (i.e. correlations of forward prices with the same maturity of commodities k and m) in Figure 5.

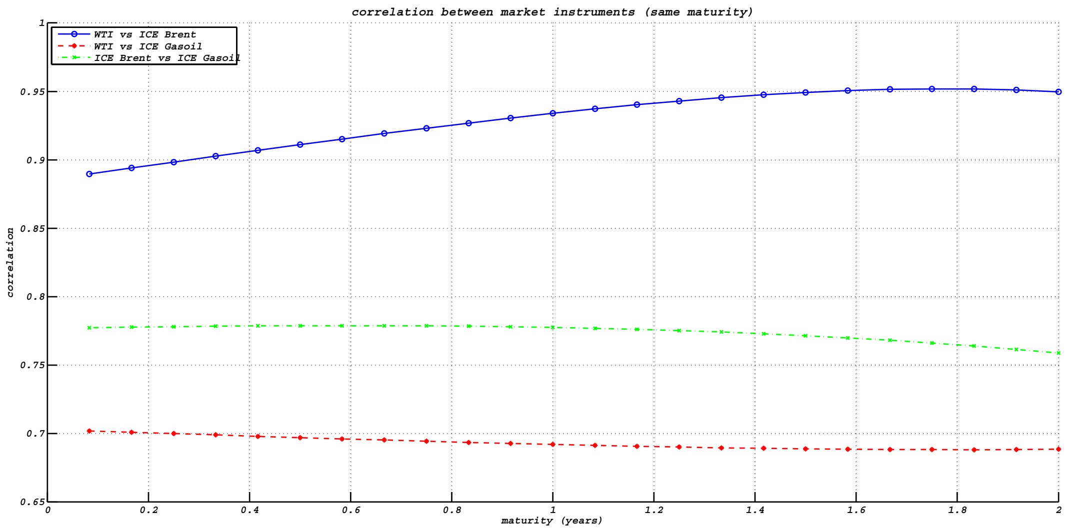


Figure 5: Correlations of forward prices with the same maturity among commodities. WTI versus ICE Brent in solid blue, WTI versus ICE Gasoil in dashed red, ICE Brent versus ICE Gasoil in dotted green.

From this latter figure, we can clearly observe a positive high correlation among all contracts with the same maturity, higher than what one may expect from the correlation matrix $\bar{\rho}$. The economic reason for this high correlation is of course that these forward contracts belong to the same market segment (energy), and in particular they all are oil subproducts. More in details, while the ICE Gasoil forwards appear with a positive correlation with respect to WTI and ICE Brent which is however far from ± 1, the forwards of the two crude oil products, WTI and ICE Brent, appear to have correlations ranging from 0.8897 to 0.9518. This can be explained by the fact that the two products have similar physical characteristics and are often used in financial markets as a proxy of one another.

9 Spot prices

We now present an application of our estimation procedure to the volatility of the spot price, which is classically defined as

$$S^k(t) = \lim_{T \to t} F^k(t, T) \tag{24}$$

From the practical point of view, there are several reason why it is difficult, if not impossible, to estimate spot parameters from the direct observation of the time series S^k, and is instead better to infer them from the forwards' dynamics:

- an official spot market of the commodity could simply not exist (ICE Brent and Gasoil) or could exist in multiple places, so that the "official" spot price is an average of these small markets (WTI);
- spot markets could follow very short term movements (for example, an important buyer which is short of the commodity in that moment), which do not influence forward prices, nor future spot dynamics;
- spot markets usually have seasonality more pronounced than the corresponding forward markets.

A natural question to ask is then whether one can infer the spot volatility from the parameters of forward contracts as obtained in the previous sections. In order to do this, the first thing is to compute the spot volatility from the definition of spot price in Equation (24) and from the model for $F^k(t, T)$.

Starting from Equation (2), for each t and $\Delta > 0$ such that $t + \Delta \leqslant T$ we can write

$$\log F^k(t+\Delta, T) - \log F^k(t, T) = \int_t^{t+\Delta} \sigma_1^k e^{-\lambda^k (T-s)} dW_1^k(s) + \int_t^{t+\Delta} \sigma_2^k dW_2^k(s) + \text{ drift}$$

By taking the limit for $T \to t + \Delta$, we obtain

$$\begin{aligned} &\lim_{T \to t+\Delta} \left(\log \left(F^k(t+\Delta, T) \right) - \log \left(F^k(t, T) \right) \right) = \\ &= \lim_{T \to t+\Delta} \left(\int_t^{t+\Delta} \sigma_1^k e^{-\lambda^k (T-s)} dW_1^k(s) + \int_t^{t+\Delta} \sigma_2^k dW_2^k(s) \right) + \text{ drift} = \\ &= \int_t^{t+\Delta} \sigma_1^k e^{-\lambda^k (t+\Delta-s)} dW_1^k(s) + \int_t^{t+\Delta} \sigma_2^k dW_2^k(s) + \text{ drift} \end{aligned}$$

and, by applying Equation (24), we arrive at

$$\log S^k(t+\Delta) = \log F^k(t,t+\Delta) + \int_t^{t+\Delta} \sigma_1^k e^{-\lambda^k(t+\Delta-s)} dW_1^k(s) + \int_t^{t+\Delta} \sigma_2^k dW_2^k(s) + \text{ drift}$$

Then, conditional to the knowledge up to time t, we have that

$$\begin{aligned}
&\mathrm{Var}_t \left[\log S^k(t+\Delta) - \log S^k(t)\right] = \\
&= \int_t^{t+\Delta} \left(\left(\sigma_1^k\right)^2 e^{-2\lambda^k(t+\Delta-s)} + \left(\sigma_2^k\right)^2 + 2\rho^k \sigma_1^k \sigma_2^k e^{-\lambda^k(t+\Delta-s)} \right) ds = \\
&= \int_0^{\Delta} \left(\left(\sigma_1^k\right)^2 e^{-2\lambda^k(\Delta-s)} + \left(\sigma_2^k\right)^2 + 2\rho^k \sigma_1^k \sigma_2^k e^{-\lambda^k(\Delta-s)} \right) ds = \\
&= \frac{(\sigma_1^k)^2}{2\lambda^k} \left(1 - e^{-2\lambda^k \Delta}\right) + \left(\sigma_2^k\right)^2 \Delta + \frac{2\sigma_1^k \sigma_2^k \rho^k}{\lambda^k} \left(1 - e^{-\lambda^k \Delta}\right) \simeq \\
&\simeq \Delta \left[\left(\sigma_1^k\right)^2 + \left(\sigma_2^k\right)^2 + 2\sigma_1^k \sigma_2^k \rho^k \right] + o(\Delta)
\end{aligned}$$

where the last asymptotics is for $\Delta \to 0$. In this way it is possible to obtain the spot volatility from the parameters of the forward prices model.

With this in mind, we try to see if the above equation succeeds in estimating the spot volatility for the WTI, which has an official spot market. More in details, with the parameters that we estimated in the previous section for WTI we plot the volatility versus maturity also for the maturity 1/252 (corresponding to one day), as seen in Figure 6. The spot volatility estimated with the forward prices' parameters is then $3.06 \cdot 10^{-4}$, while the realized volatility estimated directly from the spot market data is $2.92 \cdot 10^{-4}$, with an error of 4.7%.

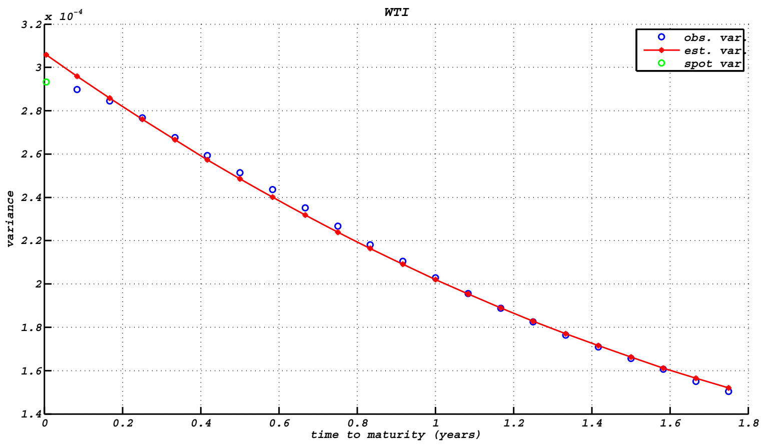


Figure 6: Volatility versus maturity for the WTI. The green dot is the spot volatility.

10 Conclusions

We present a multicommodity model for forward prices which extends the models presented in [3, 10]. 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 methods require a two-steps procedure to estimate the parameters. First one must estimate four parameters per commodity, which requires to solve a non-convex 4-dimensional optimization problem, which is hard to solve numerically: in Section 5 we show how this problem can be brought down to a quadratic 3-dimensional problem and to a non-convex 1-dimensional problem, both of which are numerically more tractable. The second step is to calibrate a global intercommodity correlation matrix: for this last step, it is numerically convenient to express the global correlation matrix via its Cholesky decomposition, which results in a quadratic minimization problem with quadratic constraints, which is numerically tractable, as detailed in Section 6. In Section 7 we test the two methods against simulated data, and conclude 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. In Section 8 we calibrate simultaneously the model to WTI, ICE Brent and ICE Gasoil, and produce four parameters for each of the three commodities above as well as their 6-dimensional correlation matrix. Finally in Section 9 we present an application to the estimation of the WTI spot volatility: by applying the entire forward curve we obtain an estimate with the forward parameters which is 10% far from the realized spot market volatility.

Acknowledgments

The authors wish to thank Fabio Antonelli, Alberto Del Pia, Paolo Foschi, Francesco Rinaldi, Wolfgang J. Runggaldier and all the participants to the Conference on Energy & Finance in Trondheim (Norway), October 4 - 5, 2012 and to the XIV Workshop in Quantitative Finance in Rimini (Italy), January 24 - 25, 2013, together with two anonymous referees, for various comments and suggestions.

Appendix: proofs of Lemmas

Lemma 2.1. **Proof.** We have

$$\begin{aligned}\int_{t_0}^{t}\left(\Sigma_i^k(t)\right)^2 dt &= \int_{t_0}^{t}\left(e^{-2\lambda^k(T_i-t)}\left(\sigma_1^k\right)^2+\left(\sigma_2^k\right)^2+2e^{-\lambda^k(T_i-t)}\sigma_1^k\sigma_2^k\rho^k\right)dt\\ &= \left(\sigma_1^k\right)^2\left[e^{-2\lambda^k(T_i-t)}\right]_{t_0}^{t}+\left(\sigma_2^k\right)^2(t-t_0)+2\sigma_1^k\sigma_2^k\rho^k\left[e^{-\lambda^k(T_i-t)}\right]_{t_0}^{t}\\ &= \frac{\left(\sigma_1^k\right)^2}{2\lambda^k}\left(e^{-2\lambda^k(T_i-t)}-e^{-2\lambda^k(T_i-t_0)}\right)+\left(\sigma_2^k\right)^2(t-t_0)+\\ &\quad+\frac{2\sigma_1^k\sigma_2^k\rho^k}{\lambda^k}\left(e^{-\lambda^k(T_i-t)}-e^{-\lambda^k(T_i-t_0)}\right)\end{aligned}$$

□

Lemma 2.2. **Proof.** This is a particular case of Lemma 3.2 below, which follows immediately by putting $m=k$ and noticing that $\rho_{1,1}^{k,k}=\rho_{1,1}^{m,m}=1$ and $\rho_{1,2}^{k,k}=\rho_{2,1}^{k,k}=\rho^k$. □

Lemma 3.2. **Proof.** The best way to proceed is to use the so-called polarization identity

$$2\langle X_i^k, X_j^m\rangle_{t_0}^t = \langle X_i^k+X_j^m\rangle_{t_0}^t - \langle X_i^k\rangle_{t_0}^t - \left\langle X_j^m\right\rangle_{t_0}^t \tag{25}$$

where the only missing term here is $\langle X_i^k+X_j^m\rangle_{t_0}^t$: in order to calculate this, first we obtain the stochastic differential of $X_i^k+X_j^m$ as

$$\begin{aligned}d\left(X_i^k+X_j^m\right) &= e^{-\lambda^k(T_i^k-t)}\sigma_1^k d\tilde{W}_1^k(t)+\sigma_2^k d\tilde{W}_2^k(t)+\\ &\quad+e^{-\lambda^m(T_j^m-t)}\sigma_1^m d\tilde{W}_1^m(t)+\sigma_2^m d\tilde{W}_2^m(t)+\text{drift}\end{aligned}$$

The variation $\langle X_i^k+X_j^m\rangle_{t_0}^t$ is then equal to

$$\begin{aligned}\left\langle X_i^k+X_j^m\right\rangle_{t_0}^t &= \frac{\left(\sigma_1^k\right)^2\left(e^{-2\lambda^k\left(T_i^k-t\right)}-e^{-2\lambda^k\left(T_i^k-t_0\right)}\right)}{2\lambda^k}+\\ &\quad+\frac{2\left(\sigma_2^k\rho_{1,2}^{k,k}+\sigma_2^m\rho_{1,2}^{k,m}\right)\sigma_1^k\left(e^{-\lambda^k\left(T_i^k-t\right)}-e^{-\lambda^k\left(T_i^k-t_0\right)}\right)}{\lambda^k}+\\ &\quad+\frac{\left(\sigma_1^m\right)^2\left(e^{-2\lambda^m\left(T_j^m-t\right)}-e^{-2\lambda^m\left(T_j^m-t_0\right)}\right)}{2\lambda^m}+\\ &\quad+\frac{2\left(\sigma_2^k\rho_{2,1}^{k,m}+\sigma_2^m\rho_{1,2}^{m,m}\right)\sigma_1^m\left(e^{-\lambda^m\left(T_j^m-t\right)}-e^{-\lambda^m\left(T_j^m-t_0\right)}\right)}{\lambda^m}+\\ &\quad+\frac{2\sigma_1^m\sigma_1^k\rho_{1,1}^{k,m}e^{-\lambda^k T_i^k-\lambda^m T_j^m}\left(e^{\left(\lambda^k+\lambda^m\right)t}-e^{\left(\lambda^k+\lambda^m\right)t_0}\right)}{\lambda^k+\lambda^m}+\\ &\quad+\left(2\sigma_2^m\sigma_2^k\rho_{2,2}^{k,m}+\sigma_2^k+\sigma_2^m\right)(t-t_0)\end{aligned}$$

and by using the polarization identity (25), Lemma 2.1 and the facts that $\rho_{1,2}^{k,k} = \rho^k$ and $\rho_{1,2}^{m,m} = \rho^m$, we have the desired result. □

References

[1] C. Alexander. *Hedging the risk of an energy futures portfolio*. In: Risk management in commodity markets – from shipping to agriculturals and energy. H. Geman (ed.). Wiley Finance series. John Wiley & Sons, 2008

[2] O. E. Barndorff-Nielsen, N. Shephard. *Econometric analysis of realized volatility and its use in estimating stochastic volatility models*. J. R. Statist. Soc. B **64**, Part 2, 253–280, 2002.

[3] G. Benmenzer, E. Gobet, C. Jerusalem. *Arbitrage free cointegrated models in gas and oil future markets*. Technical report, GDF Suez and Laboratoire Jean Kuntzmann, Grenoble, 2007.

[4] F. E. Benth, N. Lange, T. A. Myklebust. *Pricing and hedging quanto options in energy markets*. Energy&Finance2012 Conference, 2012

[5] T. Bjork. *Arbitrage Theory in Continuous Time*. Oxford University Press, 3rd edition, 2009.

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

[7] E. Fanone. *Three Essays on Modelling of Electricity Markets*. Ph.D. Thesis, University of Trieste, 2012.

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

[9] 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.

[10] R. Kiesel, G. Schindlmayr, R. H. Börger. A two-factor model for the electricity forward market. *Quantitative Finance* **9 (3)**, 279–287, 2009.

[11] B. Øksendal, *Stochastic differential equations*, 5th ed., Springer, 2000.

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