Multi-Factor Multi-Commodity Models & Parameter Estimation Processes

John Breslin
Les Clewlow
Chris Strickland
Daniel van der Zee

Lacima Group

March 2008

This article discusses a general multi factor multi commodity (MFMC) model and the process for estimating parameters from historical data. Single factor models have a wide range of applicability in energy valuation and risk management, and are relatively simple to understand and parameterize. However, the simplicity of single factor models can be a double edged sword. While these models can capture much of the dynamics of real life processes in many circumstances, by definition they only use a small amount of the potential information available from the market. In particular one major drawback of single factor models is they imply that instantaneous changes in forward prices at all maturities are perfectly correlated. Increasingly, energy risk practitioners are attracted to modelling frameworks that avoid the above simplifications. Where enough data is available, a more general multi factor model can be used to capture extra information about the price dynamics and this is the modelling framework that we concentrate on in this article. It is also relatively straightforward to extend such a multi factor model to incorporate multiple commodities. In the following we discuss a general MFMC model and describe the process of estimating parameters from historical data.

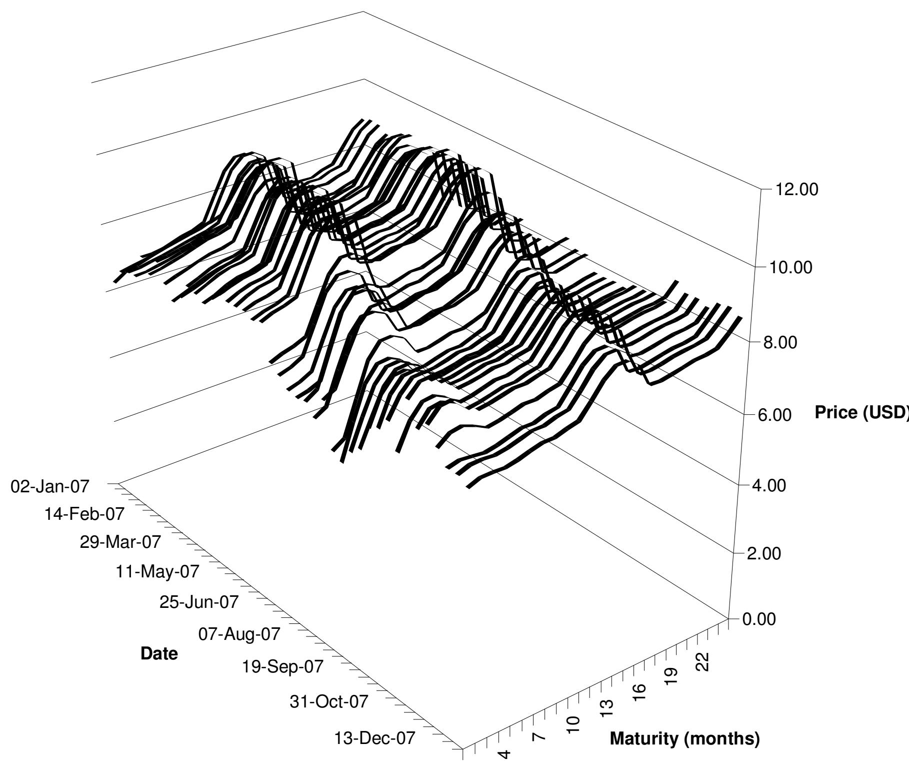

Figure 1: Forward curve for HNG during 2007

Figure 1 illustrates the historical evolution of a forward curve for Henry Hub natural gas (HNG) from 02-Jan-07 to 14-Dec-07. For clarity not all available curves in this period are shown. Each forward curve consists of 24 data points representing the next 24 monthly maturities, i.e. on each calendar date we plot the nearby contract, the second nearby, etc., out to the 24th nearby contract. The first few curves in January 08 therefore contain forward prices for contracts maturing each month from February 2007 to Jan 2009, while the last curves in December 07 contain prices for contracts maturing from January 07 to December 09.

One important observation from Figure 1 is that forward prices of different maturities are not perfectly correlated - the curves generally move up and down together, with the short end of the curve exhibiting more volatility than the long end, but they also change shape in apparently quite complex ways. In order to capture this complex interaction of different points along the forward curve we need more than a single factor of uncertainty.

A general multi-factor model of the forward curve which can be represented by the following stochastic differential equation (SDE);

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

In this formulation $F(t,T)$ denotes a forward price for delivery at time *T* (the maturity date) recorded on date *t*, and there are *n* independent sources of uncertainty which drive the evolution of the forward curve. Each source of uncertainty $dz_i(t)$ has associated with it a volatility function $\sigma_i(t,T)$ which determines by how much, and in which direction, that random shock moves each point of the forward curve. Note that it is possible to write an equation for the dynamics of the spot price that is consistent with this forward price dynamics – this is important in understanding how forward curve and spot dynamics are related, and explains the link between many popular implementations of equation (1) and some well know spot price models. We can integrate equation (1), set the maturity date equal to the current date (i.e., $T = t$), and apply a further differentiation, leading to the following SDE describing the evolution for the spot price, where $F(t,t) = S(t)$ defines the spot price:

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

(2)

The term in square brackets, which defines the drift of the spot process, involves the integration over the Brownian motions, and hence the spot price process will, in general, be non-Markovian. That is it will depend on all the random shocks which have occurred since start of the evolution at time zero. As a side note, for many energies such as the natural gas and electricity markets, seasonality in the forward price volatilities is an important feature in the evolution of the forward curve. One way to deal with this seasonality is to estimate volatility functions for each "season", where the season can be defined by the user to segregate the data to represent for example 'summer / winter' or 'summer / autumn / winter / spring' . A more elegant approach is to extend equation (1) to incorporate seasonality in the volatility functions by representing the functions as the product of a time dependent spot volatility function and maturity dependent volatility functions. The general equation (1) therefore becomes:

$$\frac{dF(t,T)}{F(t,T)} = \sigma_S(t)\sum_{i=1}^{n}\sigma_i(T-t)dz_i(t) \tag{3}$$

where $\sigma_S(t)$ denotes the spot price volatility at time *t* and $\sigma_i(T-t)$ the *n* maturity dependent volatility functions. In this way, the maturity structure of the volatility functions is normalized by the spot volatility and the volatility functions then capture the correlation between forward prices at different maturities independently of any seasonal effects. For clarity in this article we have chosen not to model the seasonality in this way but it is a straightforward extension of the analysis that we present.

Perhaps the main advantage of this forward curve modelling approach is the flexibility that the user has in choosing both the number and form of the volatility functions. The volatility functions can be determined in one of two general ways; historically, from time series analysis; or implied from the market prices of options. In this article we use the former method to illustrate estimation of the volatility functions.

Using historical forward curve data one method that can be used to simultaneously determine both the number and form of the volatility functions that drive the dynamics of the forward curve is principal components analysis (PCA) or eigenvector decomposition of the covariance matrix of the forward prices returns. The technique involves calculating the sample covariances between pairs of forward price returns in an historical time series to form a covariance matrix. The eigenvectors of the covariance matrix yield estimates of the factors driving the evolution of the forward curve.

To illustrate the process of estimating the volatility functions from historical data we consider a single commodity with n factors[1]. After applying Ito's lemma to equation (1) the forward curve dynamics is written as:

$$\Delta \ln F(t,t+\tau_j) = -\tfrac{1}{2}\sum_{i=1}^{n}\sigma_i(t,t+\tau_j)^2\Delta t + \sum_{i=1}^{n}\sigma_i(t,t+\tau_j)\Delta z_i \qquad (4)$$

Equation (4) implies that changes in the natural logarithms of the forward prices with relative maturities $\tau_j, j=1,...,m$ are jointly normally distributed. One can then compute an annualised sample covariance matrix of these forward prices (Σ) and decompose it into a series of eigenvectors and eigenvalues such that;

$$\Sigma = \Gamma\Lambda\Gamma^T \qquad (5)$$

Where

$$\Gamma = \begin{vmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ v_{21} & v_{22} & \dots & v_{2n} \\ \dots & \dots & \dots & \dots \\ v_{n1} & v_{n2} & \dots & v_{nn} \end{vmatrix} \text{ and } \Lambda = \begin{vmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_2 & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \lambda_n \end{vmatrix} \qquad (6)$$

The columns of ☐ are the eigenvectors. The eigenvalues represent the variances of the independent "factors" which drive the forward points in proportions determined by the eigenvectors. The discrete volatility functions are then obtained as

$$\sigma_i(t,t+\tau_j) = v_{ji}\sqrt{\lambda_i} \qquad (7)$$

As an illustration of the outputs from this analysis in Figure 2 we show plots of the seasonal covariance matrices for HNG forward curve data. For convenience we have defined the seasons to cover January – March, April – June, July – September, and October – December, but the user can categorize the seasons by their own definitions. Note that in order to obtain robust

[1] See Clewlow and Strickland [2000] for a detailed example of estimating volatility functions in this way.

estimates we have used a longer history of forward curves (from Jan 05 to Dec 07) and we have used the first 50 maturities.

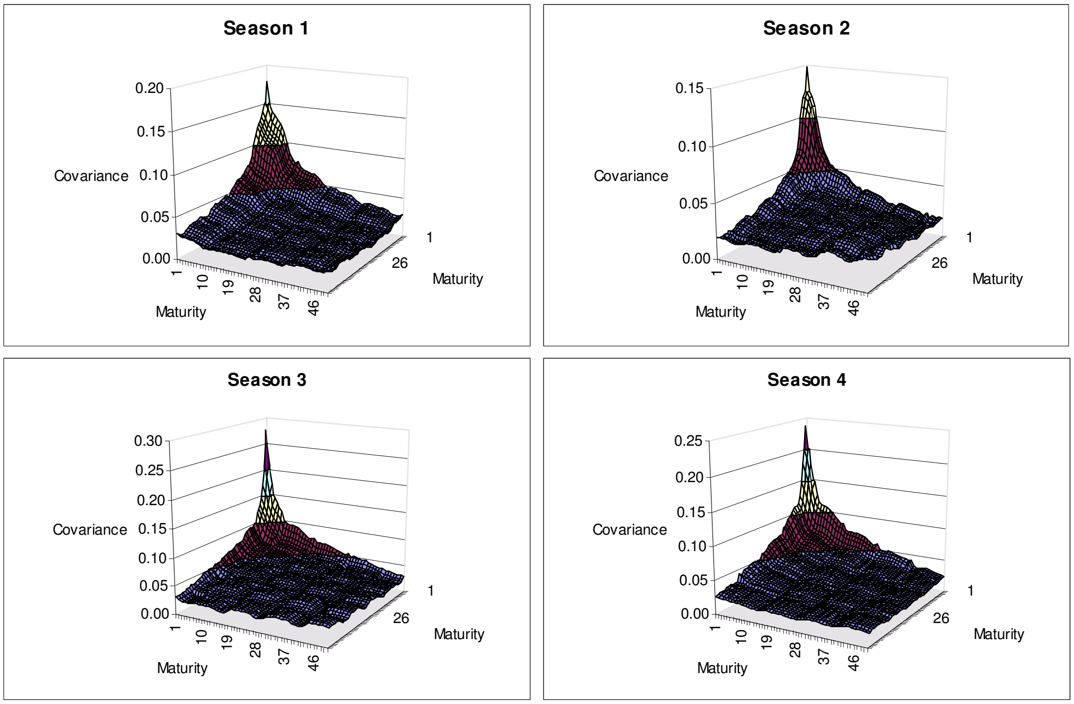

Figure 2: Covariance matrices for the Henry Hub natural gas forward curves for each season

The covariance surfaces illustrated in Figure 2 are typical of those found in energy markets. As you would expect the largest covariance is observed between price movements at the short end of the forward curve. The surface will then typically decay to a lower value for the longer dated contracts. The smoothness of the surfaces will depend on the amount of "noise" in the market, which may be due to illiquidity in contracts beyond a certain maturity, or may be due to changes in the market dynamics.

Once the covariance matrices have been calculated the form of the volatility functions in equation (1) can be obtained as described above. In Figure 3 we plot the first three volatility functions for season 1 (the results for the other seasons are very similar).

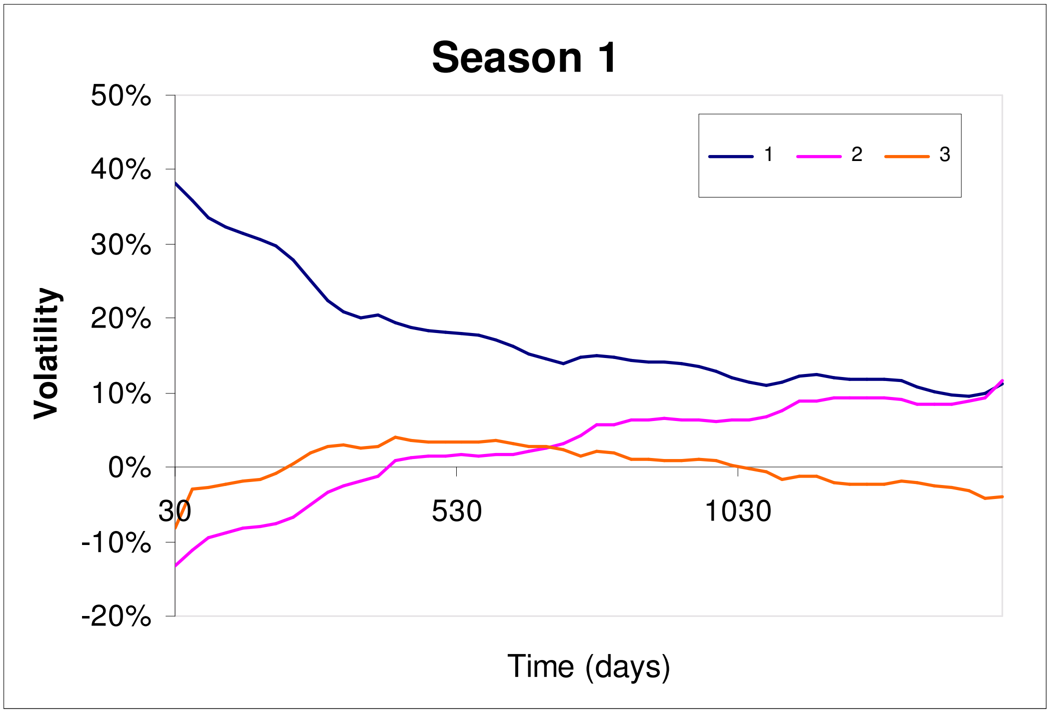

Figure 3: Plot of the first three volatility functions for the HNG data for season 1

The results shown in Figure 3 are typical of those found when analyzing energy market data. The first factor is positive for all maturities, indicating that a shock to the system will result in prices at all maturities to "shift" in the same direction. This is generally the most significant factor and is similar to the effect that would be seen in a single factor model. The second factor is a "tilt" which causes the short and long maturity contracts to move in opposite directions. The third factor is a "bending" factor, where the short and long ends of the curve move in the opposite direction to the middle of the curve. The second and third factors (and potentially others) are what distinguishes this approach from a single factor model, and allows the realistic dynamics of the forward curve to be captured in the model.

As we are dealing with 50 contracts in this example there are 50 factors that can explain the variance of the evolution of the curve, however only a few of these will be significant for explaining the variation in the forward curve. The eigenvalues obtained in the previous step (see equation (6)) indicate the importance of the corresponding eigenvectors (volatility functions). In practice we find that 2 or 3 factors are usually sufficient to explain the evolution of the observed market data. For this example the eigenvalues for the 50 factors are shown in Figure 4.

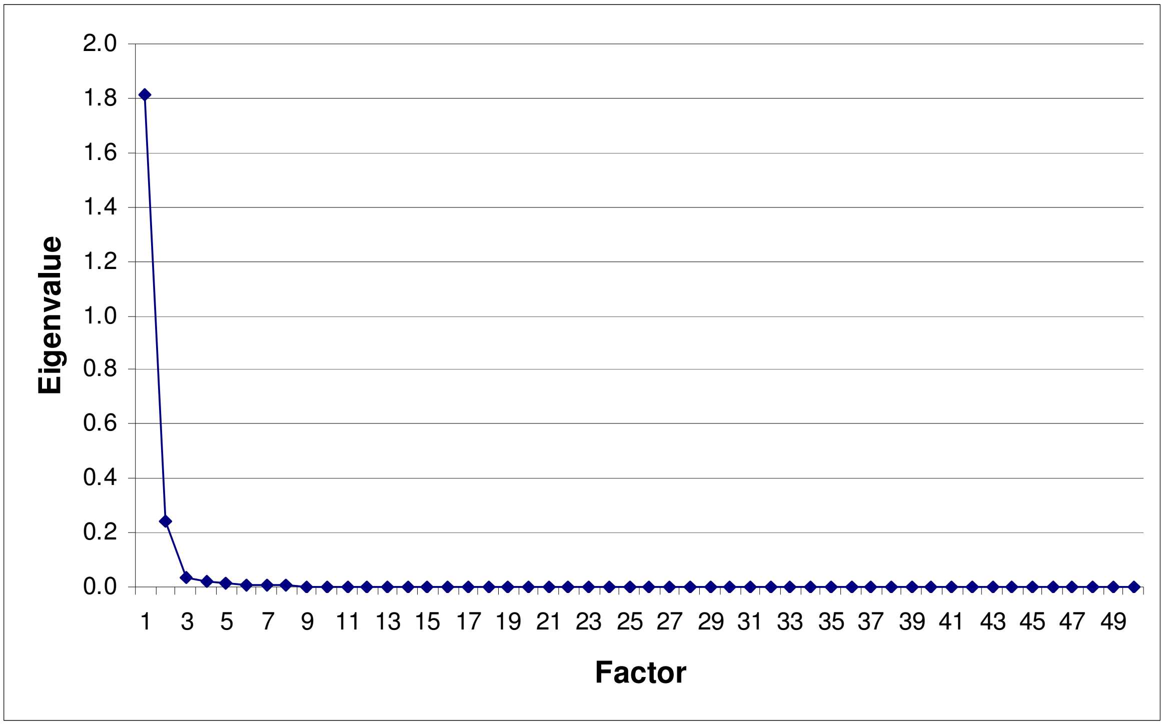

Figure 4: Eigenvalues for the 50 volatility factors derived from the HNG forward curve data

Clearly the first few factors have the largest contribution towards explaining the dynamics of the forward price. In fact for this example the first three factors explain 99.98% of the movement in the prices, so for modelling purposes it is easy to justify using only three factors to model the evolution of the forward price.

Many of the problems faced by practitioners in the energy risk management arena require the joint modelling of multiple commodities. Spreads between one or more fuels and a power price for example are key determinants in the valuation of power plants and many derivative contracts. Efficient calculation of at-risk measures, such as value-at-risk, or cashflow-at-risk type measures can be achieved by the simultaneous joint evolution of all risk factors that underlie the contracts or assets that form the portfolio. The multi factor model in (1) can be generalised further to describe the joint forward curve dynamics of multiple commodities as

$$\frac{dF_c(t,T)}{F_c(t,T)} = \sum_{i=1}^{n_c} \sigma_{c,i}(t,T)dz_{c,i}(t) \qquad (8)$$

Where:

$c = 1,\dots,m$ represents each different from the historical data and

$i = 1,\dots,n_c$ indexes the volatility functions for each commodity.

In this model the correlations between commodities are defined by a correlation matrix for the Brownian motions. The correlations between the Brownian motions driving a particular commodity are zero, while the correlations between the Brownian motions driving different commodities represent the inter-commodity correlations.

As with the case of estimating volatility functions, we can estimate the inter-commodity correlations from the joint historical forward curve data. For each commodity equation (4) describes the discrete time evolution of the forward curve in terms of the estimated volatility functions. This can be expressed in the following matrix representation:

$$\bar{x}(t) = \bar{\mu}(t) + \tilde{\Sigma}.\bar{\varepsilon} \qquad (9)$$

Where

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

$\bar{\mu}(t)$ is the vector of drift terms over the time step;

$\tilde{\Sigma}$ is the matrix of discrete volatility function terms;

$\bar{\varepsilon}$ is the unknown vector of standard normally distributed random shocks

Equation (9) can be solved to give estimates of the historical Brownian shocks which generated the evolution of the forward prices. We repeat this process for each commodity to obtain a time series of vectors $\bar{\varepsilon}$ for each commodity. The sample correlation matrix of the random shocks can then be calculated to give the inter commodity correlations for use in a multi commodity simulation.

As an example we consider the correlation between the HNG forward prices used above, and NBP natural gas forward prices covering the same period. The resulting covariance matrix is shown in Table 1, with a plot of the data in Figure 5.

	N1	N2	N3	N4	N5	N6	H1	H2	H3	H4	H5	H6
N1	1	0	0	0	0	0	0.13	0.08	-0.01	0.03	0.06	-0.05
N2	0	1	0	0	0	0	0.07	0.04	0.06	0.02	-0.04	0.10
N3	0	0	1	0	0	0	-0.02	-0.02	-0.08	0.05	0.01	-0.04
N4	0	0	0	1	0	0	0.02	0.02	-0.05	-0.04	0.00	-0.17
N5	0	0	0	0	1	0	-0.06	0.04	-0.03	0.04	0.01	-0.06
N6	0	0	0	0	0	1	-0.02	0.01	0.04	0.08	-0.02	-0.02
H1	0.13	0.07	-0.02	0.02	-0.06	-0.02	1	0	0	0	0	0
H2	0.08	0.04	-0.02	0.02	0.04	0.01	0	1	0	0	0	0
H3	-0.01	0.06	-0.08	-0.05	-0.03	0.04	0	0	1	0	0	0
H4	0.03	0.02	0.05	-0.04	0.04	0.08	0	0	0	1	0	0
H5	0.06	-0.04	0.01	0.00	0.01	-0.02	0	0	0	0	1	0
H6	-0.05	0.10	-0.04	-0.17	-0.06	-0.02	0	0	0	0	0	1

Table 1: Covariance matrix for HNG and NBP Brownian motions

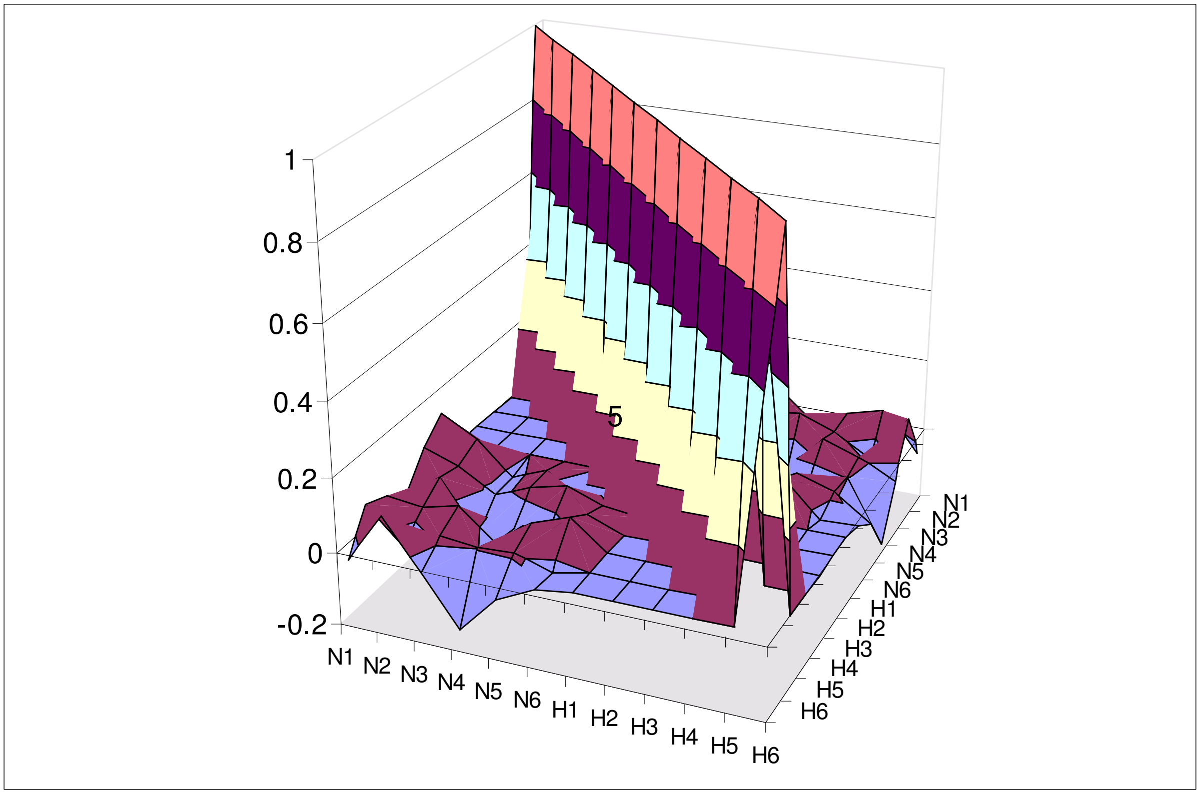

Figure 5: Covariance matrix of the Brownian motions of HNG and NBP forward prices

or this example we have considered six volatility factors for each commodity. The headers in Table 1 and the horizontal axes label the volatility factors: the factors for NBP are denoted by N1, N2, …, N6, and for HNG they are denoted by H1, H2, …, H6. As noted above, the covariance between the Brownian motions of each commodity is zero while the non-zero covariances in the off-diagonal blocks represent the correlations between the Brownian motions driving the different commodities. As noted above the correlations between the Brownian motions are used as input to a simulation model based on equation (8) to generate appropriately correlated normally distributed random shocks.

For single commodity applications the model described in equation (1) has a number of desirable analytical properties – see for example Clewlow and Strickland [2000] where we detail the analytical pricing of European options on both the spot asset and futures contracts. For most applications, however, Monte Carlo simulation is the numerical technique that the majority of practitioners turn to. One important real world application for a MFMC model is in determining the optimal location for shipping a cargo, and in the next article we will illustrate the use of the MFMC model for this type of application.

References

Clewlow, L, and C. Strickland, 2000, "*Energy Derivatives: Pricing and Risk Management*", Lacima Publications.

Clewlow, L, Strickland, C, and Kaminski V, 2001, "*The Many Shapes of Volatility*", EPRM, Risk Publications, April.

About Lacima Group

Lacima Group is a specialist provider of energy and commodity pricing, valuation and risk management software and advisory services. Based on its internationally acclaimed research in energy risk modelling, Lacima offers integrated risk management applications to address valuation, market and credit risk or the flexibility of stand-alone solutions for swing, storage and generation assets. These solutions help energy producers, retailers, distributors, large consumers and financial institutions to value and manage risk associated with complex derivative contracts and physical assets across multiple commodities and regions in a cost-effective manner.

www.lacimagroup.com