The Masterclass series continues with a discussion of a general multi-factor, multi-commodity model and the process for estimating parameters from historical data. By *John Breslin, Les Clewlow, Calvin Kwok* and *Chris Strickland*

Gaining from complexity: MFMC models

★ In our previous two articles, we discussed the modelling of swing contracts based on a single-factor model for describing the evolution of gas prices. Single-factor models have a wide range of applicability in energy valuation and risk management, and are relatively simple to understand and parameterise. 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 such 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 here. It is also relatively straightforward to extend such a multi-factor model to incorporate multiple commodities. This article discusses a general multi-factor, multi-commodity (MFMC) model and describes the process of estimating parameters from historical data.[1] In our next Masterclass article we will use this underlying model in a practical application that involves the simulation of multiple energy forward curves.

Figure 1 illustrates the historical evolution of a forward curve for Henry Hub natural gas (HNG) from January 2, 2007 to December 14, 2007. 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; that is, on each calendar date we plot the nearby contract, the second nearby and so on, out to the 24th nearby contract. The first few curves in January 2008 therefore contain forward prices for contracts maturing each month from February 2007 to January 2009, while the last curves in December 2007 contain prices for contracts maturing from January 2007 to December 2009.

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);

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

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 that 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 the 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 known spot price models. We can integrate equation (1), set the maturity date equal to the current date (that is, $T = t$), 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:

1. *This article can be seen as an extension of Clewlow, Strickland & Kaminski (2001)*

$$\frac{dS(t)}{S(t)} = \left[\frac{\partial \ln F(0,t)}{\partial t} - \sum_{i=1}^{n} \left\{ \begin{array}{l} \int_0^t \sigma_i(u,t)\frac{\partial \sigma_i(u,t)}{\partial t}du + \\ \int_0^t \frac{\partial \sigma_i(u,t)}{\partial t} dz_i(u) \end{array} \right\}\right]dt + \sum_{i=1}^{n} \sigma_i(t,t)\,dz_i(t) \qquad (2)$$

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 that have occurred since the 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:

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

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 normalised 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 a 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[2]. After applying Ito's lemma to equation (1) the forward curve dynamics is written as:

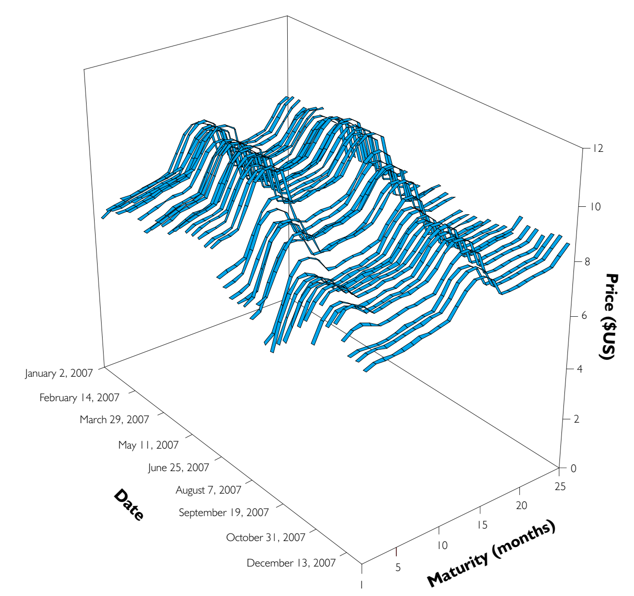
$$\Delta \ln F(t, t+\tau_j) = -\frac{1}{2}\sum_{i=1}^{n}\sigma_i(t, t+\tau_j)^2 \Delta t + \sum_{i=1}^{n}\sigma_i(t, t+\tau_j)\Delta z_i \qquad (4)$$

Equation (4) implies that changes in the natural logarithms of the forward prices with relative maturities τ_j, j=1,...,m are jointly normally distributed. An annualised sample covariance matrix of these forward prices (Σ) can be computed and decomposed into a series of eigenvectors and eigenvalues such that;

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

where

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



F1. Forward curve for Henry Hub natural gas during 2007

Source: Lacima Group

2. *See Clewlow and Strickland (2000) for a detailed example of how to estimate volatility functions in this way.*

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

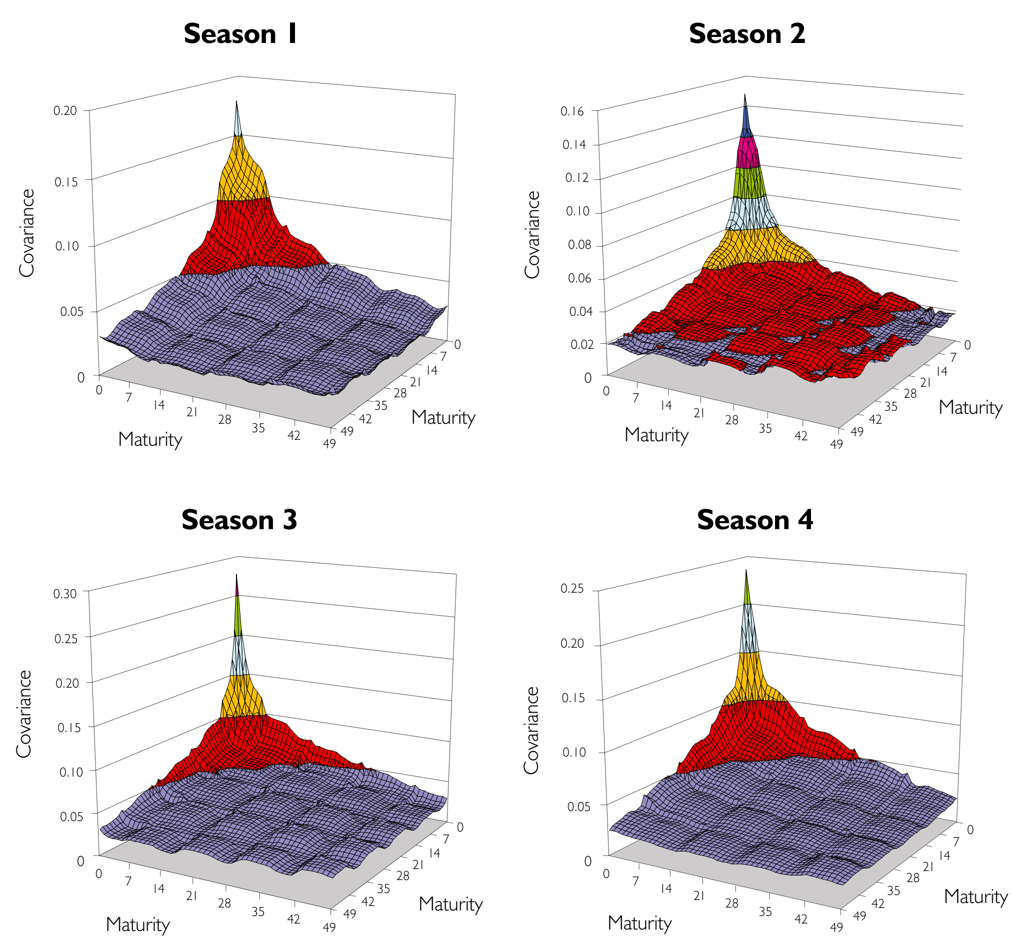
$$\sigma_i(t, t+\tau_j) = v_{ji}\sqrt{\lambda_i} \quad (7)$$

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 to March, April to June, July to September, and October to December, but the user can categorise the seasons by their own definitions. Note that in order to obtain robust estimates we have used a longer history of forward curves (from January 2005 to December 2007) and we have used the first 50 maturities.

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 changes in the market dynamics.

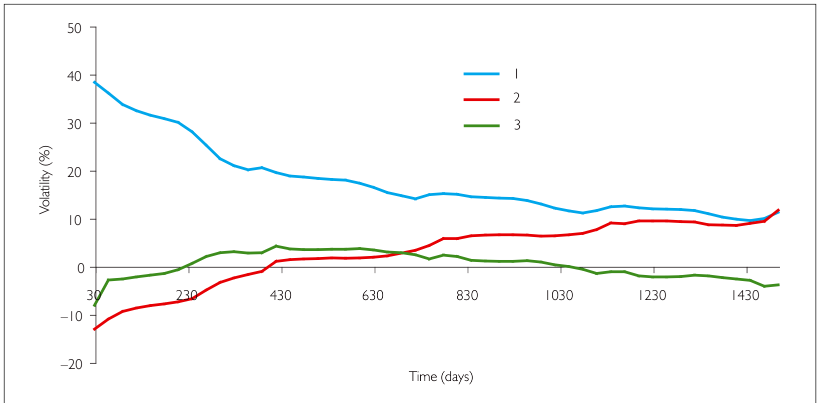
Once the covariance matrices have been calculated, the form of the volatility functions in equation (1) can be obtained as described above. In figure 3 we plot the first three volatility functions for season 1 (the results for the other seasons are very similar).

The results shown in figure 3 are typical of those found when analysing 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' that causes the short and long maturity



F2. Covariance matrices for the Henry Hub natural gas forward curves for each season *Source: Lacima Group* 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 two or three 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.



F3. Plot of the first three volatility functions for the HNG data for season 1

Source: Lacima Group

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

$c = 1,\ldots,m$ represents each different commodity from the historical data and $i = 1,\ldots,n_c$ 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 timestep;

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

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

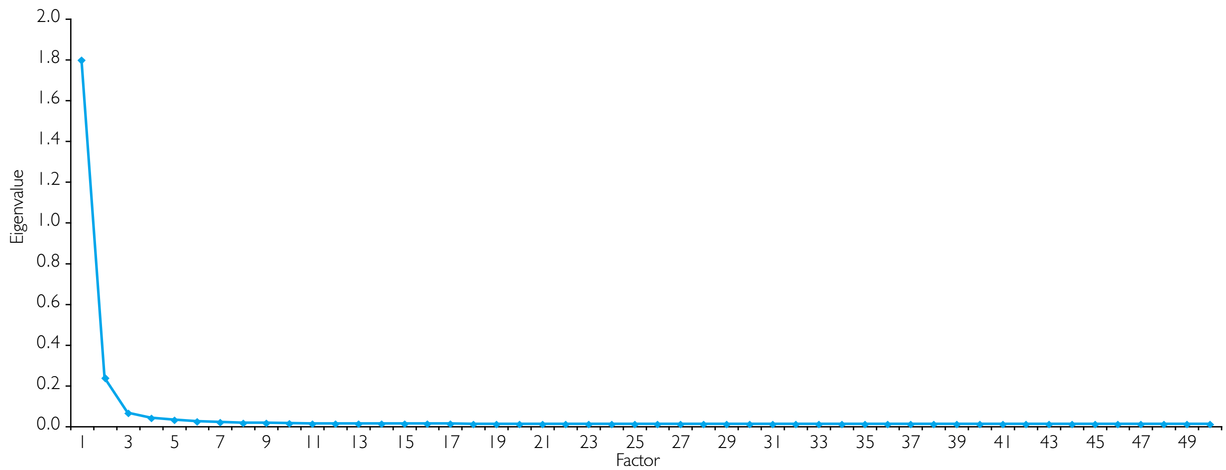
$\bar{\varepsilon}$ is the unknown vector of standard normally distributed random shocks.

Equation (9) can be solved to give estimates of the historical Brownian shocks that 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.

For 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 $N1, N2, \ldots, N6$, and for HNG they are denoted by $H1, H2, \ldots, H6$. As noted above, the covariance between the Brownian motions of each commodity are zero, while the non-zero covariances in the off-diagonal blocks represent the correlations between the Brownian motions driving the different commodities. Also, 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



F4. Eigenvalues for the 50 volatility factors derived from the HNG forward curve data *Source: Lacima Group*

in equation (1) has a number of desirable analytical properties (see Clewlow & 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 of this Masterclass series we will illustrate the use of the MFMC model for this type of application. ER

Les Clewlow and Chris Strickland are the founders and directors of Lacima Group, where John Breslin is a principal and Calvin Kwok a quantitative analyst. Email: info@lacimagroup.com

References

Clewlow L and C Strickland, 2000
Energy derivatives: pricing and risk management
Lacima Publications

Clewlow L, C Strickland and Kaminski V, 2001
The many shapes of volatility
EPRM, Risk Publications, April 2001

T1. Covariance matrix for HNG and NBP Brownian motions

	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