Multi-Factor Models for Forward Curve Analysis:

An Introduction to Principal Component Analysis

A model of the dynamics of the evolution of the forward curve is an essential building block of any energy trading and risk management methodology. Risk managers face the challenge of modelling forward curves for earnings, profits, cash flow or value at risk calculations. This is particularly important over long time horizons when seasonality and forward curve dynamics are crucial ingredients of these calculations. By **CARLOS BLANCO**, **DAVID SORONOW** & **PAUL STEFISZYN**, Financial Engineering Associates

CERTAIN INSTRUMENTS OR contracts in commodity markets have a value that depends on more than one forward price. Common examples are OTC derivatives such as options on swaps or 'Swaptions'. Other contracts that depend on multiple forward prices are certain average price and average strike options, swing and take-or-pay contracts, and gas storage contracts.

To value these contracts, it is necessary to develop a realistic model of the simultaneous evolution in time of all the forward prices; that is, a model of the forward price curve. Through the use of these models, we seek to accurately capture the joint behaviour of forward prices, and to know how they move together, rather than how they move in isolation. Using a pricing or risk management model that does not appropriately capture the evolution through time of the full forward curve often results in severe mispricing of certain instruments or large errors in risk estimations.

In this article we describe a seasonal multi-factor model for the evolution of the forward price curve in energy markets. This new approach captures the evolution of forward prices in a more realistic fashion than either the Black-Scholes type, or the mean reverting one-factor models used by most practitioners. Additionally, it is more accurate than other *Principal Component Analysis* (PCA) techniques that use the full sample data without taking into account the seasonal nature of that data.

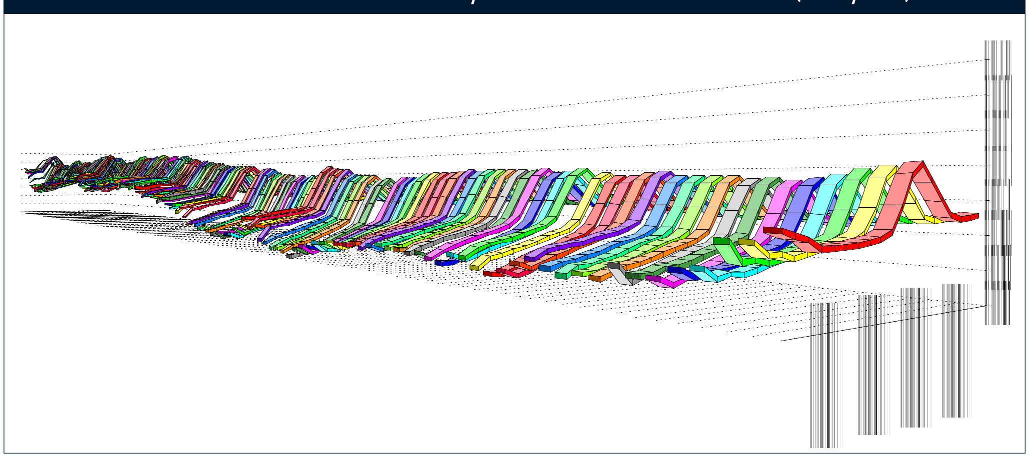
Principal Component Analysis & Forward Curves

Forward curves are just a collection of forward prices for different maturities. The forward price curve can take a whole variety of shapes depending on the market and the time period analysed. Some energy forward price curves are characterised by exhibiting strong seasonality, like the electricity or natural gas forward price curves, as we can see in the figure below with clear seasonal peaks and lows. In the case of electricity, the summer is usually the season with higher expected prices, while for natural gas, the winter months are the ones with higher expected prices as we can clearly see in chart 1 below.

Some methodologies used to capture the evolution of the forward curve require information about the volatilities and correlations of each of the different points in the forward curve. A significant problem with this type of approach is that, as the number of forward prices in the curve increases, so does the complexity of the analysis. For example, NYMEX natural gas futures contracts trade for 72 consecutive months. If we have a forward curve with 72 maturities then we need to model 72 correlated variables. We can do the analysis by estimating volatilities for each variable and a correlation between every pair of variables. With so many parameters to estimate, however, the calculation is difficult to manage and the result is liable to be non-intuitive.

We can estimate the value of many energy contracts through a Monte Carlo simulation of the evolution of the forward price curve. However, we must be careful to model the entire forward curve for natural gas in such a way that realistically accounts for changing spreads.

Historical Forward Curves for Nearby Months for NYMEX Natural Gas (Henry Hub)



How can we do this? One possible approach to this simulation problem is to directly apply the variance-covariance (V-C) matrix of the forward prices, but this is cumbersome and inefficient.

For example, natural gas prices for delivery at different future times are highly correlated but not perfectly correlated, indicating that there are fewer sources of randomness (we can think of them as 'drivers' or 'factors') than there are prices. In such a situation, a more elegant solution to the simulation problem is to apply PCA.

Using PCA, we can reduce the number of dimensions of the problem considerably. We can end up with only two or three significant uncorrelated sources of risk, while retaining most of the information contained in the original data set. In our NYMEX example, the PCA will still yield 72 dimensions (called components), but generally only a very small number of them (the principal ones) will have a significant contribution to the variability of the data.

Another advantage of the PCA process is that it yields information on the importance of each component: the first principal component contributes more to the variability of the data than the second, and so on. We can also be confident that the components are independent, whereas our original data was correlated.

PCA is widely used in many fields, including finance, to perform analysis on complex sets of data. However, it is important to understand that this type of analysis does not result in discovering the underlying economic drivers of data changes. For example, using PCA on electricity prices would not 'discover' weather and coal prices as important factors in electricity price evolution. Rather PCA uses abstract mathematical techniques to reveal the underlying structure of the data.

PCA & Energy Forward Price Curves

In the context of energy forward curves, PCA enables us to gain a much richer and more intuitive understanding of the way in which the forward curve changes. By observing historical forward curve movements, PCA can identify a set of principal components that describe those movements mathematically and rank their importance according to their contribution to the total variability of the data.

Each of the principal components can be interpreted as a source of risk, and the importance of that component (its 'score') is an expression of the volatility of that risk source. Associated with each principal component is a set of 'factor loadings', which can be interpreted as the volatilities for each maturity forward in the original dataset corresponding to the source of risk. For example, the factor loadings for the first principal component define how the price of each of the original contracts moves in response to a change in that component.

For energy forward price curves (and in financial markets as well), these uncorrelated sources of risk are highly abstract and usually take the form of:

- 1st Factor: Parallel Shift – this factor governs changes in the overall level of prices. The factor loadings usually have the same sign but slightly different magnitudes.

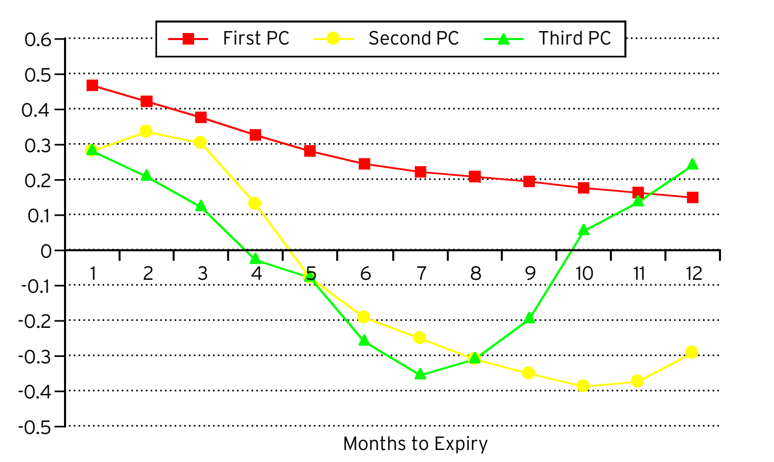
- 2nd Factor: Slope – this factor governs the steepness of the curve: it may result in the front of the curve going up and the back going down, or vice versa. It can be interpreted as a change in the overall level of the term structure of convenience yields.

- 3rd Factor: Curvature – this relates to the possibility of introducing a bend in the curve, that is the front and back go up and the middle goes down, or vice-versa.

Chart 2. above illustrates the three different components and their effect on the forward curve.

Some authors have suggested the use of PCA to model

Simulating Forward Curve Changes with Seasonal Principal Component Analysis



energy forward curves. Analysis of this type that uses the full set of sample data will produce forward curve scenarios that will capture the average variability throughout the year. However, without taking into account the seasonal dependence of those factors, the resulting curve may not be realistic. For example, the results may characterise price behaviour in a 'typical month', but that month will never exist in reality because it is a composite average of 12 real months, none of which exhibit the 'typical' behaviour.

> PCA is widely used in many fields, including finance, to perform analysis on complex sets of data <

In order to simulate movements in the forward curve through time, we need a dynamic framework that captures the most characteristic patterns that drive forward curve changes. FEA's approach allows the data to be viewed in seasonal slices so that we only compare like time periods with like.

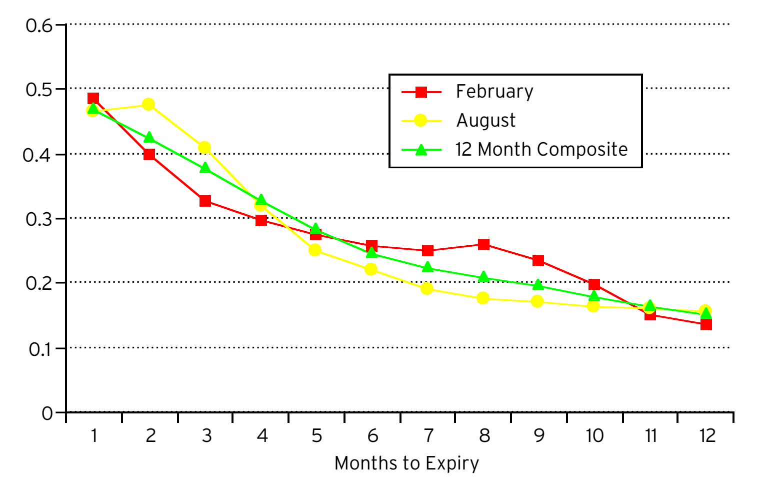
Why Should We Use Seasonal PCA?

At first glance, the method described above for simulating the evolution of the natural gas forward curve seems reasonable. But how realistically does it capture the changes through time of forward prices for natural gas? The answer is – not very well.

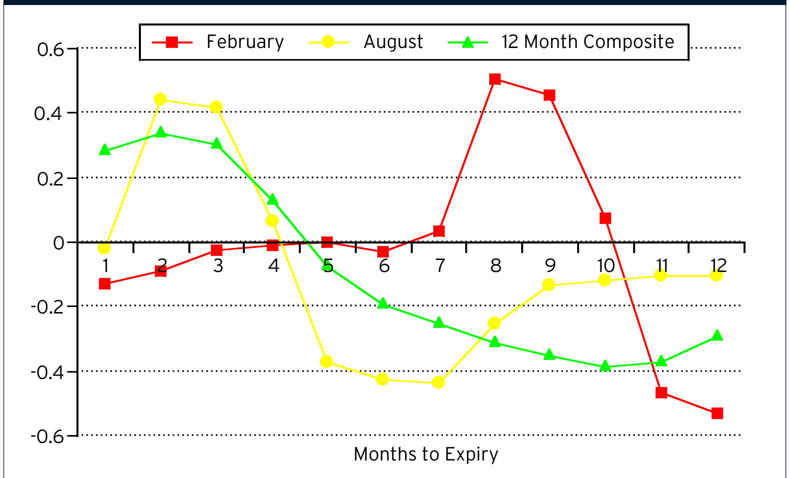
As we said, natural gas prices are seasonal, and the behaviour and evolution of the forward prices are seasonal as well. By considering the full set of historical data when applying the PCA, we produce results which capture the average variability of the forward price curve throughout the year. Using these results, we can then model price behaviour on a 'typical' day. The problem is, the typical day never really exists because the behaviour of forward prices in the summer is different from the behaviour in the winter.

Our solution is to break up the entire set of historical data into monthly groups (i.e., all the forward price observations made in January, all observations made in February, etc.). We then perform 12 separate principal component analyses, to determine the behaviour of forward prices in each of the 12 months of the year. During the simulation of the forward

First Principal Component



Second Principal Component



price curve evolution, we model each month turn by turn, applying the principal components which are specific to the month in question.

Let's examine this in more detail. Suppose we have extracted the principal components for each of the 12 'seasonal' buckets, and also for the entire data set.

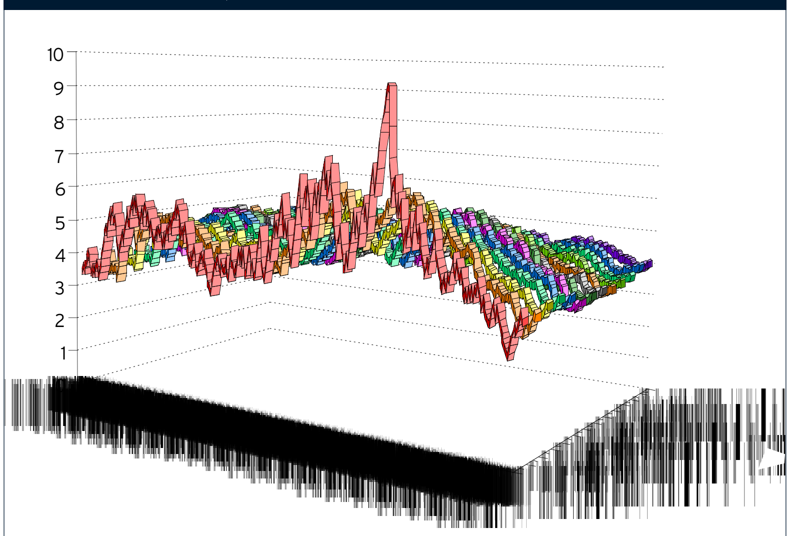
Comparing the first principal components, we see that the 'composite' first principal component derived from the entire data set would be a fair approximation to any of the seasonal first principal components (February and August are given as examples).

However, comparing the second principal components, we see that the second principal component derived from the entire data set is entirely different from the seasonal second principal components, which are themselves very different from each other. In particular, we see that a shock to the second principal components will always load onto those contracts which expire just prior to winter delivery, regardless of whether they are short-dated or long-dated contracts. In other words, the forward price volatility associated with the second principal component depends not on whether the contract is maturing soon, but whether the contract is for winter delivery.

Furthermore, for a typical sample data set which we analysed, the factor score for the composite second principal component was only 6% of the total variance, while the factor scores for those derived from the seasonal analysis ranged from 8% in the winter months to 19% in the summer months. Using the composite analysis would therefore underestimate the contribution of the second principal component to the variability of the data.

Recall that the second principal component models primarily the changes in calendar spreads between futures prices. These changes in spreads are a significant source of value for different instruments traded in the gas markets, and any modelling strategy that fails to realistically model the seasonal differences between the second principal components may seriously misprice those contracts.

Joint Simulation of Spot & Forward Prices for Natural Gas (Simulated Prices)



Simulation with Principal Component Analysis

Once we have performed the PCA, we can use the results to simulate hypothetical forward curve changes that can be used for pricing and risk measurement. A simulation model using PCA is more efficient than other multi-factor approaches, and is more realistic than single-factor models. It reduces the dimensionality of the data while maintaining most of the variance and possibilities for joint changes in the forward prices.

PCA explicitly reveals the importance of each principal component (its 'factor score'), which is an expression of the contribution of that source of risk to the volatility of the forward price. Also associated with each principal component is a set of 'factor loadings' which define how the price of each forward contract will change in response to a shock to the component ■

In our next article (September edition) we will present a new model to simulate Spot and Forward Prices simultaneously using PCA for the forward curve and different processes for the spot price such as mean reversion or jump diffusion.

CARLOS BLANCO directs the Global Support and Professional Services Group at FEA. **DAVID SORONOW** & **PAUL STEFISZYN** are Derivatives Specialists in the same group. Blanco, Soronow and Stefiszyn thank Cheryl Morgan, Dr. King Wang, and Dr. Angelo Barbieri for advice and assistance.

www.fea.com