Annual Meeting
March 18-20, 2007
Marriott Rivercenter Hotel
San Antonio, TX

AM-07-45 Crude Oil Price Forecasting: A Statistical Approach

Presented By:

Armando Lara
Senior Consultant
Turner, Mason and Company
Dallas, TX

Michael W. Leger
President
Turner, Mason and Company
Dallas, TX

John Auers
Vice President
Turner, Mason and Company
Dallas, TX

National Petrochemical & Refiners Association

1899 L Street, NW
Suite 1000
Washington, DC
20036.3896

202.457.0480 voice
202.457.0486 fax
www.npra.org

Crude Oil Price Forecasting: A Statistical Approach

Abstract

In a time when prices of crude oil and refined products have become ultra sensitive to industry developments, geopolitical events, natural disasters, and perceived risks, it is more important than ever to develop supply, demand, and price models that provide a realistic outlook of the pricing environment by considering many complex variables. Project evaluations, strategic planning, and financial transactions, among others, employ such forecasts, and their outcome is directly linked to the quality and reasonableness of the pricing parameters used.

This highly volatile pricing environment gives industry experts a new challenge as traditional methods of forecasting prices, based on simple regressions of historical data, do not do as good a job as they once did. On the other hand, the purist economists of our times are often unable to explain variations in pricing trends as a result of a lack of understanding of the inner workings of the petroleum industry.

This presentation examines an alternative for forecasting prices of crude oil employing stochastic models, decades of experience in the industry, and information about future capacity additions or deductions. The presentation will discuss the fundamental structure of a stochastic model, the intricacies of the oil industry and the basic supply, demand, and pricing relationship.

Introduction

Much has been said about the run-up of gasoline and crude oil prices in the last few years. In the past three years, the oil and refining industries experienced unprecedented events, both from the standpoint of the magnitude of their impact and the frequency with which they occurred. Geopolitical tension, hurricanes, new regulations, and other events have shaken the delicate supply/demand balance of our industry to an extent no one expected. In fact, having a realistic expectation of what the future would bring has become a lot more important than before.

As the industry seeks to meet the increasing demand with constraints on increased crude supply and refining capacity, we have come to the conclusion that the one thing that we can expect with any kind of certainty is continued price volatility.

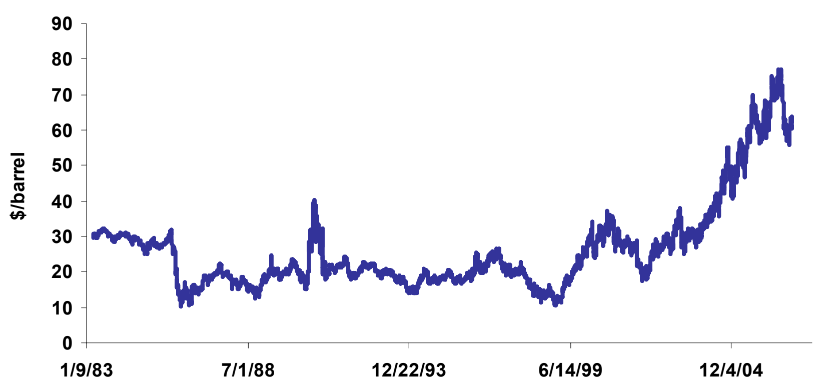
It is not new for mathematics to provide solutions and tools for engineers and scientists. Historically, in times of lower volatility in pricing, a simple linear regression could have served as an adequate representation of what crude oil pricing would be for months to come. The challenges of late have given us the opportunity to employ the tools from statistics commonly used in the study of biological, financial, and other systems. Stochastic models provide the means of studying systems with inherent volatility or randomness. This paper presents an approach that can be used to forecast the price of crude oil and refined products by employing stochastic methods as one of its elements. As a prologue to the presentation of the forecasting approach, I describe the characteristics of the industry that have displayed volatility and attempt to identify a couple of sources of that volatility.

The method presented in this work is not intended to provide a “crystal ball” picture of the future. Instead, I will demonstrate a convenient use of well-known statistical tools which, when used in combination with professional judgment and industry facts, can provide a useful pricing reference.

Supply, Demand, and Pricing in the Oil Industry

It has been noted already that the American people are addicted to oil. This addiction, however, has not been limited to the U.S. Countries around the world (China and India most notably) have also shown evidence of falling prey to the same addiction. Interestingly, the network that sustains this addiction gets more complex every year. Consolidation of majors, government regulations, technology developments, geological factors, and even individuals in positions of power or influence impact the delicate supply and demand balance of the industry. Pricing, a job that used to demand few resources, has become a complex ordeal with companies, research institutions, and government agencies dedicating substantial time and money to “getting the right pricing projection.”

Figure 1
WTI Cushing Price
1983-2006

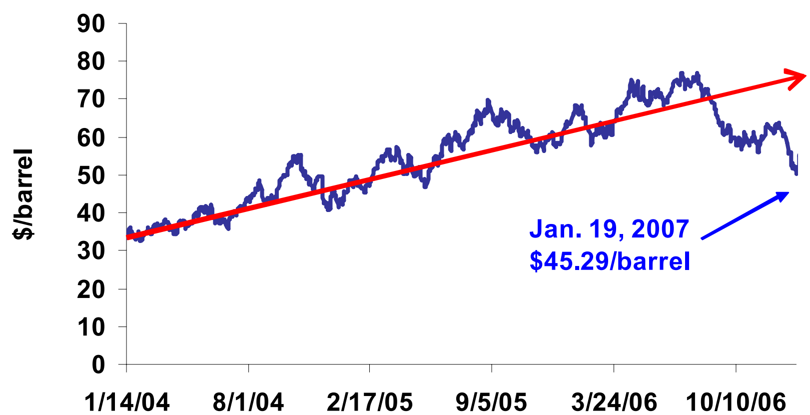


Our industry has often demanded “90 percent certainty on the long-term projection” in regard to a parameter whose value next week could be difficult to estimate with what we know today. As shown in Figure 1, the jump in price since 2004 for WTI crude, a key benchmark in the industry, is unprecedented. For most of the period between mid-1986 and mid-1999, this crude stayed pretty consistently within a band, around the $20 mark.

A sudden increase in 2001 took place as a consequence of the 9/11 events. Thereafter, until the end of 2003, the price of a barrel of oil stayed in the high $20's for a couple of years. After this, things became much more interesting, and forecasting the price of crude became a job by itself.

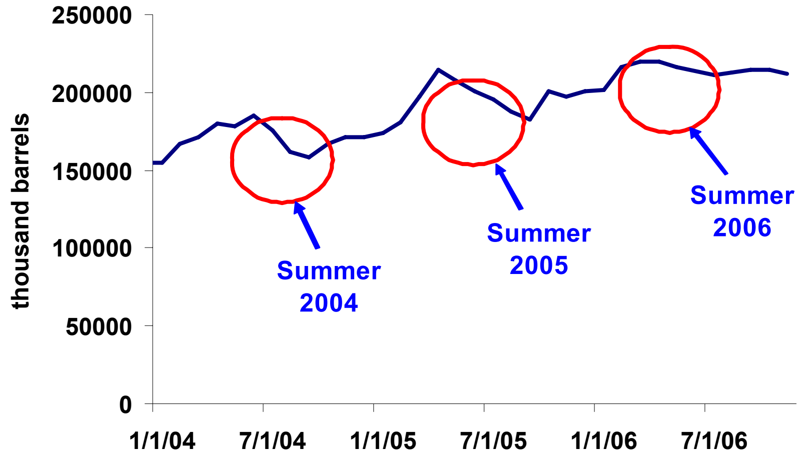
In fact, the environment is so uncertain these days that volunteering for presenting a work like this one is riskier than investing in the most volatile economies of the world. For example, in August 2006 when I was first thinking about this paper, the market had displayed a fairly steady rate of increasing price. As shown in Figure 2 for the U.S. average Spot Price, discussing the “complex” technique of linear extrapolation would have satisfied many readers as an appropriate projection for today’s environment of limited supply and what appears to be an ever-increasing demand of oil. Shortly after that, we witnessed a change in direction of a trend that had kept many in the industry up at night, trying to come up with new ways to justify capacity expansions. By January 19, that trend had taken us all the way down to $45.29 per barrel.

Figure 2
WTI Cushing Price
Jan. 2004-Jan. 2007



Looking at crude inventories in the U.S. (Figure 3), it seems that the oil market finally behaved like free markets do. If the commodity is selling at increasing prices, consumers respond by slowing down their rate of consumption.

Figure 3
U.S. Crude Oil Stocks



After years of steep declines during the consumption peaks of summer, Americans moderated their consumption, leading to inventories staying fairly flat for weeks. We still see that correction to the supply and demand dynamics in today's prices. However, this is an aspect of the market that injects volatility as well. As Americans look for more efficient alternatives to power their lives, the uncertainty in future consumption of gasoline and crude oil increases. Aspects of the economy that used to be foreign to the oil industry and refining market are now on the verge of impacting us. As discussed later in this paper, corn and the agricultural segment are examples of factors that used to be irrelevant to this discussion, but now they are no longer outside the ambit of the petroleum industry.

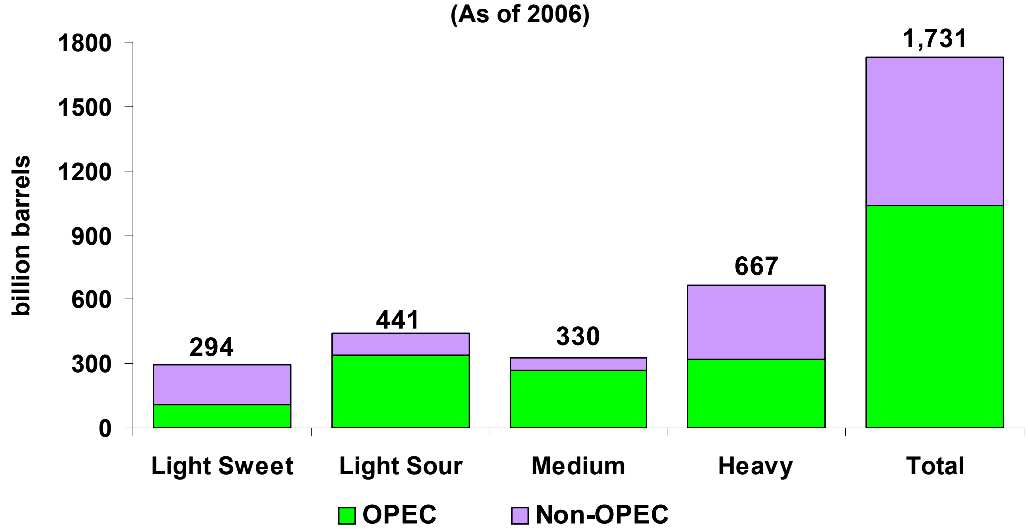
What will happen next? As the complexity of the industry grows, the variables impacting it increase in number as well as in the magnitude of their impact. As such, the only certain thing about the future is that the market will continue to display volatility in its march toward equilibrium.

Crude Supply Trends

A critical factor in determining refined product prices and refining profitability is the quantity and quality of crude oils available to refiners around the world. As a result, any uncertainty associated with the supply of this key raw material in the world economy has a major impact and leads to uncertainty in pricing.

Worldwide recoverable crude oil reserves (including Canadian and Venezuelan oil sands) amounted to 1,739 billion barrels as of 2006. Of these, 667 billion barrels, or 38%, were classified as heavy while 296 billion barrels, or 17%, were classified as light sweet. Figure 4 shows the breakdown of crude reserves by type.

Figure 4
Worldwide Recoverable Crude Reserves*
(As of 2006)



*** Including Venezuelan and Canadian oil tar sands.**

The crude oils being produced are not in proportion with the quality of crude oil reserves, however. In 2005, 13% of the crude oil produced was heavy while 38% of the oil extracted from fields around the world was light sweet. The world is obviously in much better position to process the lighter, sweeter crude. The imbalance between what can be processed and what is available for consumption in the long term adds uncertainty to the crude supply.

Another level of uncertainty for crude supply arises when one considers the source of crude oil. As shown in Figure 5, the most significant portions of recoverable crude are located in regions with major political and social turmoil. OPEC countries contain 60% of the recoverable crude, with the bulk of these in the Middle East. A significant portion of this crude can also be found in South America. OPEC features six of the top eight producing countries in the world, as displayed in Table 1. At 316 billion barrels of recoverable oil, Venezuela holds 18% of the commodity that drives the world's biggest economies.

Figure 5
Worldwide Recoverable Crude Reserves
(As of 2006)

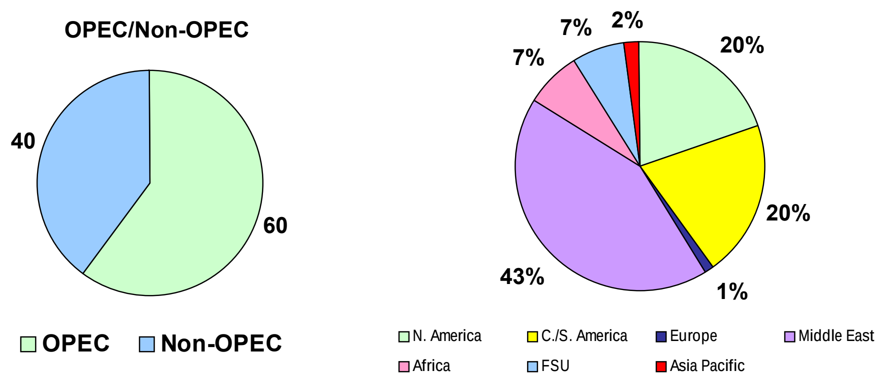


Table 1
Top Eight Recoverable Reserve Countries

	Billion Barrels	Composition
Venezuela[1]	316	95% heavy
Canada[1]	310	97% heavy
Saudi Arabia[2]	267	88% medium or light sour
Iran	133	100% medium or light sour
Iraq	115	93% medium or light sour
Kuwait[2]	104	100% medium or light sour
UAE	98	96% light sweet and light sour
Russia	75	71% light sour
	1,417	

(1) Including oil tar sands.
(2) Kuwait and Saudi Arabia split Neutral Zone Reserves 50/50.

Forecasting crude production and the quality of crudes available for consumption is a subject by itself. Nevertheless, it is an aspect that must be taken into consideration in thinking about the uncertainties one must face when forecasting oil prices. One thing can be said for sure: The crude oil production quality will shift significantly over the next 20 years with heavy crude oil increasing in production and the world experiencing the decline of light sweet crude. Apart from that, we will still face the issue of matching supply and demand.

Corn: Not An Externality Any More

For a long time I thought that corn and oil could only be part of the same discussion in meals with my chemical engineering friends. These two rarely seemed to cross paths. Today, the refiner encounters corn in more places than a bowl of chips.

Non-conventional transportation fuels, including biofuels, GTL fuels, and CTL fuels, are receiving considerable political, environmental, and even economic development push. Many see these non-petroleum based fuels as the salvation from a world crude oil supply dominated by unstable nations or the means to reduce global warming emissions from the transportation sector. Additionally, biofuels, in particular, introduce a new segment of the economy, agriculture, into the massive energy market.

Today, a whole new set of factors must be considered in the economic evaluation of infrastructure dedicated to the production of fuels. Technological risk in the agricultural segment, abrupt weather changes that may impact the production of crops, and the demand from the food industry are just examples of things that just be externalities to us in the refining world. Needless to say, the job becomes an impregnable task when one is forced to consider the government's behavior, on which the future of non-conventional fuels is so dependent.

Today, we are very limited on what we are able to say, with any degree of certainty, about the economic impact of these non-conventional fuels on the refining industry and on the consumption and pricing of oil in the long term. For the time being, these non-conventional products may exert some pressure on the price of conventional refined products if:

- Total conventional and non-conventional supply exceeds consumer demand; <u>and</u>

- Governments continue to support the price of the non-conventional products.

Predictions of the explosive growth of non-conventional transportation fuels have been around for 30 years, but that growth has yet to materialize. We may be at the threshold of such growth today, if three factors occur:

1. Oil prices stay at recent high levels;

2. Non-conventional production technology continues to improve, particularly ethanol production from cellulosic feedstocks; and

3. Governments continues to support non-conventional transportation fuel development.

If all three of these factors occur, we may observe robust growth for non-conventional fuels, for a number of years. If these factors do not materialize, we can expect growth of non-conventional fuels, but at rates no greater than overall transportation fuel demand.

<u>A Structured Approach to Forecasting</u>

For years, firms have employed internally generated estimates of prices for crude oil and refined products around the world. The task had not been as demanding until a few years or months ago, as discussed previously. With the accelerated increase in prices and their unprecedented volatility, the use of a credible forecast became more popular as people rushed to justify new projects. At the end, the objective should be to generate a solid reference that is simple to use and credible. This paper recommends a set of elements or steps to arrive at such information.

1. Develop a database of announced projects in refining and crude production.

2. Estimate the impact of added crude production and crude processing capacity on inventories using existing refining models.

3. Study projections on growth of non-conventional fuels capacity and their impact on inventories of conventional fuels.

4. Develop a rationale to estimate expected growth in consumption.

5. Reconcile the supply/demand balances.

6. Generate a stochastic model for pricing employing the historical data (assuming the historical supply/demand patterns will prevail).

7. Correct the price estimates based on the reconciled consumption patterns and employing professional judgment.

The objective of this approach is not to provide a detailed day-by-day forecast of prices as intended for daily trading and similar activities. Additionally, the approach outlined above does not intend to provide a “crystal ball” market call or the ability to predict social or political events and their impact on the oil industry. Rather, this model provides a long-term price projection with a strong mathematical basis and, more importantly, the professional judgment of industry experts employing industry facts.

Distilled to Simplicity: Statistics in Forecasting Oil Prices

What role does statistics play in forecasting oil prices? In order to achieve the objectives outlined above, the forecasting exercise requires a reference model as a basis for projection. Statistics provides this model, a model of historical performance. In a sense, it allows us to estimate what the behavior of a variable (pricing in this case) would be if consumption patterns and market disturbances had the same characteristics as in the past. While this sounds much like what one can obtain from a simple linear regression of data, a stochastic model gives us a tool to model the patterns of volatility of the past and not just the general trend of data.

This presentation does not perform a detailed study of statistics or stochastic processes, but merely an brief overview of a stochastic model that can be employed to get the forecasting job done. Additionally, as is the case of many of the applications of math and science, “there are many ways to skin the cat.” Stochastic modeling and statistics are complex subjects with many tools. We select just a few of them here for illustration purposes. For a detailed presentation of these subjects, the reader is encouraged to review the books listed in the Reference section of this paper.

In general, a stochastic model is a mathematical construct that describes a process with uncertainty in its behavior. It employs the common statistical terms, such as means and variances, to characterize past behavior and the frequency of occurrence of otherwise unexpected events. Based on the specific nature of the system under study, the user may select from the various models or families of models to use. Common in the field are Poisson processes, Markov chains, Bernoulli processes, and Gamma processes, to name a few.

There are many aspects of the oil industry that display randomness (variation) and uncertainty. Some examples are inventories, oil production rates, crude oil qualities, available refining capacity, consumption of refined products, geopolitical events, ethanol (and now corn!) availability, and natural catastrophes. Attempting to model all these variables is not only a difficult task but an extremely inefficient undertaking. After all, the ultimate goal is to provide an adequate picture of the future pricing environment and not a prediction of the occurrence of specific events. The idea is to come up with a model that represents disturbances in general, without the need of identifying specific sources of those disturbances.

A practical, simple, yet extremely powerful stochastic model is the Poisson process model. In this case, the model predicts the occurrence of events as a function of time. This model is often known as an arrival-counting process. The time between the “arrivals” has a probability (or probability distribution) of occurrence that can be modeled as an exponential distribution. Arrivals, in this model, are assumed to be independent of the time since the last event. When this model is used to represent the time between discrete, random events (as opposed to continuous), it is referred to as a Bernoulli model.

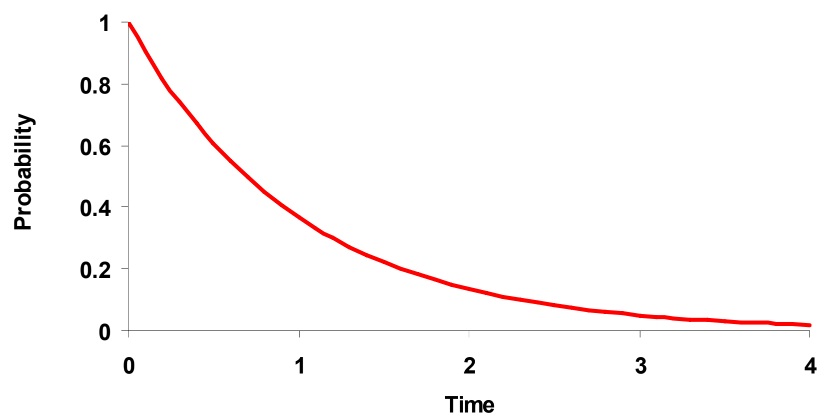
The occurrence of a major disturbance on the price of crude oil is an example of when the Poisson model can be used if the historical arrival times fit an exponential distribution. The most important assumption here is that the probability that something major may occur that will disturb the price of oil does not change from one day to the other. In other words, it assumed that the probability that a major natural catastrophe, major safety incident, or major geopolitical event may happen today is the same as it was yesterday or it will be tomorrow or the next day. The assumption is obviously not entirely correct when analyzing day-by-day events, especially since the likelihood of a major event occurring tomorrow changes if a major event occurred today. However, if the events are not coordinated, which is generally the case, this is a good approximation. Again, our overall purpose is to design a model that gives a reasonable picture of the state of things in the long run. In this case, it is sufficient to assume that in any given day, the industry has the same chance of witnessing a major price-changing event.

When this Poisson process is used to model the occurrence of a major disturbance, the probability of an event occurring in a given interval of time, as given by the Poisson process is given by the following equation (where λ is the historical rate of occurrences and x is the time between arrivals).

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & , x \geq 0, \\ 0 & , x < 0. \end{cases}$$

Graphically, the exponential probability distribution looks as displayed in Figure 6. Simple modeling allows one to determine the time between ***now*** and the "arrival" of the next event.

Figure 6
Exponential Distribution (λ = 1)



The exponential distribution is a key one because it appears commonly in modeling of pricing behavior. The normal distribution is often used to model certain events as well.

Ultimately, the forecasting exercise becomes one of breaking down the system under study to as many elements as practically possible, determining the specific stochastic model to employ in modeling each of them, and assembling the models to provide a simulation. A set of reference data is required. These data will represent past behavior and will be characterized using simple statistical metrics such as means (expected values) and variances.

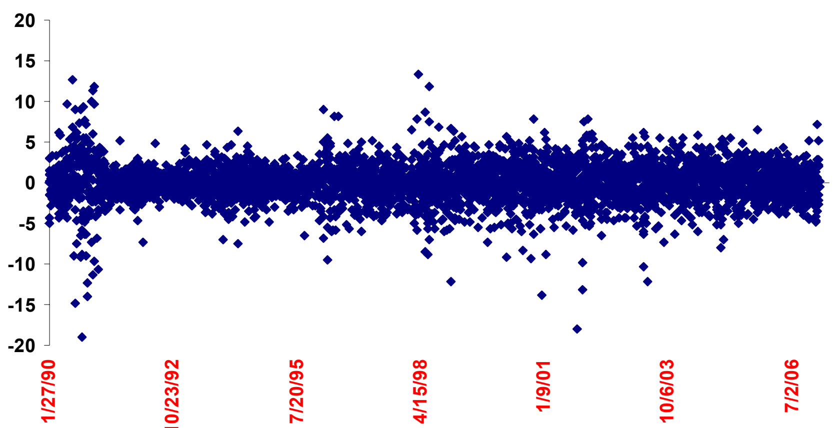
A simple spreadsheet can be used to implement an application although more advanced programming software makes the exercise a lot more enjoyable.

Case Study: WTI Cushing

An interesting example is the application of these models to the projection of the price of crude oil such as the common benchmark WTI Cushing. As discussed before, under the stochastic framework, there are many ways to approach this type of modeling. Here is one way.

First, let's study Figure 7, which plots the % change in price since 1990.

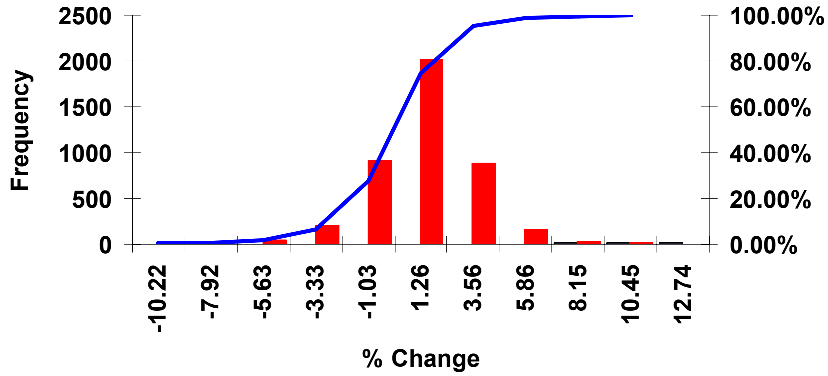
Figure 7
WTI Cushing Price % Change



From early 1990 until the end of 2006, the percent change in price of WTI Cushing stayed within a narrow band between -5% and 5%. The overwhelming evidence indicates that expecting the value to be within these two numbers was reasonable.

Figure 8 displays the histogram for the same data. As shown, a Gaussian (Normal) distribution is adequate in modeling the bulk of the randomness in price change.

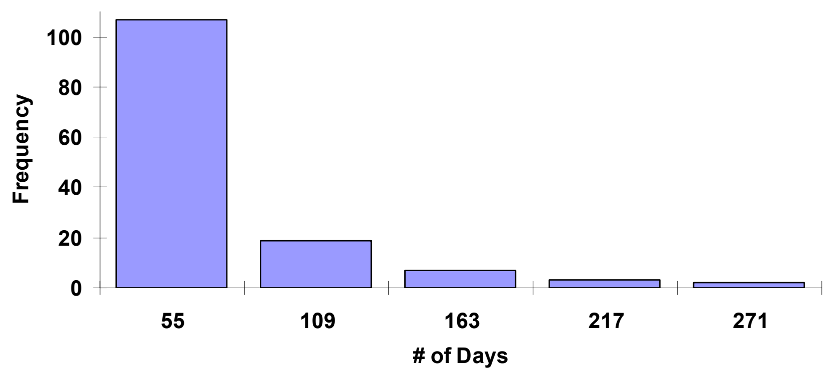
Figure 8
Histogram of % Price Change
(1990-2006)



Essentially, in the long run, it is adequate to state there is an underlying pattern of price change, perhaps dictated by the interaction between supply and demand forces in an environment without major disturbances, that can be modeled with a simple Gaussian model. The occurrence of major disturbances is a more interesting modeling case.

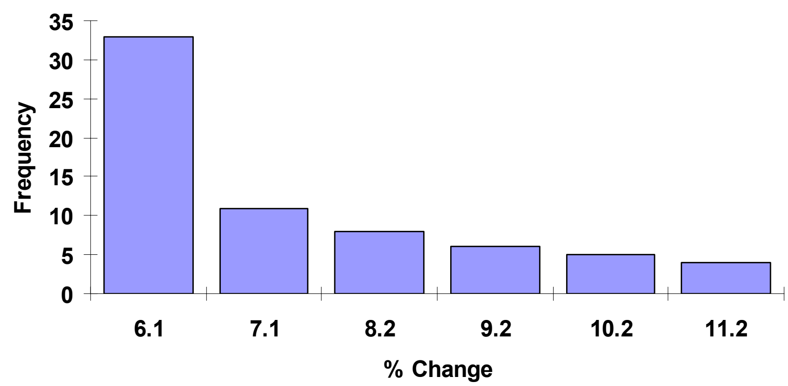
For the purpose of this discussion, I assume a major event to be any event with a price change greater than 5%. I believe this is a reasonable approximation since most of the variation between -5% and +5% appears to be explained well by the Gaussian model. By plotting the times between major events in a histogram (Figure 9), it is reasonable to state that the an exponential distribution is an adequate representation of the occurrence of these events. Therefore, assuming that these discrete events are independent from each other and the probability of their occurrence is the same on any given day, this process can be modeled as a Poisson process or the Bernoulli special case.

Figure 9
Histogram of Days Between
Major Price Changes



Additionally, by plotting the percent change in price for those occasions where major events have occurred (leading to both price reduction as well as price increases), we can conclude that they can be modeled with an exponential distribution to generate the magnitude of change when an event is predicted based on the simulation of the Poisson process. See Figure 10.

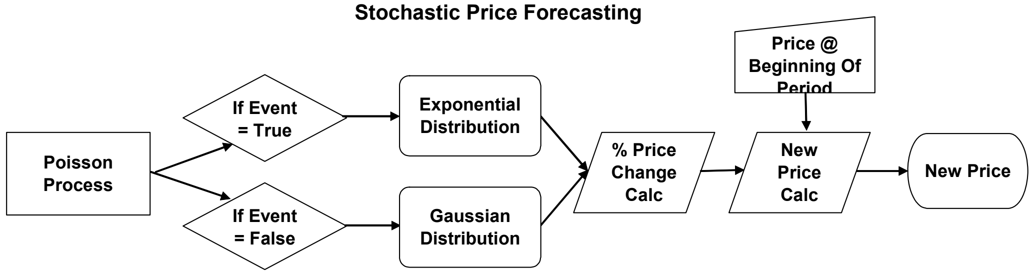
Figure 10
Histogram of % Change for
Major Price Changes



Putting It All Together

Once models for each piece in the overall pricing puzzle are determined, the models are assembled to produce a simulation of the entire pricing estimation process. Figure 11 is a flow chart depicting how this process would look in practice.

Figure 11
Stochastic Price Forecasting

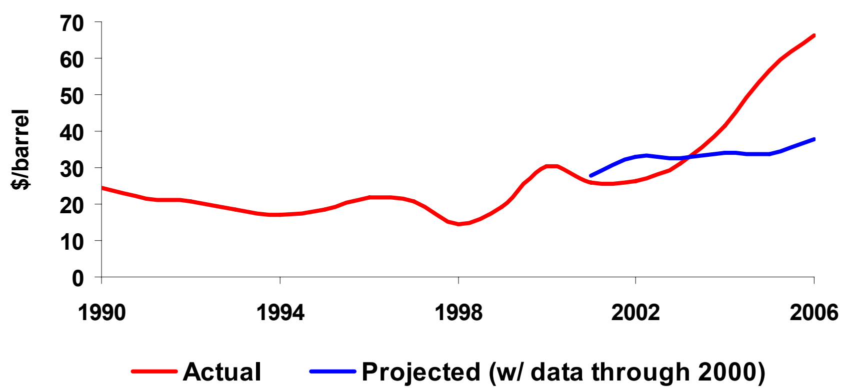


Two projections have been developed to demonstrate the forecasting tool. The first method employs data from 1990 to the end of 2000 to project pricing for the period between 2001 and 2006. The modeling parameters are displayed in Table 2. The results are displayed in Figure 12. The projected path did not follow the route established by the Rita and Katrina Hurricanes or the geopolitical tensions, but knowing the information we knew through 2000, the projection is fairly reasonable.

Table 2

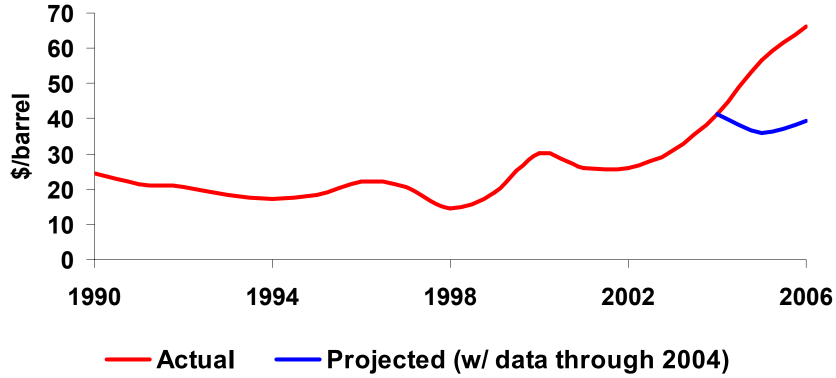
	1990-2004	1990-2000
Average Major Price Decrease	-0.079	-0.082
Average Major Price Increase	0.0703	0.0752
Average Normal Change	0.00066	0.00048
Average Standard Deviation of Normal Change	0.0184	0.0178
Average Days Between Major Price Decrease	60.74	63.48
Average Days Between Major Price Increase	87.77	89.05

Figure 12
WTI Cushing Pricing
Actual versus Projection



A second simulation was performed using the data for the period prior to 2004. I have developed the parameters displayed on Table 2 for modeling. Figure 13 displays the projected price increases for 2005 and 2006. Similar observations as before apply here. Although the projection is not an exact prediction, the results are adequate for the information known through the end of 2004.

Figure 13
WTI Cushing Pricing
Actual versus Projection



Limitations of the Models

Some of the inherent limitation of the models used in here have already been discussed. Namely, the probability of occurrence of non-random events do change after a previous event has occurred. This is not captured by the model presented here. Additionally, I have made an arbitrary selection of what an abnormal event is, and this obviously can slant the results of the simulation. Every model carries with it an assumption of certain characteristics of the data and the system in question. There is always the possibility that we have stretched the application of the model, but that is the risk with every model in mathematics and science. Finally, as the prediction runs further away from the last data point used to generate the data, it becomes less reliable because, as we have learned in the history of the industry, the environment surrounding the markets is never static and can change in a matter of months, weeks, or even days.

Summary/Conclusions

The task of forecasting prices of commodities is a challenging one. In a volatile environment, such as the oil industry, this process is often required to evaluate projects or financial strategies. For developing a sound pricing reference, one must employ a set of tools and industry facts that ensure a strong mathematical foundation while providing the essential balance of human, professional judgment.

This paper presents an approach intended to deliver such a pricing reference. It employs basic stochastic models, published information about capacity growth, official projections of demand, and the knowledge of industry experts that have witnessed and studied the intricacies of the supply/demand balance in the oil and refining industries. The model does present weaknesses as well as strengths. Ultimately, one must keep in mind that, after all, we are dealing with just a model and, therefore, the use of judgment is essential.

References

Nelson, Barry L. "Stochastic Modeling Analysis and Simulation," *Dover Publications, Inc*. Mineola, New York, 1995.

Ross, Sheldon M. "Applied Probability Models with Optimization Applications," *Dover Publications, Inc.* New Yorl, 1992.

"2007 Crude Oil and Refined Products Outlook," Turner, Mason & Company. Dallas, Texas, January 2007.