JACOBS Consultancy

White Paper on the Development of the Liquid Fuels Market Module (LFFM)

Prepared For

The Energy Information Administration

Submitted By

Vincent B. DiVita

Jacobs Consultancy Inc.
5995 Rogerdale Road
Houston, Texas 77072 USA
Phone: +832.351.7800
Fax: +832.351.7887

September 2009

Background & Overview

To give the reader a better understanding of the context in which this material was written, it is prudent to initially establish the author's point of reference and background. There are no external references cited in the paper because it is almost exclusively based on the experiences and opinions of the author. The opinions were requested by the EIA to facilitate a discussion on the development of the Liquid Fuels Market Module (LFFM), and not necessarily intended to be absolutes.

Some refinery experts develop highly sophisticated modeling techniques, include process simulators and highly advanced mathematical functionality, and equate model complexity with model integrity and accuracy. Other experts have experience in the use of LP models, yet lack the skills in code development or model debugging. They often have unrealistic expectations of LP models because they do not understand the trade-off between "wanting" the LP to perform a task and the associated code and matrix "baggage" for the task.

Other experts develop and tune their refinery-specific LP to be an exceptional tool for individual refinery planning purposes. This group often lacks experience in developing higher-level, regional, generic, or industry representations.

My LP and refinery modeling experience crosses many boundaries. I have developed models spanning the spectrum from high to low sophistication and specific single-client configurations to multi-regional representations, and I routinely develop, maintain and debug LP code. I put these models to use for analyzing downstream issues, including applications directly related to the questions being posed to the EIA.

With this as my background, my core modeling beliefs include:

- There is a delicate balance of sophistication, practicality, and versatility that must be maintained: "too much detail can be chaotic, too little detail can be catastrophic."
- A model must be usable and transparent: "the most sophisticated model in the world is useless if no one else can use it."
- LP model results should not be used in a vacuum—it is a quantitative tool that should be used in conjunction with qualitative analysis.
- To a large extent, "LP models are premise laden and assumption driven." You must understand the assumptions. When the model sophistication overwhelms the inherent LP assumptions, the degree of sophistication is probably too high.

- Because the LP is an economic tool, charged with optimizing to the highest margin for a given set of constraints and conditions, the "LP will turn the world upside down for a penny." Often, guidance is required.
- The LP is highly effective when used on a differential analysis basis.

Layout of LFMM Discussion

The Liquid Fuels Market Module (LFMM) material presented herein is somewhat of a shotgun approach. Based on guidelines and recommendations provided by the EIA, topics are highlighted, a discussion is presented, and then we move to the next topic. A few topics have been introduced by the author.

Background to the topics, as represented by the EIA, are identified in italic print and enclosed by a border.

Acknowledgments

I wish to thank Mr. John Jenkins (Director, Jacobs Consultancy) for his contributions.

Technology Representation

The EIA has been explicitly asked in the past to provide cost estimates for various effects due to EPA-proposed rules, and will likely be asked to provide such analysis in the future. Additionally, it is expected that even when not asked specifically, the EIA should have appropriate representations of refining technologies to respond to changes in specifications, particularly as the responses across world oil and other alternative cases could vary dramatically. Should the EIA rely on more complex models (e.g. GRTMPS) for this type of analysis alone, attempt build in at least a moderately sophisticated refinery processing engine, or have at least a representative unit designed and called that could inform a larger model?

It is appropriate to begin this discussion with an explanation of model types generally used to answer questions posed to the EIA.

- Regional
- Generic Configuration
- Specific Refinery Configuration

Regional Models—A regional model can take on many forms. It could be a PADD level model such as PADD 3; a sub-district model, such as Texas Gulf Coast; or a combined US refining industry model. The regional models generally take the total sum of the processing capacity of the region to establish the base configuration. Regional crude quality (API, sulfur, imports) and regional refinery production data are readily available, making the calibration case relatively easy to establish.

Regional models are well suited for supply/demand issues. These models can also be used to analyze regulatory impacts; however, the resulting answer should be considered an average result. After all, when all the specific capacities in PADD 3 are combined into a regional PADD 3 LP, that is, in fact, an average representation.

Generic Configuration Models—Some generic configuration models include a Coking configuration, a Fluid Catalytic Cracking (FCC) configuration, a highly complex coking/FCC/Hydrocracking configuration, or a low conversion hydroskimming configuration. These models are typically developed using a nominal crude throughput (100 MBPD), and generally begin with unconstrained downstream process throughputs. The downstream capacities are strongly related to the crude slate assumptions, so it is often difficult to pre-judge the downstream process capacities.

Generic configuration models are effective to analyze specific regulatory impacts. These representations bring a higher degree of fidelity and resolution to the analysis versus a regional representation. A regulatory impact on a cracking configuration will be slightly different than a coking configuration, and this method captures this difference.

Two identical generic configurations can have significantly different results due to the crude slate assumptions. In fact, the crude input assumption is often as critical to the analysis as the LP model configuration assumption. Thus, when developing configuration models it is often important to run the models using different crude mixes.

Specific Refinery Configuration Models—A specific refinery configuration is an even more detailed representation of an actual refinery operation. The specific refinery configuration essentially represents the planning LP at a specific refinery. These models are typical for single-client work. Critical operational data are confidential. Each refinery has specific parameters and constraints that impact operations. At this level, no two refineries in the US are alike. For a variety of reasons, it is somewhat impractical to suggest that the EIA develop specific configurations to represent each refinery in the US.

Each of these model configurations has merits to the questions posed to the EIA. It is my opinion that the EIA should focus on the development of a number of regional and generic configurations using an advanced refinery LP modeling system such as GRTMPS. With the advances in LP modeling (software and hardware) and the

complexity of the questions being asked, it is difficult to imagine performing an impact analysis without the use of a strong, robust LP refinery system.

A reasonable goal would be to develop these models as part of an effort to have a collection of LP tools—an LP "toolkit." It is impractical to envision a single model that could do everything, which is why I strongly support an LP toolkit approach.

Depending on the questions being asked and the analysis requirements, these specific tools could be used off-line for specific analysis. Some questions require a first pass look, others require a moderate approach, and still other questions require an extensive study—all depending on the resource (time and budget) allocations. Having an LP toolkit is a very practical approach to begin any downstream analysis, and can initiate on a simple level, and expand to a complex level. The tools and analysis can be used to support details of a larger model such as the LFFM and NEMS.

Some of the models that would be introduced to the toolkit include:

- Cracking Configuration for a light & medium crude slate
- Coker Configuration for a medium & heavy crude slate
- Coker Configuration with Hydrocracker for a medium & heavy crude slate
- Regional Configurations include PADDs 1, 2, and 3. PADDs 4 and 5 could come later, and it is probably more important to model California versus an inclusive PADD 5.
- US Regional Configuration

A diverse set of crude assays reflecting domestic and foreign crudes should be created. This is extremely important. LP models initiate with crude assays, meaning assays with suspect, incomplete, or incorrect data can potentially damage the integrity of the LP results. There are a couple of major crude assay libraries that can be purchased. Additionally, there is crude assay management software that makes crude recutting and assay modifications extremely user-friendly.

I can simplify the "universe" of US crudes to a number of assays between 20 and 40, and closely match API gravity and sulfur of the actual regional imports and domestic qualities.

Additionally, the configurations outlined above should include PADD 1, PADD 2, and PADD 3 as a minimum. The Cracking configuration above should therefore translate to three separate regional cracking configurations: CRK PD1, CRK PD2, and CRK PD3.

Over-Optimization

The question is often posed of how to deal with the over-optimization of the regional model. Reiterating, the following discussion relates to a regional model. I am convinced that the regional model should have a degree of over-optimization. By its nature, the regional model is trying to capture the trend of a number of specific refineries that have been condensed down to a single configuration. For example, a PADD 3 regional model could consist of the following configurations:

- Eight topping refineries with about 160 thousand BPD
- Seventeen cracking refineries with about 1.6 million BPD
- Twenty-six coking refineries with about 6.7 million BPD

The total of the PADD 3 regional model would represent about 51 refineries with 8.4 million BPD. Fifty-one refineries, condensed down to a single LP model, should have a degree of over-optimization because the capability of each of the refineries will be different, and over-optimization is how this is captured.

The more difficult question to answer is how much over-optimization is realistic, for which there is no straightforward answer. Ultimately, this decision is held with the model developer.

Crude slate characterization is a critical issue in the over-optimization question. Using the example above, the average crude input into PADD 3 is about 29.5 API and 1.7 sulfur, and about 75% of the crude is imported and 25% is domestic. The number of specific crudes reaching these refineries is obviously high, and it would be impractical to suggest giving the LP hundreds of crude purchase options. With this in mind, how does a modeler deal with and simplify the overwhelming number of actual crude purchase options in a regional model?

The crude purchase simplification is generally accommodated with the following techniques:

- Allow a number of specific crude purchases—for example, Saudi Lt, WTI, and Maya are all specific crude purchase options.
- Aggregating crudes by type—for example, combining Bonny Lt, Cabinda, Brent, LLS, and WTI to establish a single Light Sweet Crude.
- By blending crudes to match the actual crude characteristics of the region. This can be done in a number of ways, such as a single regional blend, a light sweet +

heavy sour blend, a domestic blend + import blend, *etc*.—the potential combinations are almost unlimited.

Regardless of how the crude purchase decision is simplified, the fact is that the simplification process has already removed a tremendous amount of "real-life" optimization of crude purchases to the region. Restated, this crude simplification step dilutes the actual crude purchase options, consequently diluting the actual "real-life" crude optimization for the region. One could therefore argue that since the crude purchase simplification has removed a degree of over-optimization, the model should have some over-optimization to compensate.

Differential Analysis

LP differential analysis topic is critical to this writing because many of the issues posed to the EIA require differential analysis.

Using the differential analysis concept, a base case or calibration case is initially developed. This can take many forms. At the PADD level, a reasonable calibration is based on PADD refinery input and production data. At the refinery level, the base case is typically the existing LP planning model. The purpose of the base case is to establish current operations.

Scenario cases are then developed from the base case. Scenario cases can take on many forms, are typically developed to analyze a specific issue, and often look into the future. Example scenarios include changing gasoline to 30 ppm, producing ULSD, changing operations in different price environments, and reducing benzene in the gasoline pool.

The analysis performed compares the scenario case to the base case, providing insight to the scenario question and how it might impact the refinery or region versus today's operations. With differential analysis, model oversights—even outright errors—tend to cancel out.

As an example, assume a 10 ppm gasoline sulfur analysis is being studied. Further assume that the LP delayed coker prediction contains a 5% coke production error. When comparing the base case to the 10 ppm scenario case, it is unlikely the 5% coke error will impact the differential analysis because it exists in both cases, and for all practical purposes the delta coke production will be negligible. Furthermore, delayed coke production is fairly independent of the 10 ppm analysis.

To drive the point home, assume the base case is running 18% Conradson Carbon (CCR—a key predictor in coke production) to the coker. If the scenario case runs 25 CCR, the LP will recognize the 7% CCR difference and predict a new coke production differential based on additional CCR in the feed. So, the differential analysis is still likely reasonable, even with the 5% coke production error in the base case. This 5% coke production error would be extremely problematic if using the results to design coke drums, but is not such an extreme issue on a differential analysis basis.

Obviously, we would never use the differential analysis concept as an excuse to develop unrealistic or sloppy LPs; in fact, in the above example, a good calibration model would have prevented the 5% error in the first place. The EIA has stated:

> *"It is very easy for simple mistakes to become part of the database for years before they are discovered."*

Part of this statement may stem from the fact that differential analysis can make mistakes difficult to uncover. This is unlikely to be the sole reason because other factors can exist: model complexity, lack of transparency, a legacy model, and lack of routine maintenance, to name a few.

Seasonal Issues

Refiners face different operational changes associated with seasonal issues. These often occur monthly, but for discussion purposes we will call them summer and winter.

For most situations, I prefer to develop an LP with two separate time periods, reflecting separate summer and winter seasons. I do this for a variety of reasons, including

- Seasonal specifications change for both gasoline and distillate
- Prices often change on a seasonal basis (strong summer gasoline prices, for example)
- Refinery operations often change seasonally, reflecting seasonal supply/demand patterns
- Complex Model Phase 2 equations have unique summer and winter calculations

One can choose to model an annual average basis; however, my experience is that running a 2-period model is worth the extra effort for accuracy of analysis and should be considered for certain applications and issues posed to the EIA.

It may be suitable to keep an average specification in larger long-term models, but for specific regulatory or industry impact analysis using generic configuration models, I recommend a 2-period LP.

Marginal Value Analysis

An LP model should calculate and report marginal values on feedstock purchases, product sales, process expansion incentives, and specifications. Marginal value analysis is fundamental to any LP analysis.

Refinery Crudes and Cuts

One of the key limitations of the existing approach is that the cut points for each crude cannot be reset without having to recalibrate all subsequent intermediate streams coming downstream from the ACU and vacuum unit. As cut point manipulation is likely key in increasing diesel yields, the model's current inability to do this may be a significant limitation.

The existing approach is complex most likely because the crude cuts and downstream processing unit feed streams are not pooled. A recursed pool combines the individual feeds into a single stream, and the LP will calculate and track the qualities of the pool based on the qualities and volumes of the individual feeds.

Additionally, this limitation might exist because of the lack of swing cuts. Swing cuts, as the name implies, allow a cut to swing up or down, based on the model optimization. Swing cuts allow the simulation of cut point optimization without the need to re-cut the crude assay.

While it is true that swing cuts are not exactly mathematically precise, it is a well accepted industry methodology. The reason a swing cut is not mathematically precise is that the swing cut carries the average quality of the cut. For example, on a heavy naphtha/jet swing cut (350˚F – 400˚F), the properties of the cut are approximately calculated at the midpoint (350˚F + 400˚F)/2 = 375˚F. In a sophisticated crude software platform, it is not quite this "simple," but this explanation will suffice for discussion purposes. The crude cutting software we use is Haverly's Crude Assay Management System (HCAMS).

If half of this cut goes to naphtha and the other half goes to jet, that would represent a naphtha endpoint and jet initial point of 375˚F. In reality, the quality of the swing cut

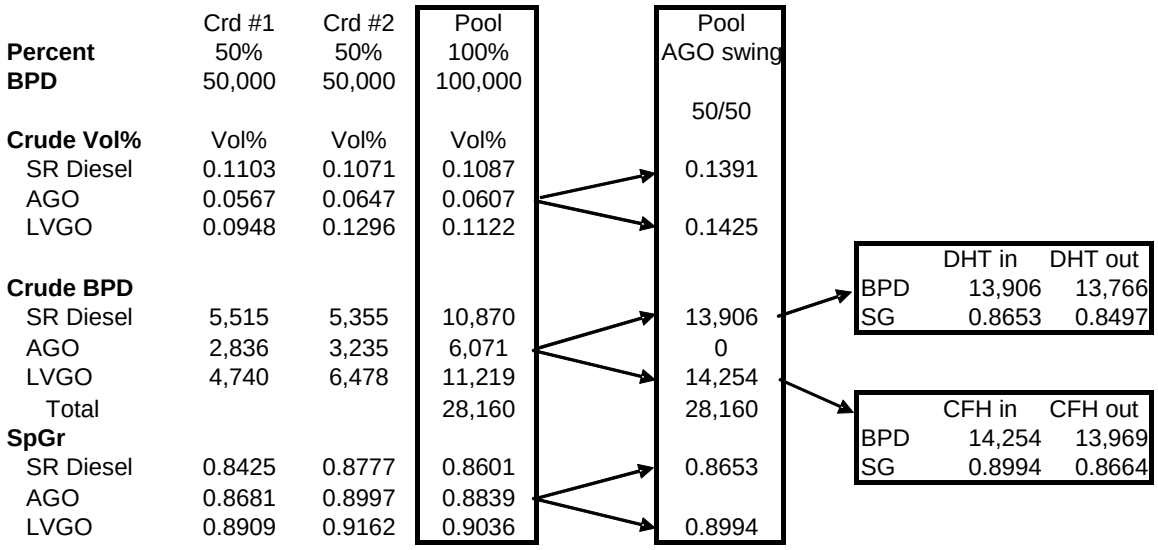
going to naphtha should represent the front-end (350˚F + 375˚F)/2 = 363˚F, and the quality of the cut going to jet should represent the back-end (375˚F + 400˚F)/2 = 388˚F. Yet the swing cut methodology calculates the quality swinging up and the quality swinging down at 375˚F. This is a slight oversimplification, but it is a standard technique that should be employed for EIA modeling purposes.

There are more sophisticated crude cutting optimization schemes; however, the mathematics behind the algorithm is extremely sophisticated, can create convergence issues, and in my opinion is best applied to a refinery-specific planning LP, and is not well suited for generalized modeling.

Below is an example of straight run (SR) cuts and swing cuts for generalized modeling. The initial boiling points (IBP) and end points (EP) are intended to represent a reasonable set of crude cuts, not necessarily the absolute recommendation for the EIA.

Swing Cut Example			
	Initial BP (F)	**Final BP (F)**	
C4 and Lighter			
Light Straight Run	49	160	
Lt Naphtha	160	220	
Hvy Naphtha	220	350	
Naphtha/Kero Swing	**350**	**400**	**Swing**
Kerosine	400	525	
Kero/Diesel Swing	**525**	**550**	**Swing**
St Run Diesel	550	650	
St Run Atm Gasoil Swing	**650**	**690**	**Swing**
LVGO	690	850	
HVGO	850	1000	
Vacuum Resid Swing	**1000**	**1025**	**Swing**
Atm Resid	700+		
Vacuum Resid	1025+		

Below is an example of recursion pooling and swing cuts, using SR diesel, AGO swing cut, and LVGO:



	Crd #1	Crd #2	Pool	Pool
Percent	50%	50%	100%	AGO swing
BPD	50,000	50,000	100,000	
				50/50
Crude Vol%	Vol%	Vol%	Vol%	
SR Diesel	0.1103	0.1071	0.1087	0.1391
AGO	0.0567	0.0647	0.0607	
LVGO	0.0948	0.1296	0.1122	0.1425
Crude BPD				
SR Diesel	5,515	5,355	10,870	13,906
AGO	2,836	3,235	6,071	0
LVGO	4,740	6,478	11,219	14,254
Total			28,160	28,160
SpGr				
SR Diesel	0.8425	0.8777	0.8601	0.8653
AGO	0.8681	0.8997	0.8839	
LVGO	0.8909	0.9162	0.9036	0.8994

	DHT in	DHT out
BPD	13,906	13,766
SG	0.8653	0.8497

	CFH in	CFH out
BPD	14,254	13,969
SG	0.8994	0.8664

This example has two crudes in a 50/50 ratio in a 100 MBPD crude tower. The LP produces 10,870 BPD of diesel, 6,071 BPD of AGO, and 11,219 BPD of LVGO. The AGO can swing up or down, and we assume 50% goes up to diesel and 50% goes down to LVGO. The recursive nature of pooling will calculate the new quantities and qualities of the diesel and LVGO. Note that after the AGO swing the SR diesel becomes heavier and the LVGO becomes lighter, precisely what should occur.

These pools go farther downstream, where the diesel goes through a ULSD HDT that increases the API gain by three with 99% recovery. The LVGO goes through a VGO HDT that increases the API gain by six with a 98% recovery. Because the feeds to the downstream units are recursed pools, the downstream units can easily deal with the quality changes across the units, provided the database and sub-modules are properly designed.

Based on the optimization, the LP will determine how much and where to swing the AGO. This decision will be driven by economic factors, operating costs, blending limits, price of products, and many other LP-related conditions and constraints.

The key point is that with properly constructed crude cuts, swing cuts, recursive pooling, LP software, crude software, crude assays, and a properly designed database, the EIA's existing challenges with crude representation are dramatically reduced. It is also likely that this methodology will not only be more user-friendly, but will be more accurate than existing methods and be far less prone to user errors.

In a previous section I stated the importance of quality assays, and that the "universe" of crudes can be simplified for generalized modeling purposes. The next two tables provide

examples of these simplifications using a “high” and “low” number of assays. At the bottom of the table, the Target and Actual API and sulfur are provided. Although all PADD regions were calibrated to similar accuracy using different ratios of the assays, the total US is shown.

LARGE CRUDE ASSAY REPRESENTATION	
FOREIGN CRUDES	**DOMESTIC CRUDES**
ARABIAN LIGHT	ANS
ARABIAN MED	BELRIDGE S.
ARZEW/SAHARAN	COLORADO SWT
AZERI LT	CYMRIC
BASRA LT	ELK HILLS
CABINDA	KERN
CANO LIMON	LLS
DOBA	LOST HILLS
FORCADOS	MARS
FURRIAL	MONTANA MIX
HAMACA	MYTON
HARDISTY LT MIX	THUMS
HIBERNIA	WTI
INTERPROVINCIAL SWT	WTS
ISTHMUS	WYOMING SOUR
KUWAIT	WYOMING SWEET
MARLIM	
MAYA	
MEREY	
OLMECA	
ORIENTE	
PEACE P/L SOUR	
QUA IBOE	
RABI LT	
SYNCRUDE	
WCS	
	TOTAL
API Calc'd - Total	30.2
API Target - Total	30.2
Sulfur Calc'd - Total	1.5
Sulfur Target - Total	1.5

SMALL ASSAY REPRESENTATION	
FOREIGN CRUDES	**DOMESTIC CRUDES**
CABINDA	ANS
FURRIAL/MESA	KERN
HARDISTY	MARS
HIBERNIA	THUMS
MAYA	WTI
MEREY	WTS
ORIENTE	
QUA IBO	
SAUDI MED	
SYNTHETIC	
WCS	
ZARZAITINE	
	TOTAL
Calc API, Total	30.3
Target API, Total	30.4
Calc SUL, Total	1.5
Target SUL, Total	1.4

Model Sophistication

For new alternative processes, how sophisticated should the models be? Are simple input/output models sufficient to represent alternative fuels processes (particularly in cases where a detailed carbon mitigation response is desired)?

The model should be sufficiently sophisticated to answer the question at hand. However, my definition of sophistication will not likely match anyone else's definition.

There are certain situations when a "black-box" simple input/output model can be sufficient. This might occur initially in the development of a process when minimal operating data are available. As a minimum, the output product qualities should carry data to calculate the final destinations. Biodiesel, for example, must carry all the blend qualities to calculate final diesel blends.

As more data are available, and if there is a significant difference in the yields and product qualities based on feed, the "black-box" method will be updated with enhanced predictive capabilities.

I recommended an LP toolkit approach that includes generic configurations and regional configurations. Generic configurations are well suited for analyzing refinery-specific issues and regional configurations are well suited to model supply/demand regional

issues, although both can be used interchangeably, depending on the issues being analyzed.

Additionally, it is easier to work with configuration models to analyze new process technologies. The code, the actual model runs, and the analysis are all often easier to perform versus a large regional LP model.

Any refinery representation used in the LFFM will be a generalized approach. This is an important concept. The EIA does not own a refinery and will not be optimizing a specific refinery location, but the EIA is interested in understanding the impacts to the industry.

The Technology Database (discussed in the next section) should include the capability to "look around the corner," either with a base and delta vector yield structure or similar process unit representation. Single yield vectors do not make sense when feed pools are employed, and are overly simplistic. Nor is it practical to include dozens of possible yield operations. The configurations are initially intended to represent the "average" operations.

From this average point, a modeler can modify the LP options to represent differences between "good" and "bad" operations. For example, one can limit the volume of AGO allowed to swing up to diesel, or recut the AGO EP crude from 690˚F to 670˚F to represent "sloppy" fractionation.

The base and delta vectors should be developed using off-line, robust engineering methods. Although it is possible to tie-in advanced simulators to the LP model, for EIA modeling I would not recommend coupling outside simulators to the LP model. This advanced technique is most practical for refinery-specific applications, not generalized modeling.

It is unrealistic to answer *"how sophisticated should the model be"* in general terms because each process unit and stream quality should be analyzed independently. Additionally, it often changes based on the question being asked. For example, a ULSD HDT can be designed across a range of pressures, such as from 700 psi to 1800 psi. Should the LP have specific vectors for 700 psi, 800 psi, 900 psi, all the way up to 1800 psi? There are various ways to construct this, including:

- Choose an "average" pressure to develop the process submodel
- Provide a low and high pressure operation and allow the LP to optimize (for example, between 800 psi and 1500 psi)

The potential issue with the low and high options is that the LP might optimize on the high pressure, but unless the capital investment cost is built into the LP code, the LP will not recognize that the high pressure unit requires more capital versus the low pressure operation (there is a pressure/capital trade-off).

For each pressure, there are potentially many different dependent variables, including hydrogen uptake, operating cost, cetane improvement, API gain, and yield. If the study question was to analyze the impact of producing ULSD, I would suggest the LP code might be updated to more rigorously model the ULSD options. If the study question was to analyze the impact of producing 10 ppm gasoline, it should suffice to keep an average/generic ULSD hydrotreater as the base LP code.

Model sophistication is a complex topic that could be the subject of a stand-alone White Paper. Rather than delve into this topic, I will conclude with the following. The EIA model should have a sound basis for the process submodels, both existing process and new / still to be developed process. The recommended predictive approach includes base and shift vectors developed from a sophisticated engineering source, and modified to suit an LP vector-based approach. If sophisticated data are not available, then the process might begin with “black-boxes” and expand in sophistication as data become available. My opinion is that an LP should not have dozens of single-yield vectors; rather, it should have a base and shift vectors to compensate for feed changes, which presumes recursive pooling.

An example of a generalized database structure for the FCC—and possibly the most complex structure in most LPs—can be represented by:

STRAIGHT RUN (SR) FEEDS	**HYDROTREATED (HDT) FEEDS**
HIGH CONVERSION BASE VECTORS (MAX GASOLINE)	
HC FCC Base	HC FCC Base
HC FCC High	HC FCC High
HC FCC Low	HC FCC Low
LOW CONVERSION BASE VECTORS(MAX DIESEL)	
LC FCC Base	LC FCC Base
LC FCC High	LC FCC High
LC FCC Low	LC FCC Low
SHIFT VECTORS	
SR Feed Sulfur	HDT Feed Sulfur
SR Fd Nitrogen Shift	HDT Fd Nitrogen Shift
SR Fd Density Shift	HDT Fd Density Shift
SR Feed CCR Shift	HDT Feed CCR Shift
SR MBP Shift	HDT MBP Shift
SR Diesel Shift	HDT Diesel Shift

The table shows a duplicated set of vectors for straight run (SR) feeds and hydrotreated feeds (HDT). These two types of feeds have significant yield and product qualities off the FCC, hence the need to take this difference into account.

For each SR or HDT mode above there is a High, Medium, and Low severity, which essentially is a change in FCC riser temperature, which will change the FCC yield pattern.

The example above goes a step farther by defining a high conversion (maximum gasoline) mode and a low conversion (maximum diesel) mode. This mode is useful to capture optimal operations in a high gasoline or high diesel price environment.

Lastly, there are a number of shift vectors representing key feed qualities that will impact FCC operations. This can be designed in a number of ways, depending on how the FCC LP code is developed. Some might prefer using a K-factor instead of Density and Boiling point. Others might employ refractive index as a shift basis. The point is that an LP should compensate the process yields and product qualities based on the feed characterization. For example, FCC feeds with high CCR, high nitrogen and high aromatics should lower the conversion.

All of these base and shift vectors effectively work in a recursive model because the intermediate steams and process feeds are pools. This is why recursive pooling is very important for a generalized database.

Technology Database

The existing PMM uses the data structure (OMNI) inherited from the ENSYS WORLD model which has been transformed to be read using Ketron DATAFORM based utilities. This representation lacks transparency and is difficult to maintain or solicit peer feedback on. Additionally, the structure has no mechanisms to query input for reasonableness

Whatever model design is ultimately chosen, some time should be spent discussing: the manner in which the technical data for various liquid fuels processing units can be stored and accessed; what attributes it should have; and if there is potential to have at least this element part of an existing model that has been vetted (and possibly maintained) by a number of knowledgeable users.

I will begin this discussion on the issue that the technology database should be evergreen: changes are inevitable. The database should undergo maintenance to keep up with computer software and hardware capabilities, as well as refinery technology.

This does not necessarily translate into “keep adding more code”; rather, it can translate into “updating code.”

The models should also undergo maintenance. For regional models, part of the maintenance issue is annually recalibrating based on the new regional balances. The US refining industry evolves and the models should evolve as well to capture changes in crude quality, product slates, capacity, new specifications and actual downstream throughputs to name a few.

This emphasizes the following:

- The database should be generalized, robust, and agile.
- The database, while generalized, might need updates from time to time as the requirements and emphasis of the LP modeling changes.
- The database and modeling techniques do not need to capture the “last decimal point.”

The EIA made the following comment regarding the database:

> *For example, a volume of 10 parts ethanol with a Reid Vapor Pressure (RVP) of 2 psi combined with 90 parts gasoline with an RVP of 2 psi could yield a product with a volume of less than 100 parts and an RVP greater than 8 psi.*

This example suggests that 10 parts ethanol plus 90 parts hydrocarbon is less than 100, which is a statement I have not seen translated into LP code. The overall product blending losses (gasoline & diesel) will probably be greater than this volume shrink impact.

The second point of the RVP example is a detail that should be captured. Even then, there are a number of different methods to deal with RVP, some more complex than others. This is a situation where the LP code must strike a balance between accuracy and complexity.

Speaking to the points above, the technology database has changed over the years to keep up with the questions being asked. When 30 ppm gasoline sulfur standards were being discussed, most LPs had to add structure and technology to deal with this analysis. The same argument can be made for other historical regulatory changes, such as ULSD production, RFG production, and MSAT compliance. Looking forward, some refiners are currently studying the issues associated with producing 10 ppm mogas; this will also likely add additional structure to the LP.

At the expert level, it is necessary to understand the basic programming nature of the LP, regardless of the code. From this perspective, technology databases do in fact lack transparency if one cannot read and interpret the code. I have found it extremely useful to keep “working sections” embedded in the technology code as well as documentation, which can take on many forms.

My preference is to keep the database in a spreadsheet format, which allows these “working sections” to exist. I have also translated spreadsheet databases into a true database environment. There are pros and cons to each approach: the database lacks transparency and is difficult to predict the outcome of the submodels without a spreadsheet; the spreadsheet is more transparent and can be modified to include “transparent” working sections, but is less forgiving on LP input and code errors.

In any event—whether the technology is stored in a spreadsheet or database format—there needs to be a mechanism to analyze and check the reasonableness of the code and predictive nature of the technology submodel. This can and must be a transparent process so that a non-LP code expert can evaluate the process technology.

The technology database must be maintained. This maintenance includes checking yields, qualities, technologies, utilities, and updating as data become available. For example, some catalyst vendors have recently developed new catalyst formulations to maximize diesel production. Consequently, the LP experts maintaining the model must make the decision whether to update the yields for the new catalyst, add new vectors for the catalyst, or whether the existing vectors are still reasonably accurate.

When refinery experts check the process units for reasonableness, there are, for the most part, some generally accepted standards for process modeling. A few examples are listed below:

- A reformer should include some form of feed characteristics relating to paraffins, naphthenes, and aromatics. For benzene prediction, some might include benzene precursors (*e.g.,* methylcyclopentane, cyclohexane, and benzene).
- A coker should include feed qualities such as CCR, density, and nitrogen.
- An FCC should properly characterize the feed, of which there are many different methods, including: K-factors, density, boiling points, nitrogen, metals, CCR, and aromatics.

It is unnecessary to carry additional qualities if it does not impact the yield or shift vectors, unless one wants to report these qualities. For example, it is not necessary to carry refractive index in the FCC feed if there is not an associated refractive index shift.

Restated, if the process submodel does not “need” the quality to predict a yield, then carrying the extra quality in the LP essentially takes up matrix space.

As a starting point, the data below suggest a reasonable property/quality set for LP streams:

QUALITY	Lt Lt St Run Naptha	Lt St Run Naphtha	Hv St Run Naphtha	Nap / Kero Swing	St Run Kero	Kero/ Dsl Swing	St Run Diesel	St Run AGO	Atm Resid	LVGO	HVGO	Vac Swing	Vac Resid
All Streams													
Specific Gravity	1	1	1	1	1	1	1	1	1	1	1	1	1
Sulfur, wt%	1	1	1	1	1	1	1	1	1	1	1	1	1
Sulfur, wppm	1	1	1	1	1	1	1	1	1	1	1	1	1
Heat Content (MMBTU/Bbl)	1	1	1	1	1	1	1	1	1	1	1	1	1
Molecular Weight	1	1	1	1	1	1	1	1	1	1	1	1	1
Gasoline Streams													
Reid Vapor Pressure, psi	1	1	1	1									
V/L	1	1	1	1									
Research Octane No	1	1	1	1									
Motor Octane No	1	1	1	1									
Road Octane No	1	1	1	1									
Dist: T10	1	1	1	1									
Dist: T50	1	1	1	1									
Dist: T90	1	1	1	1									
Driveability Index	1	1	1	1									
Dist: % Evap at 200F	1	1	1	1									
Dist: % Evap at 300F	1	1	1	1									
Benzenes, vol%	1	1											
Benzene Precursor Index	1	1	1	1									
Napthenes	1	1	1	1									
Oxygen, wt%	1	1	1	1									
Alcohol, Vol%	1	1	1	1									
Gasoline/Distillate Streams													
Parrafin	1	1	1	1	1	1	1	1					
Olefins, vol%	1	1	1	1	1	1	1	1					
Aromatics, vol%	1	1	1	1	1	1	1	1					
Distillate & Heavier Streams													
Freeze Point, degF				1	1	1							
Cetane Index D976				1	1	1	1	1					
Cetane Index 4737-A				1	1	1	1	1					
Cetane Index 4737-B				1	1	1	1	1					
Pour Point, degF				1	1	1	1	1					
Cloud Point, degF				1	1	1	1	1					
Flash Point, degF				1	1	1	1	1	1	1	1	1	1
Viscosity, CST at 122F				1	1	1	1	1	1	1	1	1	1
Viscosity, CST at 210F				1	1	1	1	1	1	1	1	1	1
Vanadium, wppm								1	1	1	1	1	1
Nickel, wppm								1	1	1	1	1	1
Naphthenic Acid kOH/mg				1	1	1	1	1	1	1	1	1	1
Concarbon								1	1	1	1	1	1
Nitrogen or Basic Nitrigen								1	1	1	1	1	1
FCC crackability (user defined)								1	1	1	1	1	1

In our regional modeling efforts and in our generalized database, we characterize our base yields as “average” yields. While it is difficult to exactly define “what is average technology,” one is forced to make this estimate, based on experience and commonly accepted industry standards. Surveys, articles, yield correlations, simulations and software databases often provide a good source for yield and product qualities.

If the technology database includes base yields (average) and shift vectors, the shift vectors capture changing feed qualities, and the robustness and agility of the model is greatly enhanced.

Product Blending

Since NEMS was designed in 1992, gasoline and diesel specifications have tightened considerably, gasoline and diesel types have proliferated, methyl tertiary butyl ether (MTBE) has been banned in eighteen states, and ethanol is increasing its share of gasoline supply. These changes have led to Congressional and Administration requests for more detailed modeling, which NEMS either cannot perform or must perform with difficulty. The problem with most petroleum product specifications is that they do not vary linearly with volumes of blend streams.

For this reason they cannot be modeled completely accurately in a model that requires all functions to be linear.

One possible way to deal with this problem would be to construct a blending sub-module. This sub-module would track how product blends must change to meet policy changes in specifications and would allow for a more accurate representation of the cost and consequences to implement such a policy.

The existing model uses several simplifying assumptions to produce gasoline of "average quality" (Octane, RVP, etc.). The blending formulations and costs do not behave in a linear fashion with respect to these properties.

Given the desire to expand the model to possible 2050, the model should also be adaptable to include speculative fuels and technologies (Algae Diesel, HCCI fuels, etc.).

Average Gasoline

The very nature of a regional model will produce gasoline of average quality. If separate configuration models are developed, one would be able to look at subtle differences in gasoline blends from a Cracking or Coking refinery—again, this will be "average" gasoline. Among other things, gasoline blend qualities will likely be more influenced by the crude slate than the configuration. Restated, gasoline from a Cracking refinery running Light Sweet Crude will blend differently than gasoline from a Cracking refinery running Medium Sour crude.

Refinery-specific planning LP models produce gasoline of average quality. Many refiners will run a 12-period model, representing each month of the year. The gasoline blend results for January will represent the average gasoline blend for the month, not the individual batch or on-line blends performed on a daily basis.

In most of my modeling efforts, I develop 2-period models to capture summer and winter seasons. These models produce "average" gasoline seasonal blends, and provide reasonable representations of actual blends because of the complete specification slate for the gasoline, the variation between seasons, CG and RFG grades, and regular and premium grades. Typical blends include:

- RFG regular grade, typically 87 (R+M)/2 (summer & winter)
- RFG premium grade, typically 93 (R+M)/2 (summer & winter)
- CG regular grade (summer & winter)
- CG premium grade (summer & winter)
- Oxygenated CG regular (summer & winter)
- Oxygenated CG premium (summer & winter)

Practically speaking, most of the US gasoline modeling under the RFS standard is 10 vol% oxygenated (ethanol) E10 CG, and 10 vol% oxygenated (ethanol) RFG. We do not typically model mid-grade gasoline, because it is a small percent of gasoline sales and can be reasonably simulated with the assumption that mid-grade is 50/50 premium/regular.

A properly developed LP does not require a separate blending sub-module, it is embedded in the LP structure.

Gasoline Qualities

RVP—It is generally accepted that vapor pressure does not blend linearly. Some models use the exponent method to accommodate this fact. An example of this is below:

	VOL %	RVP	RVP^1.15
Normal Butane	3.1	52.0	94.1
Reformate	47.7	3.5	4.2
FCC Gasoline	31.4	7.1	9.5
C5/C6 Isomerate	15.2	13.8	20.5
Naphtha	2.6	7.1	9.5
Total	100.0		11.3
Final RVP			8.2

In this method, each component is raised to a power. Often, refiners will calibrate this power to match actual refinery blends. I have seen this power range from 1.1 to 1.2. This volumetric product of the RVP's raised to the power factor is calculated above to 11.3, and then this number is inverted back by raising 11.3 to the (1/1.15) factor, to calculate a final blend RVP of 8.2.

The power factor has more impact on high RVP components, such as butane. Another reasonable RVP blending method is to modify the RVP of some components such as butane to a "blending" RVP. This modification allows the blend to be calculated volumetrically. Below, the RVP of butane is entered as a blending RVP of 62, and the RVP of isomerate is increased by 0.5 psi, to calculate a final blend of 8.2, the same as the 1.15 power factor above.

	VOL %	**RVP**
Normal Butane	3.1	62.0
Reformate	47.7	3.5
FCC Gasoline	31.4	7.1
C5/C6 Isomerate	15.2	14.3
Naphtha	2.6	7.1
Total	100.0	8.2

In the end, either method is suitable, although I prefer the blend RVP method over the exponential non-linear method.

Octane—Octane in the model should include RON, MON, and (R+M)/2. It is generally recognized that octane does not blend linearly. To offset this, there are generally recognized standards, including octane bonus values and blending interaction coefficients.

The octane bonus methodology assigns a bonus or penalty to the octane number of the component depending on the final gasoline pool. For example, light straight run (LSR) gasoline blended into premium gasoline could be assigned a +1.0 for RON and a -0.2 for MON, which represent a bonus and penalty assigned to the base RON and MON. The bonus and penalty can be different if the LSR blends to regular gasoline.

The blending interaction method assigns interaction coefficients for the components in a matrix form.

	Comp 1	Comp 2	Comp 3
Comp 1	C_{11}	C_{12}	C_{13}
Comp 2	C_{21}	C_{22}	C_{23}
Comp 3	C_{31}	C_{32}	C_{33}

This method requires extensive data since an interaction coefficient is required for all the blending components. Coefficients for C11, C22, and C33 are 1.0, because there is no interaction between a component blending to itself.

Some refinery-specific modeling efforts tie-in sophisticated blending modules simulators that run concurrently to the LP. This is a good approach for a specific refinery, but the data are typically calibrated to the refinery, and in my opinion is too sophisticated for generalized modeling.

A third approach—and the most simplistic—is to input the base octane as an average blend value octane, and perform the octane calculation volumetrically. This is usually a reasonable approach for generic modeling efforts.

In my experience, both from the development of generalized and refinery-specific models and working with other client refinery-specific models, I have seen all three methods employed. I prefer to input the average blending octanes and used bonus/penalty shifts for generalized modeling. I do not recommend the interaction method or on-line simulation for a generalized technology database because of the excessive data requirements, and the potential for non-convergence.

Distillation—In the US, gasoline distillation is generally specified with constraints on T10, T50, and T90 (Temperature at which 10%, 50%, and 90% of the sample is distilled). Distillation blending is a quality that is generally understood to not blend linearly. A more suitable method blends a percent off—such as E200 (volume distilled at 200˚F), or E300 (volume distilled at 300˚F).

The percent off temperatures are then volumetrically blended. Using E200 and E300, one can then incorporate a relationship for T50 as a function of E200, and T90 as a function of E300. Another relationship can be developed for T10 as a function of RVP.

Additional specifications that should be incorporated into the LP for gasoline blending include driveability index and vapor liquid equilibrium. These specifications can be calculated using E200, E300, and RVP relationships. These qualities are important because vapor liquid equilibrium is often more constraining in the winter versus the RVP constraint. Driveability can also be a limiting constraint, which prevents a gasoline from being to “heavy.”

Other Specifications

The following qualities are generally accepted to blend volumetrically:

- Olefins, volume based
- Benzene, volume based
- Aromatics, volume based
- Density, volume based
- Sulfur, weight based

Another quality that we also carry in the LP is BTU/Bbl. This quality has become increasingly important with ethanol blending since the energy content of ethanol is substantially lower than traditional hydrocarbon. Tracking BTUs of the gasoline pool allow the LP user to specify energy density of the gasoline pool.

Complex Model Equations should also be carried in the LP. Historically, these equations have been programmed into GRTMPS and PIMS. Because of the high degree of non-linearity in these equations, there can often be convergence issues when these equations are invoked. These equations were developed for the Simple Model, Complex Model Phase 1, and Complex Model Phase 2, for both summer and winter season, for both regions B (South) and C (North), and for conventional and reformulated gasoline.

An alternative to incorporating these equations is to simply specify a recipe that will pass the appropriate model. Using recipes will likely improve convergence issues; however, the blend will most certainly be sub-optimal. From analysis of marginal values and specifications of the blend in an LP, one could develop a new, more optimal recipe for a second LP. This technique can be effective but is not recommended as part of an LP that will be used to analyze regulatory impacts.

Going into the future, with MSAT2, the need to rigorously model toxics and NOx will no longer be required. Reformulated gasoline will still require a calculation on VOC reduction in the summer period. An alternative to calculating VOC reduction would be to specify an appropriate vapor pressure, such as 6.8 – 7.0 psi, that will ensure VOC compliance. Of course, once a solution is obtained, the results must be checked off-line to ensure VOC compliance.

The Complex Model equations add excessive structure and convergence issues. In my own models, I am contemplating removing these equations once MSAT2 is in place and rigorous toxics and NOx calculations are eliminated.

Distillate Blending

Distillate specifications also must be considered in the LP. Many qualities do no blend linearly and should be calculated using an index method. This includes viscosity, cold flow properties (pour point, cloud point, and freeze point), and flash point.

There are many different index methodologies to consider, and I have not analyzed all the methods to opine on which is "best." In fact, I am not sure there is a "best" method for generalized modeling, although a specific refinery might incorporate a specific index method that "best" reflects their actual blending operations. For reference, our generalized LP model uses Maxwell for viscosity, Hu-Burns for Pour, and Wickey-Chattendon for Flash.

Regarding cold flow properties of Pour, Cloud, and Freeze, one could certainly incorporate all of these specifications into the LP. There is a reasonable relationship between all of these qualities, as shown below:

Initial Cut Point	deg F	400.	430.	525.	
End Cut Point	deg F	430.	525.	650.	
Pour Point	deg F	-76.69	-40.10	12.34	
Cloud Point	deg F	-71.84	-36.38	17.16	
Freeze Point	deg F	-66.98	-32.66	21.98	
					AVG
Cloud-Pour	deg F	4.85	3.72	4.82	4.46
Freeze - Pour	deg F	9.70	7.44	9.63	8.92

One has to weigh the benefits of carrying three cold flow properties throughout the model versus carrying one or two and using a generic relationship to relate the other(s). Our generalized modeling typically carries pour point throughout the model, and uses a relationship on the freeze or cloud property to express them as a pour point specification. For single client work, I have often added a second cold flow property, but rarely all three.

Cetane Index methods have changed with the new sulfur standards:

- Cetane Index D976
- Cetane Index D4737-A (High Sulfur Fuels)
- Cetane Index D4737-B (Low Sulfur Fuels)

Since most of US diesel production is low sulfur, it would be practical to model using the D4737-B method.

Industry Evolution

The existing capacity expansion routine in PMM may possibly be its largest flaw. The fixed 3-year time window is often responsible for counter-intuitive results, for which adaptations and workarounds must be found. It is also likely that insufficient resolution is used to evaluate build decisions (e.g., the cost to add an incremental coker is the same as adding a coker at a refinery that has never had one).

Refiners have been burned with low margins for overbuilding for decades. Given the risk adverse nature of many of the players, should non-economic interests be modeled?

An LP will tend to have the dominant technology take over the marketplace; however, in practice it may take several years for one technology to dominate the marketplace. In the renewable fuels arena there will also likely be technological breakthroughs that shift the playing field. How (or should) these be modeled? This aspect of the LFMM will be of importance to modeling the market growth of alternative fuels (BTL, GTL, CTL, and cellulosic ethanol). The new model could look to improve upon PMM's current technology penetration/growth modeling, which is based almost solely on the Mansfield-Blackman model.

Realistic refining capacity expansion predictions and optimization using LP models is difficult. There are a number of reasons why this occurs:

- The LP is driven by optimal economics, and there are too many assumed expansion factors to let the decision be made exclusively by the LP. Guidance is suggested.
- How, when and where expansions happen in the US does not likely occur in an optimal fashion. For example, if an LP model were constructed with 5 regions, reflecting the 5 PADDS, the LP will optimize the result as if all 5 regions are coordinating to maximize the objective value. In reality, this is not how the individual refinery investment decisions are made.
- "The LP will change the world to make a penny"—that is the job of the LP. For example, I have seen dramatic regional investment swings just by a slight modification of the regional location factor, an assumed variable.

As previously stated, I believe it is highly problematic to allow the LP or any model to exclusively make the investment decision on expansions. For regional models we have found it generally useful to guide this decision based on our experience. Some of the techniques include but are not limited to:

- Keeping track of announced and planned capacity expansions. Assign probabilities to these expansions and base load the LP model with these expansions.
- Using historical capacities, one can make reasonable decisions on regional creep expansion patterns, as well as establish probabilities of where overall regional investment will occur.
- Apply qualitative analysis on the quantitative output. If, for example, the LP predicts PADD 1 investment the same as PADD 3, and PADD 1 has 20% of PADD 3's capacity, this answer would be questionable.

It is true that the LP will tend to dominate a single technology for a marketplace. When the industry was gearing up for 30 ppm gasoline, there were a number of technology choices still in their pilot stages, and the data behind these technologies were questionable for a period of time. On single-client work, we would in fact model the separate technologies when determining the best fit for a specific configuration. This is not necessarily true for regional modeling.

Regional modeling is not necessarily the best mechanism for predicting new technology choices because the answer is extremely premise laden and the assumptions can change quickly. Using the 30 ppm gasoline example, after the data assumptions were more fully vetted (octane loss versus desulfurization) the LP often chose a different technology.

To impress the point further, there are a variety of benzene reduction technologies. For a single regional configuration, it would be unlikely to see an LP choose a piece of "Technology X" and a piece of "Technology Y." Rather, you would expect the model to choose a single technology. Again, for regional analysis, the type of technology is not quite as important as long as the model reasonably depicts the overall average benzene reduction process, the associated octane loss, hydrogen uptake, *etc*. However, if the study question were specific to benzene reduction options, a number of generic configurations would likely be a more suitable modeling approach versus a regional model.

In the regional model, it seems that the goal is to allow an average technology to be implemented into the model. It does not really matter if it is Technology A or B—what

matters is the octane loss, desulfurization rates, hydrogen uptakes, yields, *etc*. are representative of the technology. After all, the regional model is an average representation, not a specific refinery.

Off-line generic configuration models are usually better mechanisms to analyze technology. Even this has shortcomings because technology choices are often a function of crude runs. This means a Cracking configuration running light sweet crude might choose "Technology X" versus the same Cracking configuration running a medium sour crude slate that would choose "Technology Y."

Every refinery is in a unique situation; therefore, it is unlikely that any regional or generic model can identify the "best" technology.

Market

Despite its name, the existing Petroleum "Market Model" does little more than compute the production cost and add largely fixed markups, based on historic data. If one of the priority goals of a new LFMM is to produce accurate product margins across cases, some additional element will likely be needed. Please discuss some of the following or other options that may be employed to achieve this: 1) Mini representation of end-use sector demand elasticity, 2) have markups a function of other model parameters (e.g. ratio of demand to production capacity), 3) Game theory analysis, and 4) Stochastic analysis.

Our approach to producing and forecasting market margins is initially done in an off-line, exogenous model. The following is intended to provide a snapshot of the methodology, not a detailed explanation.

- Forecast the benchmark crude price
- Forecast gasoline price and diesel price
- These two points define the 3-2-1 crack spread
- Forecast light heavy crude differential
- Many of the remaining refined product prices are related back to crude, gasoline, or diesel
- Qualitatively and Quantitatively (LP) check the results

Once the forecast is set, we run the prices in an LP model to analyze the impact in a Cracking configuration running the benchmark crude, a coking configuration, and often a regional configuration. The forecast begins at the USGC and uses this base to develop forecasts for other regions.

We have often incorporated probability and quantitative risk analysis into the forecast methodology to produce high and low price/margin environments. In our efforts, this is all done off-line in independent spreadsheet models, and the results are fed to the LP model.

World Oil Price and Production

Should World Oil Price continue to be an exogenous assumption?

To what level of detail should global upstream operations be detailed?
Should NYMEX Low Sulfur Light (LSL) continue to be the Benchmark crude? Due to declining US production and changes in the flow patterns to Cushing, recent history has seen significant distortion to the LSL price in comparison to its global peers (Brent, Bonny, LA Light Sweet).

We forecast World Oil Price as an exogenous assumption. To properly discuss the crude oil price environment, it is appropriate to understand the fundamentals of supply/demand and the role of OPEC in working to control crude supply and price.

Demand Elasticity

Economists define elasticity as the amount that demand changes in response to a change in price. In refined products—and therefore crude—there are two components of demand elasticity: short term and long term.

In a falling price market, there is very little immediate upside demand elasticity. It has been well documented that lower gasoline prices do not immediately encourage significantly more driving. In the same way, we are relatively unable to change demand rapidly when prices increase. Even in the oil shocks of the early 1980s and in the current environment, the number of miles driven did not change greatly.

The only way that demand can fall more rapidly in a high crude price environment is due to a recession, during which demand for petrochemicals (and thereby naphtha and LPG's) fall, people drive and fly less, and industrial demand, in general, slows. Historically, such demand changes have been relatively short-lived. The total demand curve has ratcheted down, then continued along the same trend line.

Price changes do cause long-term effects that are far more price elastic. Over the 1979 - 1985 period, world refined product demand decreased for three reasons:

- Improved efficiency in the United States, Western Europe, and Japan lowered energy usage per unit of GDP.
- Alternate energy sources including coal, nuclear, and natural gas provided all of the growth in energy use plus making up for the reduction in crude oil.
- High energy prices served to slow economic growth.

Conversely, a long period of relatively low oil prices has led to structural changes that increase demand. In the US, 1998 was the first year since 1983 that the new cars purchased during the year were less efficient than the old cars leaving the fleet. The clear reason is the growth of light trucks and sport utility vehicles. While this will likely slow in the face of high oil prices, it takes many years to change the overall efficiency of the fleet.

In summary, demand reacts to price, but slowly.

Supply Elasticity

In simple terms, there are three increments of oil supply:

- Capacity, including spare production capacity that is already available;
- Production from older existing fields that can be increased (or the decline rate slowed) via enhanced recovery methods; and
- New production from new fields.

Existing capacity is the only short-term increment of supply. OPEC and non-OPEC producers look at spare capacity in different ways. Given the economics of production, non-OPEC producers rarely cut back production, preferring current cash flow to future cash flow. OPEC, on the other hand, tries to maintain price by constraining supply. It should be noted that this system is the reverse of classic economics, in which the last increment of supply is supposed to be the most costly.

Production from older fields is price elastic. Clearly, the economics of enhanced recovery improve with higher oil price, leading companies to invest in ways to get more out of existing fields. This increment can be tapped relatively quickly once producers

decide that a crude price increase is not just a short-term spike. Increased oil field activity normally lags a price increase by six months to one year.

Finding and developing new fields also pick up as prices increase, but this increment of supply is long term—it takes several years after a crude price increase before the increment develops.

Our current crude price situation is a result of years of relatively low-priced crude, which supported increased demand while lowering the amount of new supply. Currently, it is clear that there is no spare capacity in the non-OPEC world and very little (or perhaps none) in OPEC. Crude prices should eventually decline in real terms, but unless there is a significant economic downturn, the decline will be slow.

There are those who believe that the world has run out of oil, or will very soon. Jacobs Consultancy has no expertise in geology and therefore cannot provide a cogent opinion. Suffice it to say that both the demand outlook and the outlook for refining would be very different if oil production began to decline due to supply limits.

With that background on supply/demand elasticity, it is difficult for me to envision a long-term model that has world oil price predictions embedded in the model, and not exogenous to the model. This appears to be the classic case of when the model assumptions would overwhelm any single, predictive outcome from the model. I would recommend using off-line models for world oil price. From these off-line models, coupled with risk analysis and other quantitative methods, it seems reasonable to establish scenario cases of varying crude price. These scenarios would be transferred to the LP for multiple runs.

Global upstream operations should continue to be modeled to accurately reflect long-term trends as they relate to long-term supply. The modeling and forecasting of the potential of existing fields and reservoirs, and the development of new fields has value. Currently, for example, upstream modeling might be used to express opinions of declining production and imports of heavy Mexican Maya crudes, and the increase in Canadian synthetics and bitumen-derived crude. Along with this, one should asses the long-term impact of bitumen-based crudes from Canada to arrive on the USGC.

There are legitimate reasons for NYMEX (WTI) and legitimate reasons for using a global peer, such as Brent. On the NYMEX argument, the sheer volume of trades make it a logical choice. Additionally, many other crudes are formula priced off WTI. There have been distortions between WTI and the “rest of the world,” which suggests that a highly transparent waterborne crude (Brent) should be used as the basis. The WTI distortions are typically short term and would not typically impact long-term forecasts. For argument’s sake, the long-term forecast will have a relationship as WTI as a function of

Brent, or Brent as a function to WTI, and short-term distortions in WTI will not be forecasted.

Transparency and Stability

Ideally, the new model should allow for reasonably easy insight into the connection between model inputs and outputs. The model should also be reasonably stable when slight adjustments to inputs are model parameters—e.g., avoid the LP "knife-edge" effect.

As previously stated, the model will "turn the world upside down to make a penny." Part of the stability issue revolves around the complication within the code of the LP. If, for example, the LP has 50 vectors to choose for the FCC operation, this is likely to be less stable than a representation using fewer modes. One would have to ask if the 50 vectors are on a consistent basis, or whether the vectors have accumulated over time and without a consistent basis.

For the most part, model stability is directly related to the LP code. Many stability issues can be resolved with an appropriate technology database and generally accepted and proven modeling principals. Beyond that, it often takes experts in matrix generation and solving optimization to get to the root of stability issues.

Often, the analysis on a large, complex regional model is difficult to uncover the insights that can be gained on a more simplified generic configuration model. This fact is another reason for the LP toolkit approach: use the generic configurations to "get a handle" on the issues, and then use the knowledge gained from these cases to develop and analyze the issues on a regional level.

Size/Complexity of the Model Relative to the Desired Output

How much detail is needed to produce reliable results? If the proposed model is an LP, does it need a refinery-like complexity in terms of the number of streams and interactions? Or can many of these refinery processes and interactions be "boiled down," even if an LP structure is still desired?

The LP does not need to be quite as sophisticated as a detailed representative LP for a specific refinery. Specific refinery LPs often have an additional layer of code to capture the nuances associated with that particular operation.

A regional LP should capture all the major streams in a refinery configuration, and often has more stream interactions than a specific refinery LP. This is because of the nature of a regional LP, and the fact that it represents a number of refinery configurations means that it should have "extra" interactions that may not be present in a specific refinery LP. For example, a specific refinery might not be piped up to send LCO to the residual fuel oil blending, only to a ULSD HDT; however, the regional LP should include both options, because both would be reasonable refinery options among many refiners.

Too much oversimplification can dilute the answers to a point where analysis of the situation can be difficult. As an example, if a model produced a single average octane, average quality grade of gasoline (not separate premium and regular grades or separate CG and RFG), and an analysis was done on blending E15 versus E10, the analysis would fall short because ethanol blending will impact CG differently than RFG and regular differently than premium.

The LP should have an accurate representation of volume gain since the US is volume based. Additionally, the LP should have a reasonable prediction of weight and density to ensure mass balance.

Flexibility

LFMM must be capable of delivering reliable results from sensitivity cases as well as assessing a wide range of alternative policy cases.

- *Project refinery margins.*
- *Reliably project liquid fuel production capacities and production volumes.*
- *Analyze future policy cases with sufficient ease and reliability—these include both direct energy policies (e.g., specifications on fuel products, bio fuels mandates) and indirect policies (i.e., environment and transportation).*
- *Assess the impact of alternative fuels (mandates, tax incentives, subsidies) on the liquid fuels market.*
- *Project refinery expansions and retirements—the new model should be capable of accompanying near-term planned capacity additions while allowing for economic builds and retirements.*

The requirements of the LFMM point to a modeling system that has been recommended to the EIA:

- Robust, user-friendly, and transparent Technology Database
- Accurate and sufficient crude assays

- LP toolkit with regional and generic configuration options
- Sophisticated LP software platform such as GRTMPS

These recommendations are based on many years of modeling experience. They also represent the starting point of my professional modeling work that is used to analyze the same questions above for both specific single-client work and generalized industry analysis.

Other Topics

Low Carbon Fuels Standard—*What additional considerations are necessary to allow modeling of LCFS?*

Within the LP, CO_2 should be accurately accounted for in the following areas:

- Coke Burn in the FCC
- CO_2 production off the Hydrogen Plant
- CO_2 produced from fuel combustion, including refinery fuel gas, natural gas, and fuel oil if liquid residual fuel is burned

These issues draw the CO_2 material balance at the refinery gate for most situations. Extending the CO_2 material balance box beyond the refinery, it is also possible to allocate a CO_2 contribution associated with the purchase of electricity for refinery power.

Going beyond the refinery gate, it is possible to allocate CO_2 factors for all refinery purchases and all refinery product sales. The combination of CO_2 from purchases, sales, and the refinery-generated CO_2 would result in a Well-To-Wheels (Lifecycle) tracking system that the LP can accommodate.

There are inputs and outputs that must be rigorously thought through. For example, the CO_2 loading factor for crude oil is an input. This loading factor is a complex calculation that includes but is not limited to:

- Crude / bitumen properties
- Reservoir characteristics
    - *Depth*

  - *Water to oil ratio*
  - *Gas to oil ratio*
- Flaring and venting of produced gas
- Treatment of produced water and gas
- Production method and technology
  - Primary, Secondary, Tertiary

Every crude will have a unique CO_2 loading factor. Additionally, the refined products can contain a loading factor that accounts for the distribution of the product from the refinery to the retail outlet and the combustion of the product as it is burned. This accounting represents Tank-To-Wheels.

Not only can the LP be developed to accommodate these CO_2 tracking mechanisms, but the LP can be used to optimize "new" operations for various CO_2 tax scenarios.

***Gasoline/Distillate Ratio**—What steps should be taken to accurately model increased diesel supply?*

Downstream processing splitters and swing cut methodology should be incorporated to accurately model increased diesel supply. The swing cuts off the crude tower should include:

- Heavy Naphtha/Jet Swing—this allows more jet production, which in turn can swing into the distillate pool up to a flash specification limit.
- Jet/Diesel Swing—allows the optimization of higher volume of distillate.
- Hvy Diesel (HAGO)/LVGO Swing—allows the HAGO to go directly to distillate production versus sending to FCC. This swing will be typically limited by the cold flow specification of the diesel.

The next set of units represents downstream splitters that can allow the model to "reach" deeper into the barrel for higher distillate yields.

- FCC naphtha splitter, producing Light (C5/200˚F), Medium (200˚F/350˚F), and Heavy (350˚F/430˚F) FCC naphtha. The Heavy stream should be allowed to swing to gasoline or swing to FCC LCO for a wider diesel cut. The HCN cut could

potentially have a lower initial boiling point of 350˚F (wider cut), up to distillate flash limits.

- Coker Gasoil Splitter—This becomes a function of how deep the coker gasoil is cut off the coker unit. If the coker diesel off the unit is 690˚F EP, there is not much need for this splitter, because that is about the extreme before hitting the T90 or EP spec on diesel. If the model produces a 650˚F EP on coker diesel, the gasoil (650+) can go to this splitter and dredge up the 650˚F/690˚F into the coker distillate pool.
- FCC Slurry Splitter—This is similar to the Coker Gasoil Splitter above. If the standard LCO off the FCC is 430˚F-650˚F, the 650˚F+ slurry can go to the splitter to produce a 650˚F-680˚F stream that is dredged up to the FCC LCO pool.
- Additional modifications might include a set of low conversion FCC vectors.

A unique aspect of these splitters is the user can decide how much product to enter the splitter, which determines how "deep" the refiner can cut to maximize diesel. If the coker gasoil is always cut at 680˚F, this implies all refiners can achieve this, which is not the case, especially with sloppy fractionation.

***Declining Heavy Fuel Demand**—If high sulfur resid demand evaporates, how will refineries respond?*

This question will be analyzed using the recommended techniques presented. As stated previously, depending on the question, new code and sophistication might have to be incorporated. This could include potential updates of the existing "resid" destruction technologies, including resid FCC, resid Hydrocracking, cokers combined with gasifiers, or whatever "new" technology that may be relevant. Once a qualitative assessment of this has taken place, the quantitative impacts can be implemented into the models for analysis.

International liquids market representation and its interface with the US market

Undoubtedly, there is a relationship between the international and US market. How this is captured can take a variety of forms, including but not limited to:

- LP models of international markets tied to the US market LP model
- Offline analysis of international supply/demand patterns that would be input into the US market LP

- Multiple scenario analysis will likely be performed, regardless of the method chosen

Refinery-related cost issues, such as fuel use and energy efficiency.

Refinery variable operating costs (fuel, power, water, steam, catalyst and chemicals) are all embedded into the technology database. These utility consumptions should generally try to reflect average uptakes. To modify and analyze these cost issues, one can apply efficiency factors to simulate the extremes—high energy utilization and low energy utilization.

In the end, the LP (and the refinery) typically maximizes on the gross margin. The extent to which refinery operations change (throughput, crude selection, product output) usually is not impacted by the variable operating costs. These costs are small compared to the gross margin. Restated, the refinery material balance (gross margin) does not usually change due to variable operating consumptions. An increase in efficiency will increase the variable margin, and an efficiency decrease will decrease variable margin.

Final Notes

It has been a privilege to author this white paper for the EIA. It is my desire that these opinions provide sufficient background and substance for the LFMM workshop attendees.

Often, we undertake model development projects with haste and hurry. We do not always spend the time for good honest debate over the pros and cons, the wishes and expectations, or the overall needs of the model. The modeling debate should include the good, the bad, and the ugly. I applaud the DOE/EIA for taking this effort seriously, judiciously, and patiently—it will pay off in the end.