COMMENTS ON THE DEVELOPMENT OF

A LIQUID FUEL MARKET MODULE (LFMM)

FOR THE NATIONAL ENERGY MODELING SYSTEM

Prepared for

Office of Integrated Analysis and Forecasting
ENERGY INFORMATION ADMINISTRATION

By

David S. Hirshfeld

MathPro Inc.
P.O. Box 34404
West Bethesda, Maryland 20827-0404
301-951-9006

October 14, 2009

MathPro Inc.

INTRODUCTION

EIA has undertaken the development of a new element of the National Energy Modeling System (NEMS). This new element – the Liquid Fuel Market Module (LFMM) – is intended to replace and improve upon the existing Petroleum Market Module (PMM), which has been in use for more than thirty years. In this connection, EIA has asked some people from outside their organization to identify and discuss technical issues bearing on the requirements, capabilities, and design of the new system.

I am pleased to submit this paper in response. The views expressed here are my own, and not necessarily those of EIA or any other organization.

My comments are based on discussions with EIA staff, review of the recent EIA White Paper on LFMM development [EIA1] and other EIA documents (e.g., [EIA2]), review of PMM model documentation, and my own experience in refinery modeling. My comments are intended to be suggestions for EIA's consideration and not necessarily positive recommendations for the design of the LFMM.

The paper comprises seven sections. Section 1 offers a perspective on refinery modeling that informs the entire content of the paper. Section 2 addresses requirements for the modeling platform for the LFMM. Sections 3 and 4, respectively, comment on a few aspects of the LFMM's forecasting (AEOs and IEOs) and special studies applications that are relevant to LFMM model design. Section 5 briefly delineates possible attributes of refining sector models tailored to each set of applications and comments on a few modeling issues.. Section 6 offers brief comments on model calibration. Section 7 comments on the representation of advanced bio-fuel supply within the LFMM.

1. A PERSPECTIVE ON REFINERY MODELING

My comments reflect a particular perspective on the development, support, and use of refining industry and related process models. This perspective has evolved in the course of 40+ years of consulting work, mainly related to the refining industry, and it may have relevance to EIA's development of the LFMM. It can be summarized in three principles.

- A refinery model – indeed, any model – should be as simple as possible, but not simpler.[1]

  For any given study, one wants a model that captures the technical elements and relationships essential to analyzing and illuminating the problem at hand – and no more. Additional scope and detail won't contribute to the analysis and is likely to impede or obscure it.

[1] I paraphrase Einstein and Thomas of Ockham here. Neither was a refinery modeler, but they are impeccable sources nonetheless.

Every project is unique; the model scope and content that's right for one may not be for the next one. Some projects call for a highly detailed and comprehensive representation of refining operations (perhaps multi-regional); others may need a simpler or highly aggregated model. Hence, a group doing continuing analysis of the refining sector benefits significantly from access to multiple model templates and the capability to quickly adapt an existing template to the requirements of the study at hand – where "adapt" may mean adding new elements to the selected model template and/or removing some existing ones.

- It is better to be approximately right than precisely wrong.

One corollary of Parkinson's First Law[2] is that models and modeling platforms tend to grow with the resources available for operating them. Models tend to get bigger and more complicated over time, mainly out of a well-intentioned desire to achieve more precise representations of the "real world" and thereby obtain more precise solutions. However, precision is not the goal in forecasting or in analysis; accuracy is.[3]

Accuracy in a policy analysis context means providing new insight into how some part of the "real world" – the refining sector in this instance – would respond to policies, economic drivers, and other external forces over the *range of values* that these elements might reasonably take on in the future. The value of an analysis usually lies in the range of various parameters that it explores and the insights that these explorations produce, not in the level of detail or the number of significant figures in the numerical results the analysis produces – especially if these results are wide of the mark in the first instance.

- Analysis is done by analysts, not by models.[4]

Not infrequently, a client (stakeholder?) will ask what our model says about some issue. I always take exception to such questions. A model has no more to say about an issue than a slide rule or a desk calculator would.

An analyst working with a model develops understanding and insight into the problem at hand by making many – perhaps hundreds – of exploratory model runs. Ultimately, these probes reveal the critical elements of the study, the parameters whose values largely determine the results and findings of this particular analysis. Every analysis has its own critical elements. Capturing them may require gathering new technical data, modifying the model, and then running many more cases to develop the results of the analysis.

---

[2] During the '50's and '60's, C. Northcote Parkinson was a popular observer and satirist of large organizations. Parkinson's First Law is "Work expands to fill the available time; expenses rise to match revenues."

[3] *Precision* is "refinement in a measurement, calculation, or specification, e.g., as represented by the number of digits given."

Accuracy is "the degree to which the result of a measurement, calculation, or specification conforms to the correct value or standard." [NOAD]

[4] An alternative phrasing is, "Analysts are for thinking; models are for computing."

One cannot conduct analysis in this manner if the model and modeling platform at hand are cumbersome, slow, and opaque.

This general approach to analysis depends on what one might call “agility” – the capability to shape a model to the analysis at hand; the use of models that are as parsimonious and transparent as the application will allow; the execution of large numbers of model runs to identify the essential elements of each study; and the ability to organize results that span multiple scenarios. Achieving this kind of agility depends mainly on the capabilities of the modeling platform.

2. Some Thoughts on the Modeling Platform for LFMM

Regardless of the specific requirements, design philosophy, and analytical scope that EIA establishes for the LFMM, EIA should create a completely new modeling platform for the LFMM. Having such a modeling platform is more essential to achieving EIA’s overall objectives for the LFMM than any particular capabilities in the model(s) themselves.

EIA should not try to implement the LFMM by further modifying, revamping, or enhancing the current PMM software. That software has grown unduly costly and time-consuming to operate. It has reached the end of its useful life, with respect the wide range of applications and analytical issues that EIA seeks to address with the LFMM.

The LFMM modeling platform should employ current best practice for such systems, with the aim of speeding, simplifying, and reducing the cost of the on-going development, enhancement, maintenance, and operation of LFMM models. The LFMM modeling platform should be capable of routinely creating and processing numerous different model instances (e.g., cases representing different assumptions) in a given analysis. The platform also should facilitate the updating, documentation, query, and display of the model(s) and the modeling data.

None of these capabilities requires the development of new system development tools and methods. The LFMM modeling platform should exploit proven tools and methods, including commercially available modeling languages, software implementation languages (Java, C++), and solvers.

In broad terms, the modeling platform should have these attributes:

- The system should reside, in its entirety, on a central server.

  This attribute would permit (1) centralized maintenance and updating of system software and (2) establishment and maintenance of a common data store for all system applications (forecasting exercises, one-off studies, etc.).

- All system elements should be accessible through a highly-structured user interface based on Web facilities and tools (i.e., browser/HTML based).

The user interface should contain "Help" facilities supporting both input screens and results screens, with the Help including identification and explanation of linkages/logical connections to other screens and fields.

- The system should employ a "high end" relational database system, such as Microsoft SQL or ORACLE, for data management.

  These systems are fast and offer a full array of data manipulation, query, and validation; pre-defined and ad hoc report writing, and back-up/restart capabilities. The relational data model minimizes the need for repetitive entry of new or updated data elements.

- The data store should accommodate different classes of modeling data, such as
  - Semi-permanent techno-economic data (e.g., crude assays, refining process yields, refining cost elements, etc.) to be used in all model variants and updated only on a centralized basis ("facts"); and
  - Transient, study-specific data, assumptions, and exogenous inputs – such as crude supply curves, product demand volumes, policy representations, etc. ("choices").

- The models themselves should be expressed, if possible, in a symbolic modeling language, not in a procedural language such as FORTRAN, C, DATAFORM, or OMNI.

  A number of symbolic modeling languages are available as commercial software products. Most of them represent model variables and constraints in some variant of meta-algebraic notation. Algebraic notation was not designed with refinery models in mind, and therefore may not be suitable here. But other forms of symbolic model representation exist that are more compatible with refinery modeling. In general, symbolic modeling languages are likely to have considerable advantages over a suite of programs written in general-purpose procedural languages such as those in which PMM was implemented.

  Alternatively, if (as I understand) EIA already licenses the GRTMPS modeling system, then EIA should explore using GRTMPS for expressing and generating models.

- The system should include a standard crude oil assay library and a crude oil assay manager.

- The models should be solved with an advanced commercial solver with MIP, SOS2, and nonlinear capabilities.

  The system's optimization capabilities should include stochastic programming – implemented either through the commercial solver or through recursive procedures built with the relational database facilities.[5] Stochastic programming capability would be needed if and when EIA were to more rigorously address the uncertainty inherent in economic forecasting.

[5] Though not in wide use now, stochastic solution capabilities are sufficiently amenable to generalization that they can be incorporated in commercial solvers.

- The system should have a model analysis component that makes the models visible to the analyst for development, debugging, verification, and explanation.

  This component would include, for example, facilities for viewing selected sectors (variables, constraints, coefficients) of a model to verify that it has the intended structure and numerical content, tracing the interactions between specified classes of variables and constraints, identifying sources of infeasible solutions when they occur, etc.

- The system should have advanced case management capabilities, implemented by means of the relational database system and accessible through special features in the Web-based user interface.

  Advanced case management is essential for any study involving the creation and solution of numerous cases. The case management component would organize the various model instances (cases) in a given analysis into a hierarchical "case tree" by giving the user the capability to "case" as a distinct attribute of the data, in the relational sense. A hierarchical case tree structure permits (1) creation, rapid processing, and selective retention (or deletion) of new model instances, (2) automatic propagation of new inputs in a model instance to all "descendants" of that instance, and (3) the "side-by-side" viewing and reporting of solutions values from multiple model instances.

These are "state-of-the-art" attributes now; but, in my view, a modeling platform with these elements and capabilities is essential to the success of the LFMM endeavor. More than the LFMM model or models themselves, the modeling platform will determine the practical analytical capability of the LFMM and the resources required to operate it.

3. Forecasting (for AEOs and IEOs)

Like the PMM, the LFMM would have two primary and distinct areas of application:

- *Forecasting*: Preparation of the liquid fuels portion of the AEOs and IEOs

- *Special studies*: One-off analyses of the refining industry and of techno-economic developments and proposed policies, legislation, and regulations that would affect it.

This section offers some comments on the forecasting applications; Section 4 comments on the special studies applications.

The comments here bear on three aspects of the LFMM's forecasting applications that warrant consideration in the design of the corresponding refining sector model: the dominant role of the crude oil price in the formation of product prices; the linkage of the U.S. refining sector to foreign sources of refined product supplies; and the linkage of the U.S. refining sector to other energy supply sectors.

3.1 Crude Oil Prices, Refining Costs, and Product Prices

Within the NEMS forecasting framework, the refining sector model (PMM and, ultimately, LFMM) serves primarily to return estimates of delivered prices, by year and by region, for refined products and other transportation fuels produced by U.S. refineries to meet specified demand volumes.[6,7] In the iterative solution procedure for NEMS, these prices are key inputs to the various regional and sectoral demand models, which estimate volumes of refined products and other fuels demanded at the prices returned by the refining sector model.

For each refined product and region, the end-use price conveyed to the sectoral demand model equals the sum of:

- Average delivered price of crude oil
- Refining margin, by product, computed by PMM/LFMM as the sum of
  - Refinery energy use
  - Other variable refining costs (catalysts and chemicals, etc.)
  - Fixed costs
  - Capital charges and return on refinery investment
- Transportation cost from refinery to terminal
- Distribution costs and mark-ups from terminal to end-use point
- Federal and state taxes

Of these, the refining sector model generates only the refining margin.[8] The average delivered price of crude oil is largely determined by the world oil price, an exogenous input. The distribution system costs and taxes are also exogenous inputs.

Two aspects of the forecasting applications appear to warrant special consideration in the design of the corresponding refining sector model. One is the dominant role of the crude oil price in the formation of product prices; the other is the linkage of the U.S. refining sector to foreign sources of refined product supplies.

The (exogenous) crude oil price is the largest component of end-use liquid fuel prices. It accounts for more than 50% of these prices, even before factoring in the cost of refinery energy use, which is tied to the crude oil price.[9] Crude oil prices are essentially impossible to forecast with accuracy even in the medium term (say, 3 and 5 years), let alone the long term (10, 15, and 20 years). EIA's crude price forecasts usually miss the mark, and do by substantial margins.

---

6 PMM also produces estimates of refinery energy consumption, as presumably LFMM will as well.

7 The demands are net of product imports, estimated by NEMS on the basis of supply functions generated by WEPS+.

8 More precisely, the refining sector model generates marginal costs of production (shadow prices), with which one calculate a corresponding refining margin.

9 In 2007, the crude oil price accounted for about 55% of the weighted average end-use price of refined products, as reported in *AEO2009*. The higher the crude oil price, the higher its share of end-use fuel prices.

EIA's own retrospective analysis [EIA3] indicates that since 1995, the AEO 3 year and 5 year forecasts of crude oil prices have missed the mark by averages of 30% and 38%, respectively.[10]

Moreover, real crude oil prices – and hence real refined product prices – fluctuate over time, whereas all AEO forecasts indicate crude prices increasing monotonically over time. For this reason alone, neither the refining sector model nor the rest of NEMS can fully capture the effects of crude oil price on capital investment in either the refining sector or the end-use sectors.

The components of end-use fuel prices estimated exclusively by the refining sector model – variable refining costs (including refinery energy cost, which depends in large part on crude oil and natural gas prices) and capital charges – constitute ≈ 20% of the end-use liquid fuel prices that NEMS conveys to the sectoral demand models. So, mis-estimation by the refining sector model of variable refining costs (ex refinery energy) and capital charges has only a small effect on estimated end-use fuel prices and (because demand elasticity is low) an even smaller effect on the liquid fuel demands estimated by the sectoral demand models.

3.2 Foreign Sources of Supply

Few options exist in the end-use sectors for substantial volumes of inter-fuel substitution between liquid fuels and other energy sources. At the same time, the U.S. imports and exports substantial volumes of refined products (gasoline, gasoline blendstocks, jet fuel, and diesel fuel), and some segment of the imports may be the marginal supplies. Consequently, the U.S. refining sector has stronger interactions with foreign sources of crude and product supply than it does with the domestic sources of non-petroleum energy supplies.

One can view the sources of product imports as constituting three groups:

- *Short-haul export refineries*: a handful of large, near-by refineries that produce large volumes of refined products on-purpose to supply U.S. markets and are effectively part of the U.S. refining sector (i.e., refineries in Maritime Canada, the Virgin Islands, and the Caribbean)

- *European export refineries*: refineries, primarily in Western Europe, that produce surplus gasoline volumes in the process of meeting local distillate fuel demand and up-grade those volumes to meet U.S. specifications, on a continuing basis[11]

- *Opportunity suppliers*: more remote refineries that supply occasional cargos of refined products when market conditions are suitable

[10] This is an observation, not a criticism of the forecasting methodology.

[11] By and large, the product volumes exported to the U.S. are a residual of the refining operation; only the up-grading to meet U.S. standards is done on-purpose. Consequently, estimating a supply function for these volumes is particularly difficult.

Exhibit 1 shows the average daily volumes of U.S. imports of gasoline, jet fuel, distillate fuels, other refined product, and refinery inputs contributed by each of these groups of refineries, from 2005 through 2008. Together, the Atlantic Basin exporters – short-haul and European – accounted for 85%–90% of gasoline imports, 45%–65% of jet fuel/kerosene imports, 85%–90% of distillate imports, and 65%–70% of all petroleum imports during this period. The balance of imports came from opportunity suppliers.[12]

Exhibit 1: Imports of Refined Products, by Region of Origin and Year (K Bbl/day)

Product Type	2005	2006	2007	2008
Finished Refined Products	**2,243**	**2,134**	**2,139**	**1,821**
Finished Gasoline & Gasoline Blendstocks	**1,113**	**1,144**	**1,165**	**1,090**
Canada, Virgin Islands, & Venezuela	376	340	363	334
Europe	555	658	629	604
All Other	181	146	174	152
Jet Fuel & Kerosene	**197**	**190**	**220**	**105**
Canada, Virgin Islands, & Venezuela	79	85	96	69
Europe	8	2	0	1
All Other	111	103	124	34
Distillate	**329**	**365**	**304**	**213**
Canada, Virgin Islands, & Venezuela	252	258	222	172
Europe	34	58	33	19
All Other	43	48	48	22
Residual Oil and Asphalt	**572**	**400**	**411**	**374**
Canada, Virgin Islands, & Venezuela	177	154	140	130
Europe	105	56	120	105
All Other	291	190	151	138
Specialty Products[1]	**32**	**35**	**38**	**39**
Canada, Virgin Islands, & Venezuela	16	14	16	16
Europe	2	2	4	3
All Other	13	18	18	21
Refinery and Petrochemical Feedstocks	**892**	**995**	**961**	**979**
Unfinished Oils	**582**	**689**	**717**	**763**
Canada, Virgin Islands, & Venezuela	66	82	86	80
Europe	222	275	295	388
All Other	295	332	336	295
Petrochemical Feedstocks[2]	**310**	**306**	**244**	**216**
Canada, Virgin Islands, & Venezuela	14	20	15	13
Europe	62	46	29	20
All Other	234	240	200	183
Total Refined Products	**3,135**	**3,129**	**3,100**	**2,800**
Canada, Virgin Islands, & Venezuela	980	953	938	815
Europe	988	1,098	1,110	1,139
All Other	1,167	1,078	1,051	845

Note: This table deals with refined products; hence it excludes the following petroleum products: liquified petroleum gases, pentanes plus, oxygenates, and petroleum coke.

1 Includes special naphthas, aviation gas, aviation gas blendstocks, lubes, waxes, and miscellaneous products.

2 Includes naphtha and other oils used as petrochemical feedstocks.

Source: Derived from Table 24, *Petroleum Supply Annual, Vol. 1, 2005-2008*, DOE/EIA.

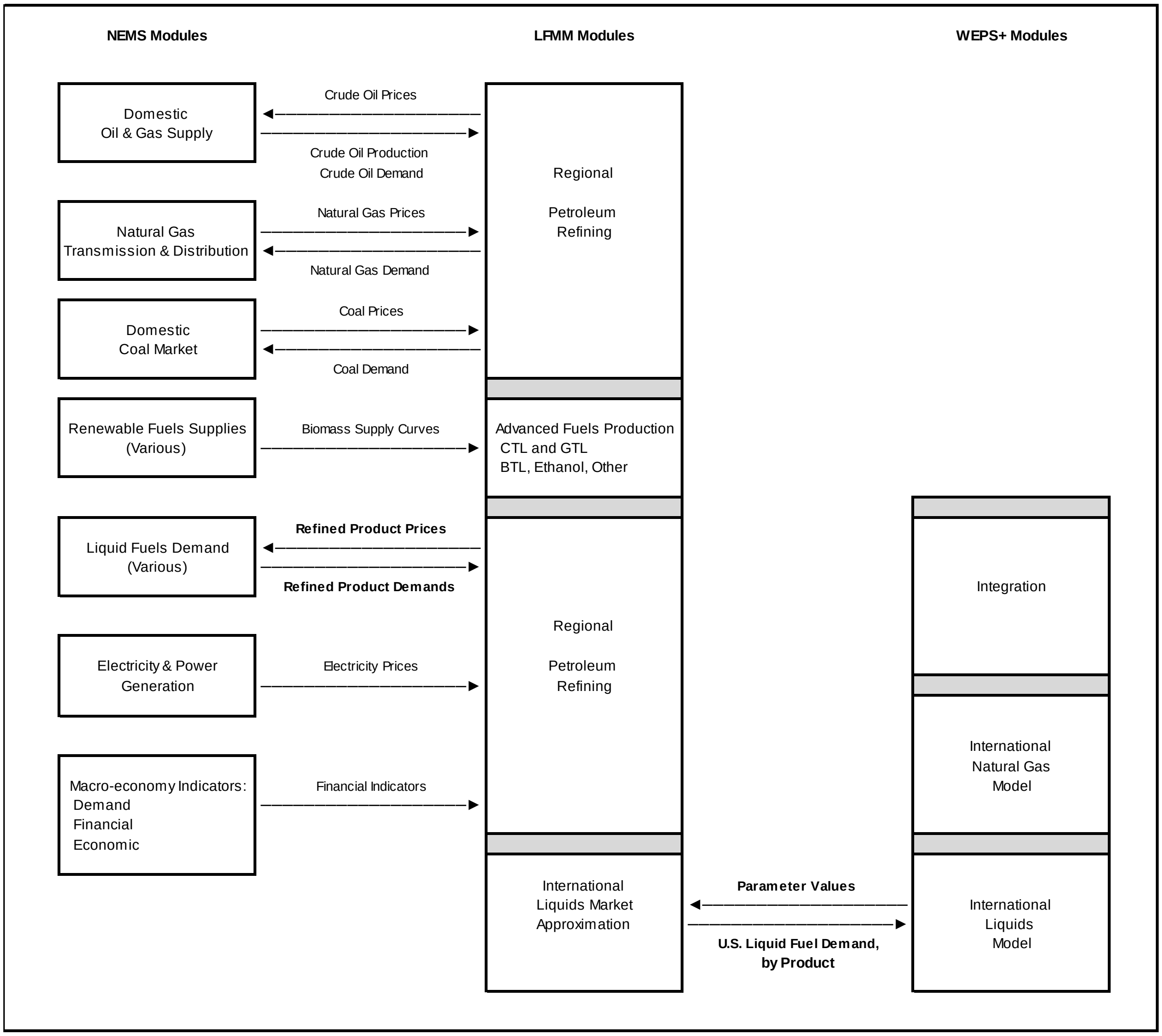
[12] The market share held by opportunity suppliers could well decline in the future, as the result of capital projects now in progress in U.S. refineries that will expand domestic gasoline-making capacity and the continuing increase in the volume share of ethanol in the U.S. gasoline pool, mandated by EISA.

3.3 Interactions of the Refining Sector with Other Sectors in NEMS

As **Exhibit 2** indicates, NEMS includes linkages not only between the refining sector and the demand sectors but between the refining sector and certain other energy supply sectors. NEMS includes linkages to other supply sectors because the refining sector, by virtue of its size, accounts for a significant segment (≈ 3%) of total U.S. energy consumption. In addition, the refining sector accounts for a comparable share of total U.S. GHG emissions.

These aspects of refining sector performance are not primary outputs of the refining sector model in NEMS in the sense that refined product prices are; but they are significant in their own rate and are of continuing interest to both analysts and policy makers.

Exhibit 2: Linkages Between LFMM, NEMS, and WEPS+



Source: "*A White Paper on the Development of a Liquid Fuel Market Module (LFMM) for the National Energy Modeling System*";
Energy Information inistration; May 10, 2009

3.4 Some Prospective Attributes of a Refining Sector Model for Forecasting Use

The above observations suggest that, for LFMM's forecasting applications:

- A relatively simple and compact refining sector model would be adequate for preparation of AEOs and IEOs, given that (1) crude oil price, not refining cost, is the primary component of refined product prices; (2) crude oil price is subject to large forecasting errors; and (3) the portion of end-use product prices computed by the refining sector model is relatively small.

- The refining sector model should encompass not only U.S. refineries but also the short-haul refineries and European export refineries that are primary continuing sources of U.S. imports of the key refined products: gasoline and distillate fuels.

- The refining sector model should produce estimates of (1) net refinery consumption of purchased fuel and electricity, (2) refinery emissions of CO_2 equivalents, and (3) the energy density (M BTU/Bbl) of each refined product, to support other elements of NEMS.

4. Special Studies

EIA documents [EIA1, EIA2] indicate that EIA often is called upon to conduct special studies on legislative and policy issues that bear on refining and that call for thorough techno-economic analysis of the U.S. refining sector. These analyses may include:

- Estimating investment requirements and refining costs for complying with new environmental regulations affecting liquid fuel properties (e.g., on-road and off-road ULSD, MSAT2, DI standards, low-sulfur marine fuel, possible future low-carbon fuel standards, renewable fuels mandates, etc.);

- Assessing the refining sector's overall response to significant changes in U.S. requirements for refined products, such as
  - Decreased gasoline/distillate ratio in response to dieselization of the vehicle fleet
  - Demand for HCCI fuel, ULSG, high octane gasoline, high cetane diesel fuel, or other new fuel types needed for prospective new engine technologies
  - Decline or elimination of demand for conventional residual fuel
  - Further restriction or elimination of local "boutique fuels" in favor of national fuel standards

- Analyzing the technical and economic effects on the refining sector of introducing new refining technology (e.g., "green diesel" processes, resid or coke gasification, etc.) or new types of crude oil (e.g., Canadian SCO, dilbit, and synbit)

- Estimating refinery energy use and CO_2 emissions associated with processing specific crude oils, in connection with lifecycle ("well-to-wheels") analyses

- Identifying individual refineries, refinery types, or refining centers that would be advantaged or disadvantaged by a new regulatory program

Such analyses call for modeling refining operations and economics in greater detail than that needed to generate end-use refined product prices adequate for the NEMS iterative solution procedure. However, because a special study of a proposed policy usually is not likely to address a specific U.S. refinery, such studies do not call for modeling refining operations at the level of detail and analytical rigor that U.S. refiners employ in their own refinery operations planning, investment planning, and crude evaluation applications.

Conducting one-off studies, such as those delineated above, in a timely and economic manner places special requirements on the modeling platform and the refinery model. Three such requirements are discussed briefly below: the ability to create and analyze many cases, the ability to analyze refining operations at different levels of aggregation, and the representation of different kinds of investments in additional refining capacity.

4.1 Analyzing Multiple Scenarios in a Study

As suggested in Section 1, studies of refining-related issues usually involve numerous critical elements; that is, technical or economic parameters (e.g., crude or product prices, performance of certain processes, product slate, etc.) whose values – often assumed – strongly influence or determine the results and findings of the analysis. Every analysis has its own critical elements, which one usually discovers in the course of exploratory model runs.

Our experience, also indicated in Section 1, is that an analysis usually covers many cases (each case defined by a unique combination of parameter values). Some of the cases are baseline ("business as usual") cases, which do not contain the legislative or policy initiative of interest; most are policy cases, which do contain the initiative of interest. The results of the analysis usually are determined by the differences between solutions returned by the model in the policy cases and solutions returned in the corresponding baseline cases. (Such analyses are sometimes called "differential analyses.")

Clearly, conducting analysis in this manner is feasible only with a modeling platform that is supple, efficient, and designed for multi-case processing and a refining sector model that can be set up quickly and solved readily. The model's representation of refining operations should be comprehensive but need not be particularly complex, because the results of most analyses are determined by differences between cases, not by solutions returned for any individual cases.

4.2 Analyzing Refining Operations at Different Levels of Aggregation

Special studies may be conducted at different levels of aggregation of the refining sector, depending on the nature of the proposed policy being analyzed, the objectives of the analysis, and the time and resources available for the analysis. For example, analyzing energy use and CO_2 emissions in the refining sector could be done at the national level; estimating the refining cost of proposed new environmental standards might best be done at the PADD level; analyzing refinery investment requirements might call for a more disaggregated refining representation.

Hence, the modeling platform should be capable of supporting different variants of the same underlying refining model:

- An aggregate national refining model, representing all U.S. refining capacity as though it were a single refinery;

- Regional refining models, each representing all refining capacity in a given PADD as though it were a single refinery; and

- A single-refinery model, configurable to represent either
  - a *notional* refinery; that is, an analytical artifact having capital stock and performance characteristics typical of a group of similar refineries (e.g., coking refineries in PADD 3, California refineries, etc.); or
  - an actual refinery (using operating data provided by the refiner).

All such variants can, and should, be created from the same symbolic model statement, or model template, by specifying the appropriate refinery process capacity profile, crude oil and other input volumes, and refined product volumes.

4.3 Analyzing Refinery Investment Requirements

Frequently, one of the main effects of a new policy or regulation affecting the refining industry is to call out investment in new refining process capacity. Accordingly, a stated priority for the LFMM is that it be capable of analyzing refining sector investments in process capacity [EIA1, EIA2].

Additions to refining capacity can occur in various ways, each with its own investment economics:

- Installation of a new process unit ("grass roots" economics)
- Expansion of an existing unit
- Retro-fitting an existing unit to a new service
- Capacity creep (debottlenecking economics)

(This list is in *decreasing* order of investment cost per unit of capacity.)

For all of these investment routes, investment cost per unit of capacity (K$/Bbl/day) varies widely by process (e.g., conversion processes require higher investment than upgrading processes), and the relationship between on-site (ISBL) and off-site (OSBL) investment varies by process as well.

All of these elements of refinery investment can be captured in a notional (single-refinery) model, but not with an aggregate regional model.

4.4 Some Prospective Attributes of a Refining Sector Model for Special Studies

These observations suggest that, for LFMM's special studies applications:

- A purpose-built refining sector model template is required that is capable of representing refining operations (notional or actual) in moderate detail – greater than that required for forecasting applications; less than that used by refiners in their own operations planning.
- The modeling platform must provide the capability to quickly and efficiently create, process, interpret, and report on numerous modeling cases.
- The modeling platform and the refinery model must be able to represent refining operations at various levels of aggregation and to capture a range of refinery investment routes.

5. A PROPOSED APPROACH TO REFINERY MODELING FOR DIFFERENT APPLICATIONS

As the comments in Sections 3 and 4 indicate, the two areas of LFMM application – forecasting and special studies – impose significantly different requirements for the modeling of refining sector operations and economics. Using one refining model, even a new one, for both sets of applications is likely to prove an inefficient use of resources and less than satisfactory in terms of analysis capability for both forecasting and special studies.

5.1 Different Applications; Different Models

Accordingly, EIA should consider developing two distinct LFMM models of the refining sector: one for the standard forecasting applications, the other for special studies. The forecasting model would be a single, multi-regional formulation. The special studies "model" in fact would be a model template, which would be used to create multiple variants at different levels of aggregation, as discussed in Section 4. The two would be at different levels of technical detail, tailored to the particular set of applications. The two models would be created, maintained, and operated using the same techno-economic data on refining operations and the same advanced modeling platform, provided the modeling database and platform were designed with this purpose in mind.

For purposes of this discussion, I assume that both models would be process-oriented optimization models (though other approaches might be considered for the forecasting model).

Both would be conventional refinery LP models, with explicit representations of standard refining processes and of product blending to industry and regulatory standards. The models would have a common symbolic model statement and be maintained and operated with the same modeling platform.

In addition, EIA may wish to consider developing an additional model: a non-optimizing, spreadsheet-based model that represents each U.S. refinery in highly simplified fashion. A number of firms have found such models useful in studies aimed at estimating costs incurred by individual refineries in meeting a proposed new standard – for example, to identify refineries that would be advantaged or disadvantaged by a new industry or regulatory standard bearing on transportation fuel (e.g., ULSD, MSAT2).

Exhibits 3 and 4 show possible sets of characteristics for two refining sector LP models, for use in forecasting and special studies, respectively. The exhibits are intended to serve only as a basis for discussion; they are not intended to be prescriptive in any sense.

5.2 Refining Regions in the Forecasting Model

Exhibit 3 indicates seven refining regions: five domestic regions (PADD 1, PADD 2, PADD 3, PADDs 4 and 5 (ex CA), and California) and two exporting regions (short-haul exporters, European exporters).

PADD 4 is combined with PADD 5 (ex CA) because PADD 4 is much smaller (in terms of refining capacity) than the other PADDs. California is represented separately because (1) it has its own, more stringent, standards on gasoline and diesel fuel, (2) the California market has only limited interaction with the other regional markets in the U.S., and (3) California refineries have some unique characteristics that differentiate them from other PADD 5 refineries.

The two export regions would provide the linkage in NEMS between the U.S. refining sector and refining sectors elsewhere in the world (a linkage that influences estimated prices of refined products.) As indicated in Section 3, the two indicated export regions account for roughly 2/3 of all U.S. imports of refined products and about 90% of gasoline and diesel imports – and these shares seem likely to increase in the future.

5.3 One Aggregate Refinery Model per Region in the Forecasting Model

At present, the PMM comprises two aggregate refining representations in each PADD – one denoting the PADD's complex refineries, the other the PADD's simple refineries. The complex refinery aggregate handles most of the crude and product volumes; the simple refinery aggregate serves to estimate marginal refining costs. My understanding is that the rationale for this dual-refinery approach is that simple refineries always incur higher refining costs than complex refineries and therefore must be the marginal sources of supply.

Exhibit 3: Possible Characteristics of a Refining Sector Model for Long-Term Forecasting (AEOs and IEOs)

Characteristic	Description	Comments
Geographic scope	Seven (7) refining regions	
	-- Five (5) domestic	
	PADD 1, PADD 2, PADD 3	
	PADDs 4 and 5 (ex CA), California	CA is large, unique, isolated
	--Two (2) foreign	These regions account for 80%-90% of gasoline and distillate imports
	Short-haul export refineries (Eastern Canda, Caribbean)	
	European export refineries	
Level of aggregation	One aggregate model per region, covering *all* refineries, both simple and complex	
Seasonality	Annual average	
Representation of. . .		
Refining operations	Process-by-process (neither extreme point nor base/delta formulation)	
	Limited to major processes	E.g., coking: delayed coking only, no flexicoking
	Swing cuts included in crude running representation	To accommodate possible changes in G/D ratio
	Parsimonious sets of process operating modes and intermediate streams	
	Operating modes mass-balanced	Mass balancing via stream yields and densities
Investments in new capacity	Process-by-process	
	Standard regional average investment values ($K/Bbl/day) for each proces	
Stream and blendstock properties	Standard stream properties; no recursive pooling	
	Standard energy content values (MM BTU/Bbl) for each blendstock	To meet product demands specified in energy tems
Refinery energy consumption	Process-by-process	
	Nat gas, still gas, FCC coke; calibrated to EIA-reported energy use	
Refinery emissions of CO_2e	Based on refinery energy consumption, by energy source	
Crude oil slate	Five (5) standard NEMS crude types	
		Blending Representation
Product slate and blending	CG, RFG, CaRFG (one grade per gasoline type)	Specification blending: octane, RVP, sul, bnz, oxygen
	E85	Recipe blending
	Kero jet	Specification blending: smoke pt., flash pt., sulfur
	ULSD, No. 2 heating oil	Specification blending: sulfur, cetane, cloud pt.
	Low sulfur marine fuel	Specification blending: sulfur
	LPG, Av gas, petchem feedstocks	Recipe blending
	Lubes & waxes, asphalt, pet coke	Recipe blending

Exhibit 4: Possible Characteristics of a Refining Model Template for Special Studies

Characteristic	Description	Comments
Geographic scope		
Level of aggregation	As dictated by the study at hand -- Regional aggregate (e.g., all PADD 3 refineries) -- Notional refinery representation (e.g., typical PADD 3 coking refinery) -- Individual "real" refinery	All models purpose-built from the same model template
Seasonality	As dictated by the study at hand: summer, winter, annual average	
Representation of. . .		
Refining operations	Process-by-process (not extreme point or other simplified representation) Base/delta representation of key refining processes Expanded set of refining process representations available Operating modes mass-balanced	 Recursive solution; delta coefficients determined off-line E.g., coking: delayed coking, fluid coking, flexicoking Study requirements determine which are active Mass balancing via stream yields and densities
Investments in new capacity	Process-by-process; with each process having -- multiple investment options: grass-roots, expansion, retro-fit economics -- investment cost functions (requiring MIP or SOS2 solver), as well as standard average investment values ($K/Bbl/day)	
Stream and blendstock propertie	Stream properties established by recursive pooling Standard energy content values (MM BTU/Bbl) for each blendstock	 To meet product demands specified in energy tems
Refinery energy consumption	Process-by-process Nat gas, still gas, FCC coke	
Refinery emissions of CO_2e	Based on refinery energy consumption, by energy source	
Crude oil slate	Selected from a library of crude assays covering main crudes in commerce Cut points adjustable	Uses a crude assay library and crude assay manager
		Blending Representation
Product slate and blending	CG, RFG, CaRFG (PRM and REG grades for each gasoline type) E85 Kero jet ULSD, No. 2 heating oil Low sulfur marine fuel LPG, Av gas, petchem feedstocks Lubes & waxes, asphalt, pet coke	Specification blending; CM and PM available Recipe blending Specification blending: extended set of specs Specification blending: extended set of specs Specification blending: extended set of specs Recipe blending Recipe blending

The refining sector model for forecasting suggested in Exhibit 3 uses a single aggregate refinery representation for each PADD instead of the PMM dual-refinery representation, for two reasons. One is the principle of modeling parsimony discussed in Section 1; the other is that I do not agree with the marginal price rationale for the dual-refinery approach, and for a number of reasons. (Further discussion of this issue is beyond the scope of this paper, but I would be pleased to address it elsewhere, should EIA wish.)

6. Model Calibration

Regardless of the design and content of the LFMM's refining sector model(s), I urge EIA to adopt the practice of *calibrating* these models annually to corresponding refining operations reported for a recent prior time period.

Calibrating a refining model involves adjusting some of the model's internal technical coefficients – such as yields of refinery streams from certain refining processes, blending properties of refinery streams, or process capacity utilization rates – as needed so that solutions returned by the model closely approximate key measures of refining operations and economics reported for the calibration period(s) – usually the summer and winter gasoline seasons.[13]

Calibration is an iterative procedure. It involves (1) establishing model inputs corresponding to reported inputs to the refining region in the period of interest, (2) solving the model with those inputs, (3) comparing the model's outputs to the reported outputs of the refining facilities being analyzed, (4) adjusting certain technical coefficients in the model, and (5) repeating the preceding three steps until model outputs match with desired accuracy the reported measures of refining operations.

The most important of these reported measures include seasonal average values of:

- Production rates of the primary refined products;
- Capacity utilization of key refining processes (especially the conversion processes and the octane-enhancing processes);
- Properties of the gasoline and distillate pools; and
- Gasoline and distillate product prices.[14]

Every regional refining model should be (re-)calibrated annually.

Model calibration is time-consuming and can be tedious, but it is an essential element of sound modeling practice. Calibration demonstrates that a regional refining model represents with desired accuracy regional refining operations in the prior year – and indeed in the current year, absent significant new regulatory requirements affecting refined products. It not only establishes

[13] This discussion deals specifically with regional refining models, but it applies as well to models of specific individual refineries.

[14] The product prices are the marginal values, or "shadow prices", returned by the model and represent spot prices at the refinery gate.

a sound technical foundation for analytical studies but also enhances the credibility of such studies. Unless a model can be shown to replicate the results of the prior year's refining operations, why should stakeholders accept its representations of prospective future operations?

7. Representation of Advanced Bio-fuels Supply

A number of the high-priority recommendations by LFMM stakeholders [EIA1, EIA2] deal with representing advanced bio-fuels production and uses, representing GHG emissions and other externalities in liquid fuels production, and providing the capability to analyze prospective policy initiatives related to these issues.

It is probably both useful and necessary to establish some sort of "placeholder" modules for such features in the LFMM design. But I suggest that EIA defer efforts to fully implement such modules in the LFMM until the relevant technologies and economics are well-enough defined to be modeled.

All bio-fuels other than corn ethanol and sugarcane ethanol have uncertain technical and economic feasibility and unknown time to achieve commercial production. Moreover, if any advanced bio-fuels achieve commercial status, the resulting industry will be highly fragmented.

Consider, for example, production of cellulosic ethanol – the advanced bio-fuel of primary interest. Cellulosic ethanol production by nature would be more difficult, more complex, and more capital intensive than corn or sugarcane ethanol production. Hence, should it prove to be feasible at commercial scale, cellulosic ethanol will be more costly to produce than corn ethanol or sugarcane ethanol. As yet, no one knows how costly; estimates vary widely. Moreover, the time until initial commercial-scale operation, let alone wide-spread deployment of commercial plants, is unknown.

From an analytical standpoint, cellulosic ethanol is intrinsically different from corn ethanol (as are all other advanced bio-fuels).

- When EPAct2005 established the first RFS (intended for corn ethanol), commercial-scale corn ethanol production had been practiced for more than thirty years. Corn ethanol's production technology, costs, and overall economics were well established, and the primary feedstock (corn) was in commerce and readily available. One could construct an ethanol supply model.

  By contrast, neither the technology nor the economics of large-scale commercial cellulosic ethanol production, assuming it proves feasible, are yet established. No cellulosic ethanol has been produced in sustained, commercial-scale operations. No commercial-scale (≈ 50–60 MM gal/yr) cellulosic ethanol plants are in operation or under construction, and the proposed biomass feedstocks for such plants are not now in commerce. There is nothing that can be modeled for purposes of the LFMM.

- Corn ethanol is produced from a single feed by one of two processes: dry milling (the process of choice) and wet milling. Corn ethanol plants are concentrated in the Mid-west and most produce 60 to 120 million gallons per year.

  By contrast, prospective feedstocks for cellulosic ethanol production include corn stover, wheat straw, rice hulls, other agricultural waste materials, wood chips, other lumber wastes, municipal solid waste, and on-purpose crops such as switchgrass, jatropha, and poplar. Scores of processes are in research and development; most involve bio-chemical conversion, some involve bio-gasification (BTL). The number of prospective feedstock/process/region combinations is formidable, and each combination would have its own supply function. For reason of feedstock availability, commercial plants would be widely dispersed in various regions and would be designed to produce only 40 to 60 million gallons per year.

Even if current R&D activities solve the formidable technical problems associated with cellulosic ethanol production, its rapid deployment could have adverse effects on plant construction costs, which could be difficult to represent in a supply model. For example, the surge in cellulosic ethanol capacity that would be required to achieve the RFS2 mandate volumes – construction of ≈ 250–350 plants in less than a decade – would limit the economic benefits that normally accrue from accumulated experience in designing and operating new facilities. And, the surge would likely trigger bottlenecks in engineering, procurement, and production, with resulting inflation in plant design and construction costs (as happened in the corn ethanol boom from 2003 to 2006).

These considerations have important implications for EIA and the LFMM endeavor.

- No verifiable data will be available to support formal modeling of advanced bio-fuels supply in the LFMM until advanced bio-fuels achieve commercial status, and reliable techno-economic data become available from sustained operation of commercial-scale plants.

- Once advanced bio-fuels production were to attain commercial status, building an endogenous representation of advanced bio-fuels supply comparable to the other supply modules in the LFMM would require capturing bio-fuel production's extreme granularity in terms of process, feedstock, and region. This would entail an extensive and costly initial development effort, involving scores of different supply functions, followed by continuing efforts to stay abreast of developments and update the bio-fuels module(s).

  These efforts will be considerably more costly and labor-intensive than comparable efforts to stay abreast of refining technology. Reliable and accessible sources of information on refining technology exist (e.g., technology providers). Refining technology is mature; advances over the next twenty years are likely to incremental: new catalysts, improved reactor designs, etc., for existing processes. Such incremental changes, coupled with investments in new capacity, are likely to be sufficient to handle even large changes in the refined product slate (such as HCCI fuels, increased diesel/gasoline ratio, low sulfur marine fuel, etc.)).

- Without benefit of data based on commercial operations, forecasting advanced bio-fuels supply involves nothing more than *assuming* the result: either the investment requirement, operating cost, feedstock supply and cost, and geographical dispersion of advanced bio-fuels production capacity or – more simply – an annual production rate profile. The result is a speculation, not a forecast.

- Placing an advanced bio-fuels module in the LFMM before advanced bio-fuels are in sustained commercial production would mislead NEMS stakeholders. The existing NEMS modules all represent existing supply or end-use sectors, with existing capital stock that turns over only gradually, and that generate substantial operating data. An advanced bio-fuels module would represent what is now a non-existent sector. Including such a module in the LFMM would convey the impression that advanced bio-fuels production was as "real" as coal mining, power generation, and oil refining.[15]

I am not suggesting that EIA refrain forever from including advanced bio-fuels production in the LFMM framework. Rather, I am suggesting that EIA refrain from formal modeling of advanced bio-fuels supply until adequate data, based on *sustained commercial operations*, become available. Until then, EIA should use only the most parsimonious set of assumptions regarding future supplies and prices of advanced bio-fuels (e.g., assumed supply volumes and prices by year) and clearly state those assumptions in its forecasts and other publications.

These comments apply equally to CTL and GTL, if the technologies of interest are to incorporate carbon capture and sequestration.

With respect to externalities, it is not yet clear which ones – life-cycle GHG emissions, water and land use, "sustainability," and others – and which prospective policies will emerge as being important and enduring enough to warrant representation in the LFMM.

Finally, various EIA documents mentioned the prospect of establishing formal linkages between LFMM and models of the agricultural sector (whether for representing crop budgets, including energy use, water use, land use effects). In my view, any such effort would be exceptionally complex and expensive, with many organizational complications and technical difficulties, and of questionable value.

8. Accuracy and Precision, One Last Time

Achieving accuracy – and therefore relevance and value – should be the objective of all analysis; achieving precision is the objective of computation. Without accuracy, precision is pointless.

[15] For example, in Exhibit 2 the block representing the virtual sector looks just like the blocks representing the real sectors. Similarly, in EIA publications, forecasts of future supplies from the virtual sector look just like forecasts of supplies from the real sectors.

REFERENCES

[EIA1] "*A White Paper on the Development of a Liquid Fuel Market Module (LFMM) for the National Energy Modeling System*"; Energy Information Administration; May 10, 2009

[EIA2] "*Guidelines for White Papers*"; Energy Information Administration (W. Brown); July 2009

[EIA3] "*Annual Energy Outlook Retrospective Review: Evaluation of Projections in Past Editions (1982-2008)*"; Energy Information Administration;

[EIA4] "*NEMS Petroleum Market Model Documentation; Appendix I: Ethanol Supply Model*"; Energy Information Administration

[NOAD] "*New Oxford American Dictionary, Second Edition*"; Oxford University Press; 2005