This is a working document prepared by the Energy Information Administration (EIA) in order to solicit advice and comment on statistical matter from the American Statistical Association Committee on Energy Statistics. This topic will be discussed at EIA's spring 2006, meeting with the Committee to be held April 6 and 7, 2006.

Improving the Petroleum Refinery Formulation in the SAGE Model

John Staub, OIAF, john.staub@eia.doe.gov, 202-586-6344,
Phillip Tseng, SMG, phillip.tseng@eia.doe.gov, 202-287-1600,

April 7, 2006

Abstract

The United States Energy Information Administration (EIA) completed initial development of the System for the Analysis of Global Energy markets (SAGE) in early 2003. The model is built on a liner programming platform and is solved for the least cost of meeting a predetermined set of energy service demands (e.g. residential lighting, light duty vehicle transportation, etc.). Service demands are driven by exogenous assumptions about population and gross domestic product (GDP). To reduce computation requirements, the model time horizon is divided into five year periods centered on 2005, 2010, etc. The model solves each period sequentially without foresight (i.e. SAGE does not take 2015 service demands into account when solving for the 2010 time period.) SAGE includes sixteen world demand and supply regions. For each region, there are region specific specifications for four end-use demand sectors (commercial, industrial, residential, transportation), petroleum refining, power generation, and supply of both fossil fuels and renewable energy. The structure of SAGE is generic; in the sense a modeler can relatively easily increase the number of demands for energy services and introduce new technologies into the system.

One limitation with the current petroleum refinery formulation in SAGE is its simplistic nature. The current oil refinery formulation in SAGE is a single, extremely flexible technology with six input commodities and 17 output commodities. The only constraint on output commodities is they may not be larger than 50 percent of total input to the refinery. The following paper discusses initial work to enhance the refinery formulation. Work conducted thus far has concentrated on the structure, not the data accuracy. For simplicity, the work has been conducted outside of SAGE with the intention of eventually inserting the enhanced formulation into the SAGE model.

Background on the Original SAGE Refinery

The original oil refinery formulation in SAGE is a single, extremely flexible technology with six input commodities and 17 output commodities (Figure 7). The only constraint on output commodities is they may not be larger than 50 percent of total input to the refinery.

Inputs

BIOLIQ	Biofuel liquids
GANNGA	Natural gas
OINCRD	Crude oil
OINFEE	Refinery Feedstock
OINADD	Additive (MTBE, etc.)
OINNGL	Natural gas liquids

Outputs

GANRFG	Refinery gas
GANETH	Ethane
GANLPG	Liquid petroleum gas
OINGSL	Gasoline
OINAVG	Aviation gas
OINJTG	Jet fuel gas
OINJTK	Jet kerosene
OINKER	Other kerosene
OINDST	Distillates
OINHFO	Heavy fuel oil
OINNAP	Naphta
OINWSP	White spirit
OINWAX	Paraffin wax
OINLUB	Lubricants
OINASP	Asphalt
OINPTC	Petroleum coke
OINNSP	Non specified oil

The lack of detail in the original refinery formulation is repeated in the oil supply formulation. Place holders were originally created for three crude oil types (light, mid & heavy), but supply curves were only developed for one type of crude oil. Moreover, the supply curve for oil production has only three steps with large differences between price steps.

Trends in crude oil prices and petroleum products

Crude oil prices are affected by product demand, API gravity, sulfur content, and refining capacity. Crude oil prices for different grades (light vs. heavy crude oils) diverge as prices rise (Figure 1). This is driven by two main factors. First, light crude oil production, as a share of total crude oil production, has declined since the late 1970's. Figure 2 shows the share U.S. imports of light crude oil declining as heavy crude imports have risen. The drop in light crude oil imports and rise in heavy oil imports imposed constraints on refiners' capability to produce light products. As a result, light crude oil commanded bigger premium. Secondly, not all petroleum refineries are able to process heavy crude oil. Thus a price premium for light crude oil develops when demand is high. Investments in down stream conversion units provide added capability in processing heavy crude oils. The observed price differences between light and heavy crude oils reflect changes in the composition of oil supply, additional investment in down stream conversion units, and the removal of excess teakettle refinery capacity during the 1980s and 1990s (Figure 3).

Demand for petroleum products has also shifted toward lighter products, i.e. gasoline and diesel fuels (Figures 4a and 4b). Strong demand growth in China and the U.S. is expected to sustain higher oil prices over time. Assuming all petroleum product prices will rise to some extent, then demand for heavier petroleum products should fall as traditional consumers of heavier petroleum products (such as electricity generators), switch to coal.

Projected changes in crude oil quality, demand for petroleum products, and refinery capacities are of great interest to energy analysts and policy makers. Furthermore, world oil consumption is projected to increase by about 50% in the next 25 years (Figure 5). Converting crude petroleum to end use products will require significant investments.

Figure 1.

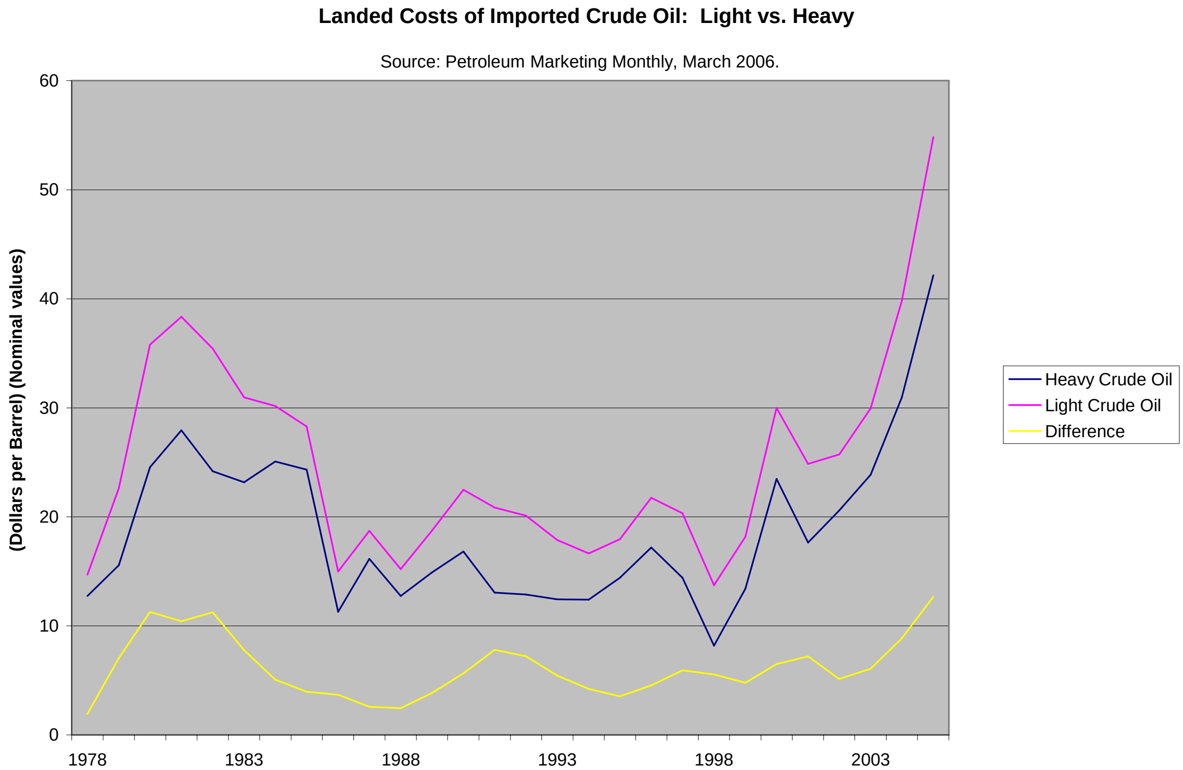


Figure 2.

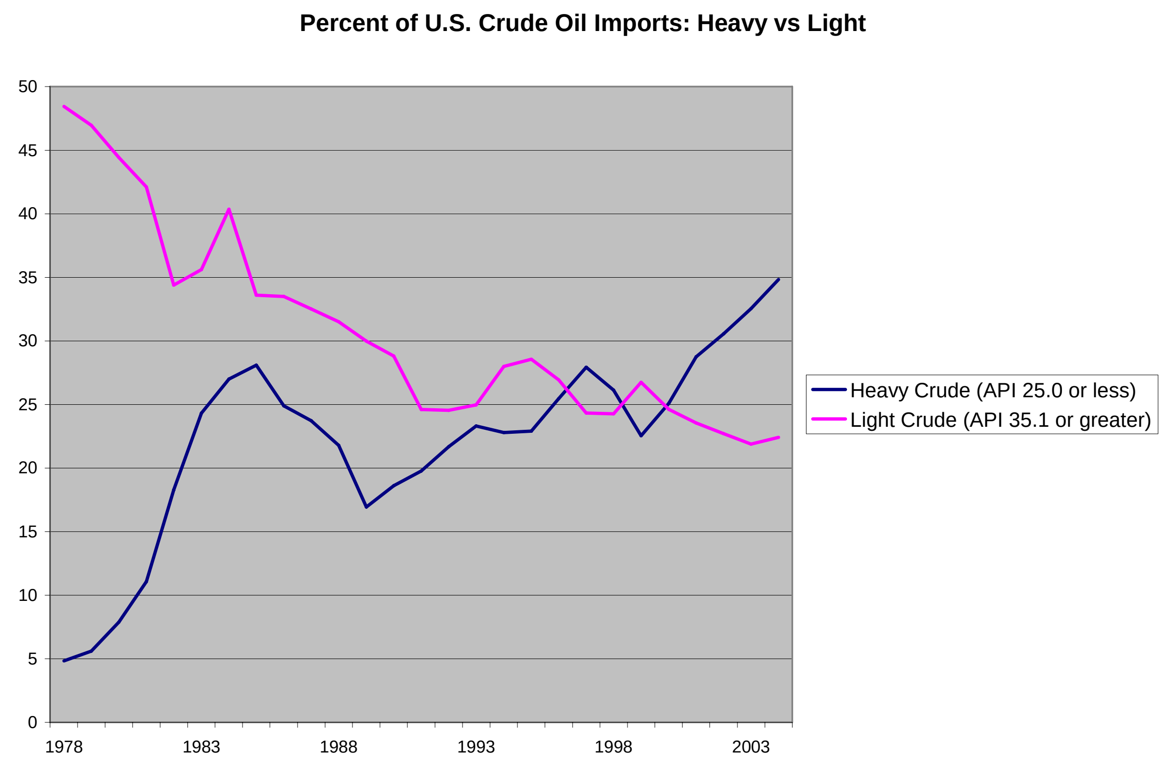


Figure 3.

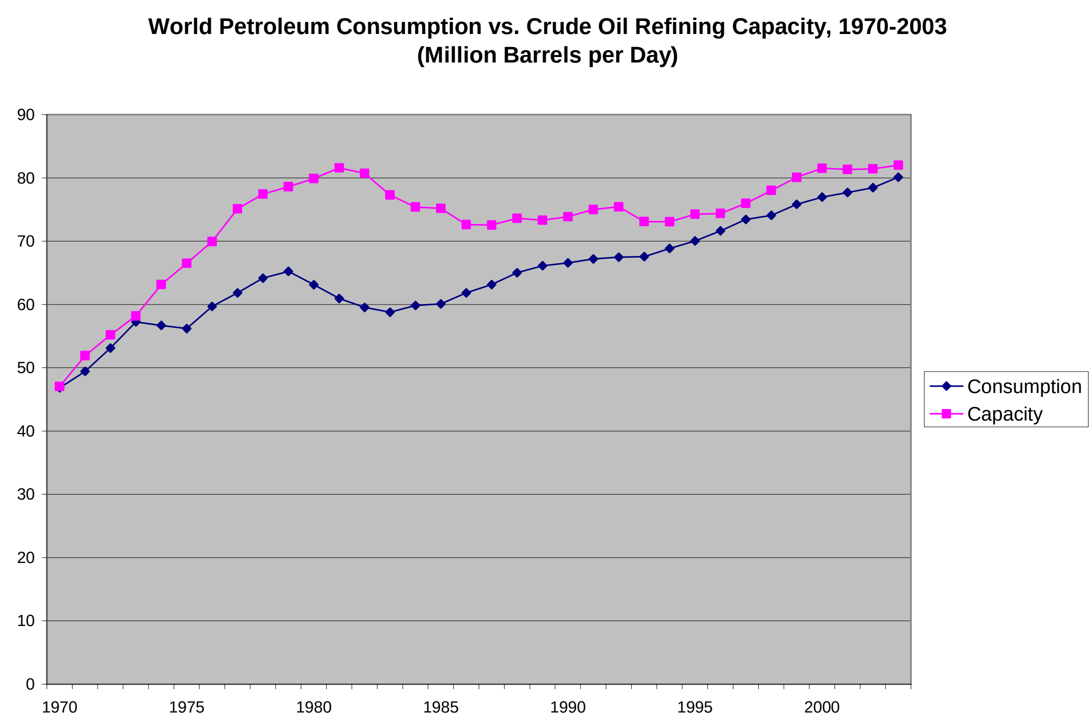


Figure 4a.

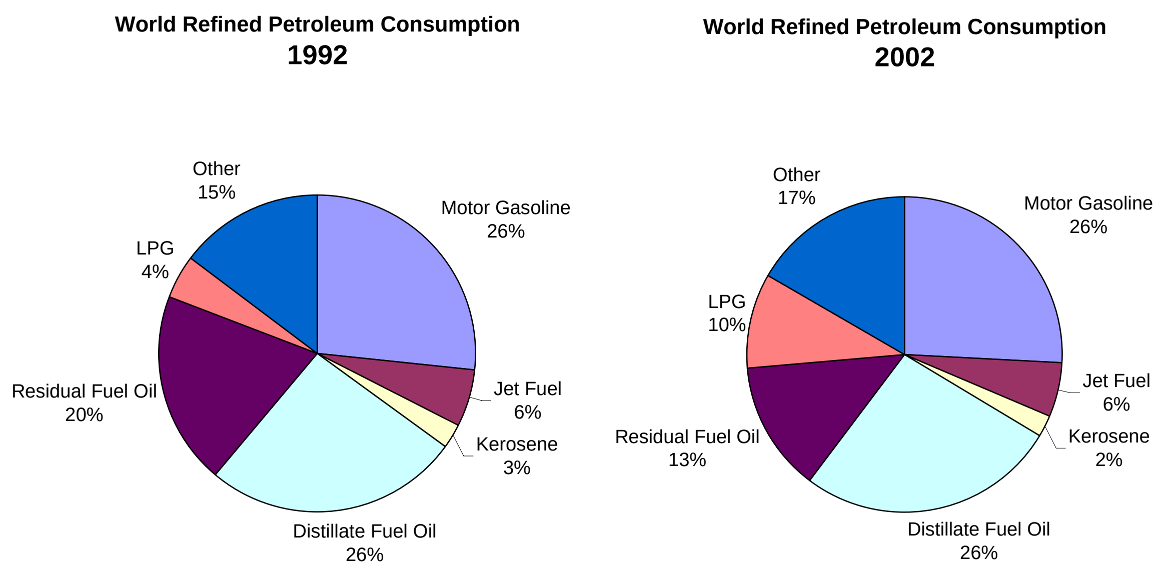


Figure 4b.

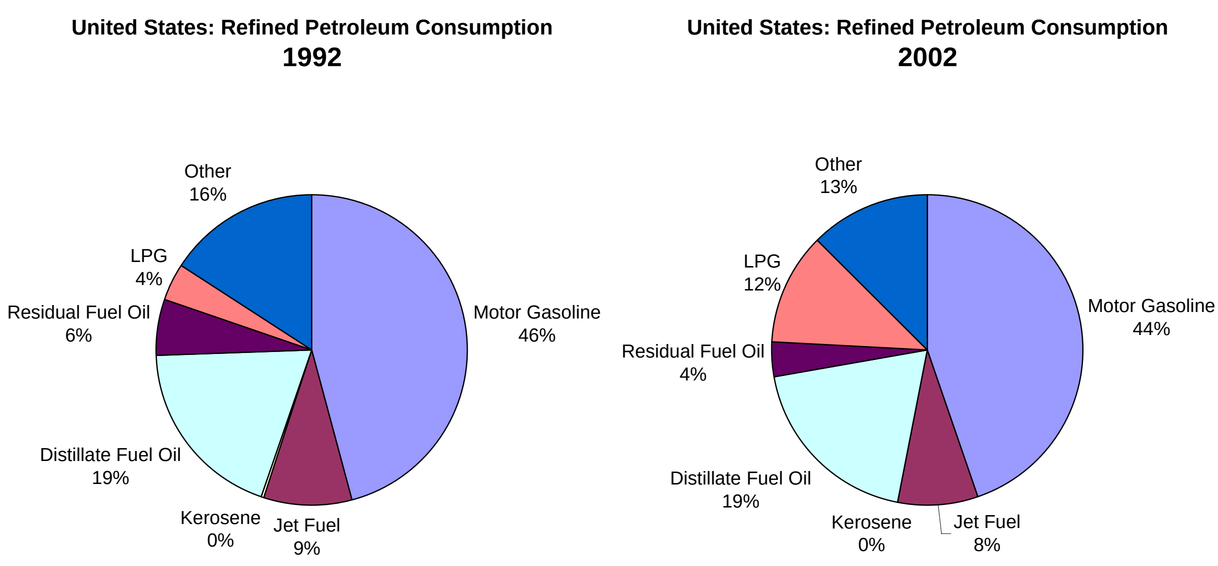
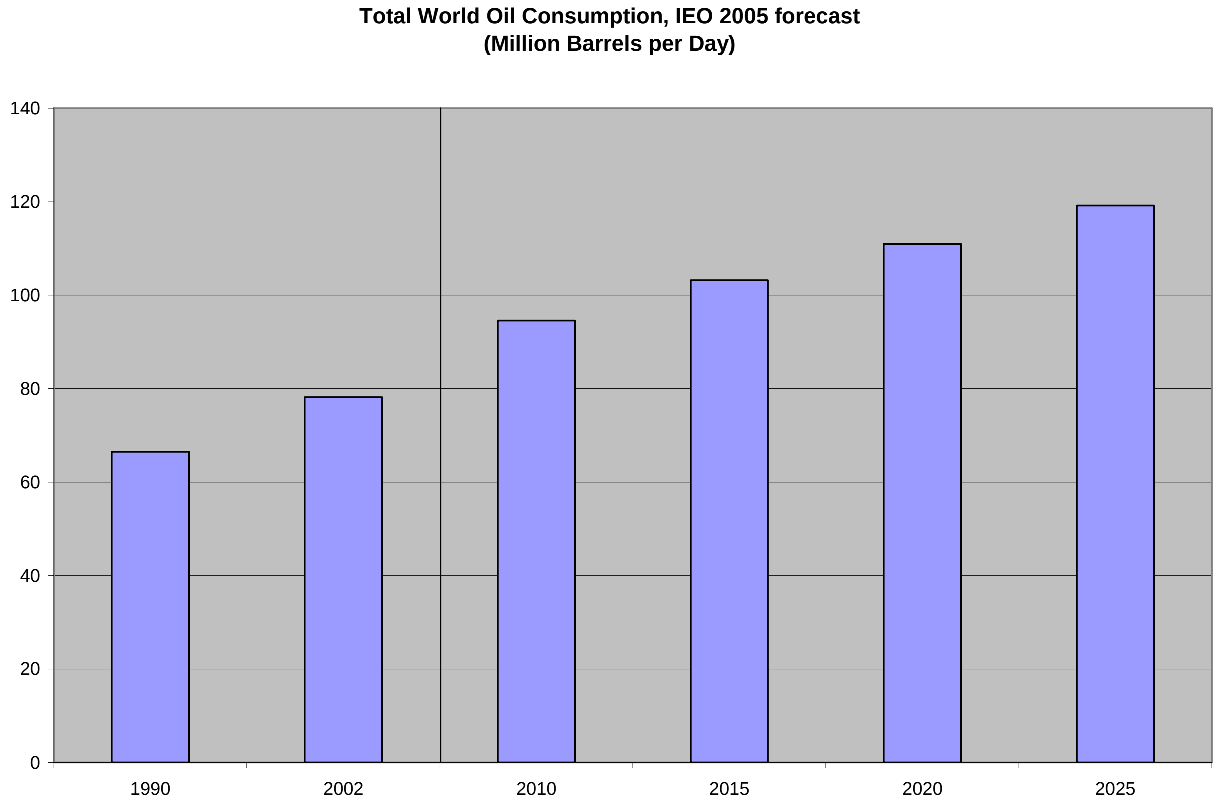


Figure 5.


Total World Oil Consumption, IEO 2005 forecast
(Million Barrels per Day)
140
120
100
80
60
40
20
0
1990
2002
2010
2015
2020
2025

Purpose of the new refinery representation

A new refinery representation is needed for four main reasons. First, we need to correctly account for the cost of adding refining capacities and observe investment decisions. This added capability makes it possible to simulate and assess the impacts of unexpected changes in demand.

Second, the improved representation of petroleum refining allows us to capture interactions between demand for end-use refined products and demand for light or heavy crude oil. With this added capability, the model provides results, which reflect the interactions of fuel switching by technology, fuel and region. For example, as population and GDP rise in China, will the demand for transportation reduce the petroleum available for electricity production or will coal-to-liquids meet the transportation fuel requirements?

Third, this new formulation allows us capture interactions in the international crude oil trade, international refined petroleum product trade and expansions in refinery capacity.

Fourth, with the improved representation of petroleum refining, model results will show more realistic pricing of refined products. These prices can help us determine technology choices in the end use sectors; it can also provide insights when we run scenario analysis that may be useful to policy makers. Representing the cost of petroleum refining is also a pre-requisite for correctly modeling technology choices on the supply.

Figure 6. Flow chart of typical refinery operations

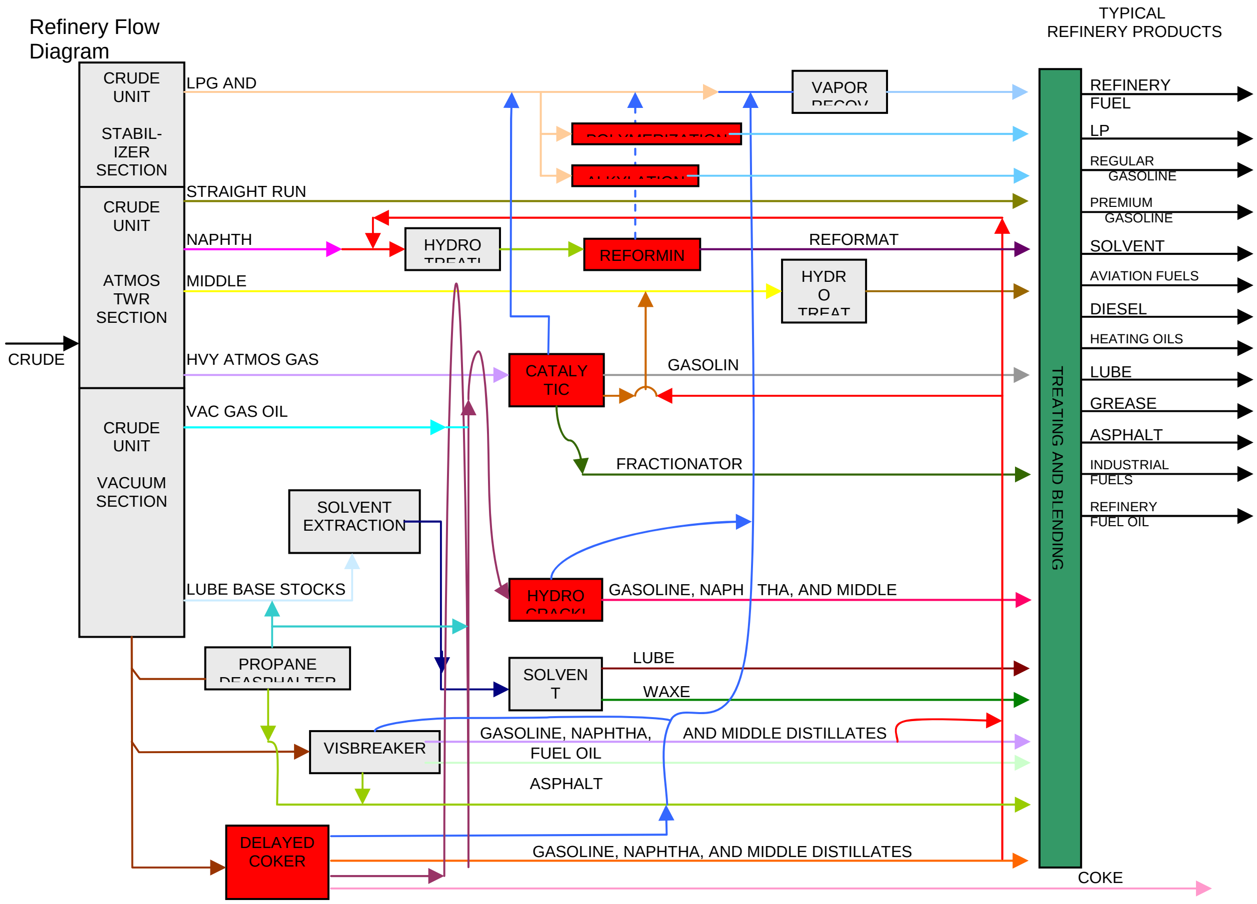


Figure 7. Original SAGE "flexible" refinery

SAGE "flexible" refinery inputs

crude X
Natural Gas Liquids
Natural Gas
Refinery Feedstock
Biofuel Liquids
Additives (MTBE, etc.)

Any combination produces up to 50 percent of any of these →

17 SAGE "intermediate" petroleum products

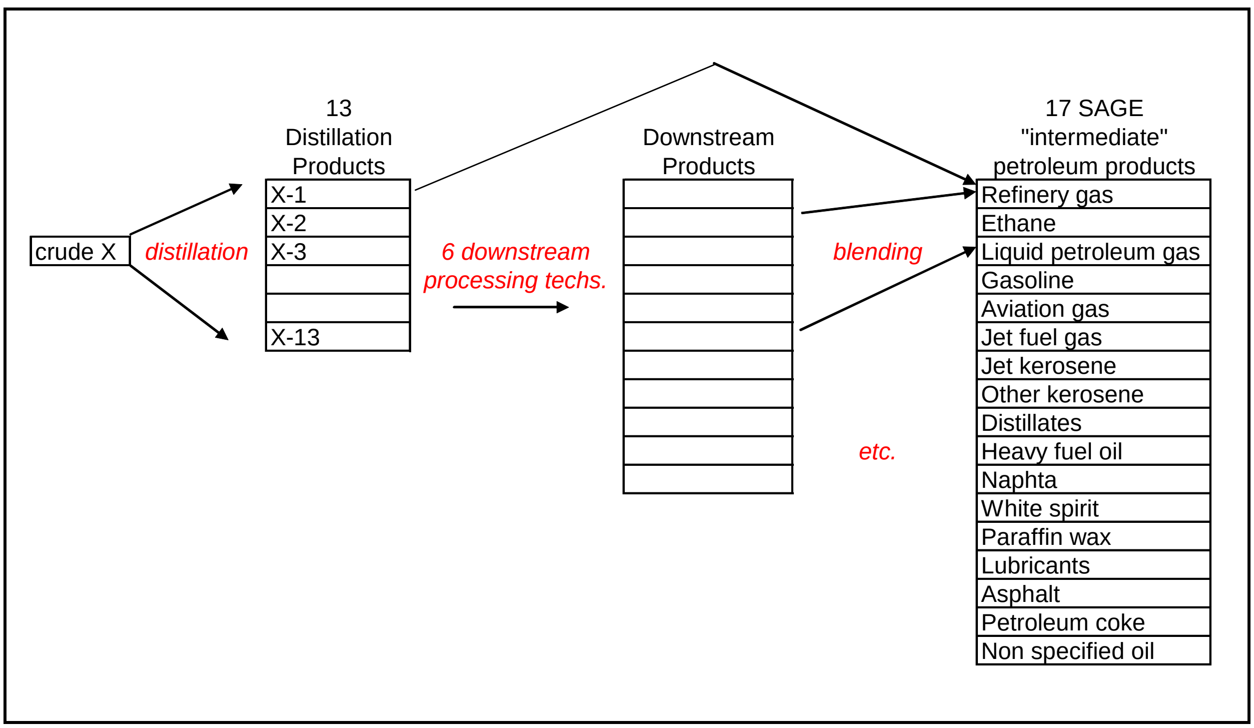
Refinery gas
Ethane
Liquid petroleum gas
Gasoline
Aviation gas
Jet fuel gas
Jet kerosene
Other kerosene
Distillates
Heavy fuel oil
Naphta
White spirit
Paraffin wax
Lubricants
Asphalt
Petroleum coke
Non specified oil

Model structure

We have built a preliminary enhanced refinery for one of the sixteen regions (Figure 8). This was done outside of the SAGE model, because of the lengthy solution times when solving the whole SAGE model. It has been structured so that eventually it can be inserted directly into the SAGE model, replacing the existing flexible refinery. We have run the enhanced refinery model only to the extent of making sure it reaches a solution, and changes to input data move the solution in the expected direction. Note: The single crude oil in the old SAGE refinery is first broken into the five crudes with fixed shares, by analyst judgment.

The structure of the enhanced refinery includes five types of crude oil instead of one. Each crude oil type is assigned a specific distillation process, which captures differences in API gravity and sulfur content. The output of the distillation process is 13 distillation streams. These flow to six downstream processing technologies (hydro-cracker, fluid cat-cracker, coker, reformer, alkylation, and polymerization) for further processing and upgrading. Each of the six downstream processing technologies has a lower and an upper operating bound, which provides flexibility to the refinery. Finally, a set of blending recipes are used to mix outputs from the distillation and six downstream processing technologies, into 17 "intermediate petroleum products." Demand for the 17 SAGE intermediate petroleum products is what drives the refinery operation.

Figure 8. Schematic of Enhanced Refinery



Analysts may adjust many of the refinery parameters on a region by region basis. For example, the ratios of distillation output products for the five crude oil types in the model may be adjusted to reflect different compositions of crude oils. Light-low-sulfur crude oil when distilled produces more low density oils and less residual oil than heavy-sour-crude oil does. To the extent the model is permitted to switch between the crude oils, this is one way it can adjust the final quantities of refined product outputs.

Analysts are also free to adjust the input and output product ratios of the six downstream processing technologies. Furthermore, from the perspective of the model, there are 12 downstream processing technologies because each technology has a high and low "operating severity formulation" (i.e. setting). This provides a second level of flexibility in producing refined petroleum products. One challenge we have encountered is the current data is based on units of volume while the model tracks units of energy. The volume data ideally needs to be converted to an energy only basis.

A third lever available to analysts is the blending recipe used to mix distillation and downstream processing outputs into the 17 SAGE intermediate petroleum products. The 17 products are consumed by end-uses.

Finally, analysts are able to attach distillation costs onto crude oil just before it enters the distillation unit, and the cost of operating downstream processing units per unit of liquid processed.

Data Sources

Four sets of data are required by the enhanced refinery model formulation.

1. Distillation product ratios specific to each crude oil type. These are often referred to as "crude oil assays." (*Oil and Gas Journal*)
2. I-O (input/output) coefficients for downstream processing units (*Petroleum Engineer*)
3. Product blending/recipe specifications (*Petroleum Engineer*)
4. End-use demand for refinery products (IEA, EIA, and *Oil and Gas Journal*)

For the preliminary runs of the enhanced refinery, we used five distillation assays from twenty year old EIA data. The downstream units' input and output shares are also from twenty year old EIA data. They are currently in terms of input barrel per output barrel. This will need to be converted to energy units to properly handle refinery volume gains, because SAGE is in energy units.

Final product demand was based on the *EIA Annual Energy Review 2004*, table 5.8, using data for 2003. The *EIA Petroleum Marketing Annual 2004*, table 3, year 2003 average was also consulted.

Data Issues

- A major challenge is the availability of detailed data on refinery output.

- Calibration of refinery output to match historical demand.

Questions for the Committee

1. Is the enhanced formulation described adequate for the petroleum product trade analysis we desire to conduct? Should we have more or less details?

2. Several preliminary runs of the enhanced refinery model demonstrated a problem in which it used too much butane to produce gasoline. One solution is to include octane constraints. The drawback to this solution is SAGE would need additional equations. Are there other simpler solutions to this problem?