Analysis of Petroleum Refining Energy Efficiency of U.S. Refineries

Hao Cai[a], Jeongwoo Han[a], Grant Forman[b], Vince Divita[c], Amgad Elgowainy[a], Michael Wang[a]

[a] Systems Assessment Group, Energy Systems Division, Argonne National Laboratory
[b] Sasol North America Inc, USA
[c] Jacobs Consultancy, USA

1. Introduction

The overall energy efficiency of a petroleum refinery that converts crude oil into many finished products is key to estimating the energy consumption, greenhouse gas (GHG) and air pollutant emissions of the refinery and its finished products. U.S. petroleum refineries receive crudes oil from various countries and regions, in addition to domestic crude oil. The U.S. had 134 operating refineries with a total annual operating capacity of 6.13 billion barrels (bbl), or 17.9 million bbl per stream day in 2012 (EIA, 2013a). The crudes received and refined by the refineries differ in their API gravity and the sulfur contents, among other characteristics (EIA, 2013b). Refinery configurations, including the complexity of the processing units, also differ (EIA, 2013c). Furthermore, product yields of refineries and the slate of finished products differ significantly by the Petroleum Administration Defense District (PADD) region (EIA, 2013d), as shown in Table 1. As a result, the refinery overall efficiency can vary among refineries and among the PADD regions, because the refining energy consumption are known to be functions of crude oil properties, refinery configurations, finished product slate, among others (Hirshfeld and Kolb, 2012).

Table 2 shows shares of crude oil sources and the crude oil properties by region in 2012 (EIA, 2013e). Table 2 also presents average transportation distances from various regions to the U.S. receiving ports. The transportation distance information is used in GREET modeling of crude oil transportation. On the basis of EIA's projection of domestic crude production and the imports of foreign crudes by country and region (EIA, 2013e), we estimated the crude source mixes up to 2020, as shown in Table 3. In particular, the volumetric share of the imported Canadian oil sands in historical years (up to 2012) was estimated based on the amount of the heavy crude from Canada that is imported by PADD II (CAPP, 2013) and the total amount of U.S. crudes that go into refineries (EIA, 2013e). This approach is likely to result in a conservative estimation of the share of the Canadian oil sands. For future years, we assumed that the volumetric amount of the imported Canadian conventional oil will remain constant because the Canadian conventional oil production will stay almost at the same level (CAPP, 2013), with the rest amount of imported Canadian oil being oil sands.

In this memo, we document the analysis of the linear programing (LP) modeling results of 43 U.S. refineries, which represent about 70% of the total U.S. refining capacity in 2012, as shown in Table 4. With this data we estimated the overall efficiency of each individual refinery. Moreover, we investigated the impacts of crude oil properties, refinery configurations, and the finished product slate on the refinery overall efficiency by exploring the correlation between the refinery overall efficiency and these factors. Our aim is to develop a regression model, in which parameters that are known to have a cause-effect relationship with the refining efficiency are incorporated. Only parameters that are statistically significant predictors of the overall refinery efficiency were considered for the regression analysis. This sought correlation by the regression analysis is then used to predict refinery overall efficiencies in the future based on such factors as crude API gravity, sulfur content, finished product yields, etc., for incorporation into the Greenhouse gases, Regulated Emissions, and Energy use in Transportation (GREET^{TM}) model developed at Argonne National Laboratory (ANL, 2013).

Table 1. Refining production and product slate (in vol%[a]) by PADD region in the U.S., 1990-2012

	Total production of all refining products					**Gasoline**					**Diesel**				
	PADD 1	**PADD 2**	**PADD 3**	**PADD 4**	**PADD 5**	**PADD 1**	**PADD 2**	**PADD 3**	**PADD 4**	**PADD 5**	**PADD 1**	**PADD 2**	**PADD 3**	**PADD 4**	**PADD 5**
1990	9.9%	21.7%	46.7%	3.0%	18.7%	0.0%	27.6%	49.7%	3.8%	19.0%	12.0%	24.2%	44.6%	4.3%	14.8%
1991	9.9%	21.8%	46.5%	3.1%	18.7%	0.0%	27.7%	48.9%	3.8%	19.6%	12.1%	24.3%	44.7%	4.3%	14.5%
1992	10.1%	21.9%	46.6%	2.9%	18.5%	0.0%	27.6%	49.7%	3.5%	19.2%	12.3%	24.5%	44.5%	4.1%	14.7%
1993	10.4%	21.4%	47.1%	2.9%	18.3%	10.7%	23.7%	45.3%	3.2%	17.1%	12.8%	24.5%	44.7%	4.0%	14.2%
1994	10.2%	21.4%	47.1%	2.9%	18.3%	9.3%	24.3%	45.2%	3.4%	17.8%	12.6%	24.7%	44.8%	4.0%	13.9%
1995	11.0%	21.3%	47.0%	2.9%	17.8%	4.6%	25.6%	48.0%	3.5%	18.3%	13.1%	24.4%	44.3%	4.1%	14.1%
1996	10.2%	21.5%	47.5%	3.0%	17.8%	4.1%	25.8%	48.1%	3.6%	18.4%	11.6%	24.9%	46.1%	4.1%	13.3%
1997	11.1%	21.2%	47.4%	2.9%	17.4%	4.8%	25.6%	47.7%	3.6%	18.4%	12.5%	24.8%	45.4%	4.1%	13.3%
1998	11.2%	21.2%	47.5%	2.9%	17.2%	4.9%	25.3%	47.8%	3.5%	18.5%	12.9%	25.4%	44.3%	4.0%	13.4%
1999	11.3%	20.8%	48.1%	3.0%	16.7%	5.1%	24.8%	48.5%	3.6%	18.0%	12.7%	24.5%	45.5%	4.2%	13.1%
2000	11.2%	20.5%	48.4%	3.0%	16.9%	5.0%	24.0%	48.7%	3.7%	18.5%	12.8%	24.2%	45.8%	4.1%	13.0%
2001	11.1%	20.2%	48.4%	3.0%	17.2%	5.2%	23.8%	48.4%	3.6%	19.0%	12.6%	23.4%	46.7%	4.1%	13.2%
2002	11.2%	20.1%	48.0%	3.1%	17.5%	5.2%	24.1%	47.6%	3.7%	19.3%	12.7%	22.9%	46.3%	4.4%	13.7%
2003	11.4%	19.7%	48.2%	3.1%	17.6%	5.3%	23.8%	47.6%	3.8%	19.5%	12.2%	22.6%	47.2%	4.2%	13.8%
2004	11.7%	19.4%	48.6%	3.2%	17.1%	5.4%	23.5%	48.1%	3.8%	19.3%	11.7%	22.4%	48.1%	4.4%	13.4%
2005	12.5%	19.6%	46.9%	3.2%	17.7%	6.3%	22.9%	46.8%	3.8%	20.1%	12.5%	23.0%	46.4%	4.3%	13.8%
2006	13.8%	19.4%	46.2%	3.2%	17.4%	6.5%	23.4%	45.6%	3.9%	20.6%	12.0%	22.6%	47.7%	4.2%	13.5%
2007	14.3%	19.1%	46.5%	3.1%	17.0%	6.6%	23.4%	45.8%	4.0%	20.2%	12.0%	22.4%	48.7%	3.9%	13.0%
2008	16.2%	20.5%	43.1%	3.2%	17.1%	10.4%	26.4%	39.4%	4.0%	19.8%	11.1%	23.0%	48.8%	3.9%	13.2%
2009	17.8%	20.5%	41.7%	3.3%	16.6%	15.4%	26.7%	33.9%	4.1%	19.9%	9.6%	22.2%	51.9%	4.2%	12.2%
2010	18.7%	21.3%	40.7%	3.1%	16.2%	19.1%	27.9%	29.5%	3.8%	19.7%	8.7%	22.8%	52.6%	4.1%	11.8%
2011	19.4%	21.6%	39.6%	3.1%	16.3%	21.5%	28.5%	26.6%	3.7%	19.8%	8.2%	22.0%	53.8%	4.0%	12.0%
2012	18.9%	21.9%	39.7%	3.3%	16.2%	21.6%	28.5%	25.9%	4.0%	20.0%	7.6%	22.3%	54.3%	4.2%	11.5%

	Jet					Heavy products[b]					LPG				
	PADD 1	**PADD 2**	**PADD 3**	**PADD 4**	**PADD 5**	**PADD 1**	**PADD 2**	**PADD 3**	**PADD 4**	**PADD 5**	**PADD 1**	**PADD 2**	**PADD 3**	**PADD 4**	**PADD 5**
1990	1.8%	14.3%	58.1%	3.5%	22.3%	14.4%	7.7%	38.5%	1.1%	38.3%	8.2%	20.7%	58.3%	1.4%	11.3%
1991	2.2%	14.3%	55.1%	3.4%	24.9%	15.1%	8.4%	40.9%	1.1%	34.6%	7.5%	20.4%	59.2%	1.2%	11.7%
1992	3.7%	12.0%	56.9%	3.7%	23.7%	14.3%	9.1%	40.7%	1.3%	34.7%	7.8%	20.5%	56.8%	1.2%	13.6%
1993	2.6%	12.4%	56.5%	3.5%	25.0%	15.7%	19.3%	36.4%	3.4%	25.2%	7.8%	20.4%	56.3%	1.2%	14.3%
1994	2.0%	11.7%	57.4%	3.2%	25.7%	16.7%	20.2%	35.4%	3.2%	24.4%	7.4%	19.0%	58.7%	1.1%	13.7%
1995	1.0%	14.2%	57.8%	2.3%	24.6%	19.1%	19.5%	36.1%	3.6%	21.7%	7.3%	19.9%	59.3%	1.1%	12.5%
1996	1.2%	19.2%	54.7%	2.8%	22.0%	17.5%	20.2%	35.5%	4.1%	22.6%	6.5%	19.0%	62.2%	0.9%	11.4%
1997	0.5%	19.5%	56.0%	2.2%	21.7%	17.3%	20.9%	35.4%	4.4%	22.0%	6.7%	19.0%	62.8%	0.9%	10.5%
1998	0.6%	22.4%	56.3%	2.3%	18.5%	17.7%	19.9%	38.0%	4.2%	20.2%	6.8%	17.0%	65.2%	0.6%	10.4%
1999	1.7%	22.9%	64.1%	2.1%	9.2%	16.3%	20.2%	36.8%	4.6%	22.1%	6.0%	18.0%	65.6%	0.9%	9.4%
2000	1.2%	24.8%	59.2%	2.9%	11.8%	17.3%	19.9%	39.1%	4.7%	18.9%	6.4%	17.6%	64.0%	1.0%	11.0%
2001	1.1%	22.6%	58.6%	2.4%	15.2%	16.7%	19.6%	40.1%	4.5%	19.0%	6.9%	17.9%	63.8%	0.9%	10.5%
2002	1.2%	24.7%	57.7%	2.2%	14.2%	17.5%	20.9%	36.3%	5.4%	19.9%	7.3%	17.1%	64.2%	0.9%	10.5%
2003	0.0%	24.8%	57.5%	2.3%	15.4%	19.4%	20.4%	37.4%	5.3%	17.5%	7.6%	16.2%	63.9%	0.7%	11.6%
2004	0.0%	22.3%	58.6%	1.8%	17.4%	18.8%	21.2%	36.5%	5.6%	17.9%	6.8%	17.2%	64.7%	0.7%	10.5%
2005	0.0%	23.6%	63.4%	1.8%	11.3%	18.0%	21.7%	36.7%	6.0%	17.6%	7.2%	18.3%	61.9%	0.7%	11.9%
2006	0.0%	19.6%	65.4%	2.2%	12.8%	18.5%	22.8%	35.4%	5.4%	17.8%	6.9%	21.1%	58.7%	1.2%	12.1%
2007	0.0%	19.0%	68.0%	2.9%	10.1%	18.2%	20.4%	37.3%	5.5%	18.6%	8.2%	19.6%	59.7%	1.2%	11.3%
2008	0.0%	20.2%	65.3%	2.7%	11.8%	19.0%	22.1%	37.1%	4.3%	17.6%	8.3%	18.4%	60.2%	1.4%	11.7%
2009	0.0%	17.1%	67.0%	2.8%	13.2%	18.1%	20.3%	40.3%	4.3%	17.1%	7.5%	17.1%	62.9%	1.6%	10.9%
2010	0.0%	17.9%	69.5%	2.5%	10.1%	13.7%	22.0%	44.5%	4.2%	15.6%	6.1%	17.3%	65.0%	1.6%	10.0%
2011	0.0%	15.5%	71.2%	3.0%	10.2%	13.3%	23.6%	41.1%	4.4%	17.6%	7.1%	18.5%	64.1%	1.3%	9.0%
2012	0.0%	18.6%	68.3%	3.5%	9.6%	13.5%	24.4%	38.8%	5.5%	17.9%	6.2%	19.4%	65.0%	1.5%	8.0%

[a]: volume/volume %;

[b]: heavy products includes residual oil and asphalt.

Table 2. Shares of crude oil by region that is received by U.S. refineries in 2012

	U.S. Domestic[a]	Canada (Oil Sands)	Canada (Conv. Crude)	Mexico	Middle East	Latin America	Africa	Others
Crude source	44.1%	7.3%[b]	8.7%[c]	6.4%	14.0%	11.3%	6.8%	1.4%
API gravity	31.1	28.4	26.5	26.5	31.8	24.8	38.3	32.0
S content (wt%)	1.41	1.26	2.95	2.20	2.26	2.80	0.29	0.76
Transportation distances (mi)		1,708	1,708	797	14,596	4,620	7,280	6,128

[a] EIA, 2013e and EIA, 2013f;

[b] CAPP (2013) reported these heavy crude consumption rates in the U.S.: 0.122, 1.100, 2.214, 0.195, and 0.384 M b/d heavy crude in PADD I, II, III, IV, and V, respectively. We assumed that all heavy crude consumed in PADD II was oil sand products, resulting in an estimated oil sands consumption of 1.10 M b/d by U.S., which is 7.3% of total US consumption in 2012. CAPP (2013) estimated that the consumption of oil sands by Canadian refineries was 0.52 M b/d. Thus, total U.S. and Canadian oil sands consumption is 1.62 M b/d. For comparison, CAPP (2013) estimated total production of Canadian oil sands products was 1.80 M b/d in 2012;

[c] Total Canadian oil export to the U.S. in 2012 was estimated to be 16.0% of total U.S. crude consumption (EIA, 2013e and EIA, 2013f). The Canadian conventional crude export share of 8.7% is the difference between the total export share and our estimated oil sands export share.

Table 3. Crude source mixes (in vol%[a]) and the shares of Canadian oil sands as inputs in the U.S. refineries in 2010, 2015 and 2020

	2010	2015[b]	2020[b]
US	35.3	44.9	46.7
Canada	13.3	16.5	22.4
Oil sands	5.7	8.1	13.9
Conventional	7.6	8.4	8.5
Mexico	7.8	6.8	6.5
Latin America	13.5	9.9	8.4
Middle East	11.4	8.4	5.7
Africa	14.8	10.9	8.1
Others	3.9	2.6	2.2

[a]: Volume/Volume %;

[b]: Based on the forecast of the total crude oil production and total crude oil demand in Canada by CAPP (2013), and on an assumption that the net surplus of Canadian crude oil production will all be exported to the U.S.

Table 4. Coverage of annual refining capacity (thousand bbl/day) by LP modeling by PADD region

	Capacity evaluated with LP modeling	**Total refinery capacity[a]**	**Coverage of LP modeling**
PADD I	404	940	43%
PADD II	2,150	3,468	62%
PADD III	5,983	7,765	77%
PADD V	1,956	2,330	84%
Total	**10,493**	**14,990**	**70%**

[a]: from EIA (2013e) for year 2012.

2. Method and data

2.1 Overall energy efficiency of 43 individual refineries

We estimated the overall energy efficiency of individual refineries based on the total energy input and total energy output on the lower heating value (LHV) and higher heating value (HHV) basis, using Equations (1) and (2), respectively.

$$\eta_{i,LHV} = \frac{\sum_n (P_n \times LHV_n)}{\sum_m (C_m \times LHV_m) + \sum_0 (OI_o \times LHV_o) + NG_{purchased,LHV} + H_{2,purchased,LHV} + Electricity_{purchased}} \quad (1)$$

$$\eta_{i,HHV} = \frac{\sum_n (P_n \times HHV_n)}{\sum_m (C_m \times HHV_m) + \sum_0 (OI_o \times HHV_o) + NG_{purchased,HHV} + H_{2,purchased,HHV} + Electricity_{purchased}} \quad (2)$$

Where:

$\eta_{i,LHV}$ and $\eta_{i,HHV}$ are the LHV- and HHV-based overall efficiency of refinery *i*, respectively;

P_n is the amount of refining product *n*;

C_m is the amount of crude input *m*;

OI_o is the amount of other input material *o*, e.g., normal butane, isobutane, reformate, alkylate and natural gasoline;

$NG_{purchased,LHV}$ and $NG_{purchased,HHV}$ are the LHV- and HHV-based energy of purchased natural gas;

$H_{2,purchased,LHV}$ and $H_{2,purchased,HHV}$ are the LHV-and HHV-based energy of purchased hydrogen;

$Electricity_{purchased}$ is energy in purchased electricity;

and LHV_m, HHV_m, LHV_n, HHV_n, LHV_o and HHV_o are the LHVs and HHVs of crude input *m*, refining product *n*, and other input material *o*, respectively.

The crude sources and refining products vary by refinery, and so do their LHVs and HHVs, resulting in differences in the LHV- and HHV-based overall efficiencies.

2.2 PADD-specific refinery energy efficiencies based on EIA statistics

The U.S. Energy Information Agency (EIA) publishes annual statistics on volumetric crude oil and blending stock inputs, the captive hydrogen consumption, consumption of fuel gas and process fuels like marketable petroleum coke, catalyst petroleum coke (carbon that deposits on a process catalyst and subsequently is removed via a combustion process that is a source of refinery energy), natural gas, residual fuel oil and coal, consumption of purchased electricity and steam, and refining products outputs of the U.S. refineries on the PADD level. The latest data available are for year 2011 (EIA, 2012a; EIA, 2012b) and were used for the estimation of the refinery overall efficiencies by the PADD region, using Equation (1). We note that the EIA statistics does not include merchant hydrogen consumption by refineries. We collected the 2006 data on merchant hydrogen consumption from the Chemical Economics Handbook (CEH, 2007), and we assumed that the hydrogen consumption per bbl of crude remained constant since 2006. This appears to be a reasonable assumption, since the sulfur contents of crude oil did not change considerably between then and now, and the sulfur content requirement for gasoline (30 ppm) and diesel (15 ppm) remained the same during this period (EPA, 2013a; EPA, 2013b).

2.3 Linear programing modeling results of 43 U.S. refineries

For this study, we analyzed the LP modeling results of 43 U.S. refineries provided by Jacobs Consultancy. LP is a rigorous, widely used mathematical modeling approach for evaluating different refineries operation strategies. Refinery LP modeling is process-oriented representation of refinery operations, the material flows between refining processes, the prices of inputs and outputs, and refinery technical and economic responses to changes in requirements of petroleum products, such as sulfur contents of gasoline and diesel and higher octane number of gasoline (Hirshfeld and Kolb, 2012).

The LP modeling results of 43 refineries provided by Jacobs Consultancy included material and energy flows of the crude oil, blending stocks, fuel gas, purchased hydrogen, electricity and steam, and refining products of individual refineries. The refineries differed in their processing capacity, the sources and properties of crude oil, the refinery complexity, and the finished product slates, as shown in Table 5.

Table 5. Operational characteristics of the 43 refineries with LP modeling

PADD region		Capacity[b]	API gravity	Sulfur (wt%)	WP ratio[c]	G/D ratio[d]	HP ratio[e]	Complexity index
PADD I	Weighted average[a]	202	34.7	1.2	0.9	2.2	0.07	9.4
	Minimum		32.9	0.5	0.8	1.7	0.06	9.1
	Maximum		35.9	2.1	0.9	2.5	0.09	9.9
PADD II	Weighted average[a]	239	28.6	2.0	0.9	2.1	0.14	10.1
	Minimum		21.5	0.5	0.7	1.6	0.05	8.8
	Maximum		39.1	3.0	0.9	2.7	0.19	12.6
PADD III	Weighted average[a]	299	28.1	2.1	0.9	1.8	0.08	10.8
	Minimum		20.3	0.5	0.8	1.1	0.04	7.8
	Maximum		36.4	3.5	1.0	2.7	0.12	13.4
PADD V	Weighted average[a]	163	24.9	1.3	0.8	2.7	0.17	11.9
	Minimum		16.2	0.9	0.7	1.5	0.07	7.0
	Maximum		33.1	1.7	0.9	5.4	0.25	15.4

43 refineries	Average	244	27.9	1.9	0.9	2.0	0.11	10.8
	Minimum		16.2	0.5	0.7	1.1	0.04	7.0
	Maximum		39.1	3.5	1.0	5.4	0.25	15.4

[a] On the volume of crude oil inputs basis;
[b] In thousand bbl of crude oil inputs per stream day;
[c] The volumetric ratio of total white products (WP, i.e., gasoline, diesel and jet fuel) to the total product yield;
[d] The volumetric ratio of gasoline yield to diesel yield (G/D);
[e] The energy ratio of heavy products yield to the total product yield.

The API gravity is a measure of the density of crude oil. Crude oil that has higher API gravity is less dense (i.e., lighter) and contains higher proportions of small molecules, which refineries can readily process into gasoline, jet fuel and diesel. Most of the sulfur in crude oil must be removed in the refining process, primarily through hydrotreatment to meet product sulfur specifications. Therefore, crude properties such as API gravity and sulfur content have significant effects on the refinery energy use and therefore on refining overall efficiency, because they influence the extent and severity of refinery processing required to meet product volume and quality requirements.

Each refinery is unique in terms of physical configuration at process unit level and operating characteristics. The physical configuration of a refinery can be expressed by complexity index, which is a numerical score that represents the extent, capability, and capital intensity of a given refinery's process units (Nelson, 1976). In general, the higher a refinery's complexity index, the greater the refinery's ability to convert more of the heavy fractions of crude into lighter, high-value products, and to produce light products that meet more stringent quality specifications (e.g., ultra-low sulfur fuels) (Hirshfeld and Kolb, 2012).

2.4 Regression analysis

We conducted both linear and nonlinear multivariate regression analysis to explore the potential correlation between the refinery overall efficiency calculated by using Equation (1) and crude oil properties, refining products slate and the refinery configurations. We employed Stata™ statistical tool package to perform the regression analysis using the LP modeling dataset of the 43 refineries. Equations (3) and (4) show the linear and nonlinear regression models, respectively.

$$\eta_i = c_0 + c_1 \times P_1 + c_2 \times P_2 + ... + c_n \times P_n \tag{3}$$

Where:

η_i is the overall efficiency of refinery i;

P_1, P_2,,, and P_n are the predictor variables (e.g., crude API gravity, S%, etc.);

c_0 is a constant;

and $c_1, c_2, \ldots$ and c_n are the coefficients of predictors P_1, P_2,,, and P_n, respectively.

$$\eta_i = c_0 + c_1 \times TP_1 + c_2 \times TP_2 + ... + c_n \times TP_n \qquad (4)$$

Where:

η_i is the overall efficiency of refinery *i*;

TP_1, TP_2,… and TP_n are transformed forms of predictor variables, e.g., the logarithms of the predictors or the predictor raised to some power;

c_0 is a constant;

and c_1, c_2,… and c_n are the coefficients of predictors TP_1, TP_2, … and TP_n, respectively.

The LHV-based overall refinery efficiencies of the 43 refineries served as data points of the dependent variable, η_i, and the potentially influencing factors, e.g., the crude oil properties, the refining products slate, and complexity index of the refinery configuration, were tested as independent predictors of various types of nonlinear regression models, e.g., the polynomial model, the power model, the logarithmic model, the exponential model, the logistic model, the Gompertz model, etc., as well as the linear regression model. A regression model that is statistically acceptable must have statistically significant P values (i.e., <=0.05) for each individual predictor. For a multivariate regression model, the potential multicollinearity among the individual predictors was tested, and the regression model had a variance inflation factor (VIF) of less than 5 to tolerate minimum possibility of having multicollinearity. Then, the regression model with a highest Pearson coefficient (R-square) that met the acceptance criteria as mentioned before was constructed as a statistically robust prediction model.

3. Results and Discussion

3.1 Refinery overall efficiency based on EIA statistics

Table 6 shows the LHV- and HHV-based overall refining efficiency by PADD based on EIA statistics in 2010

Table 6. Overall efficiencies of U.S. refineries by PADD region in 2010

	PADD region					**U.S. average**
	I	**II**	**III**	**IV**	**V**	
LHV-based efficiencies	**93.6%**	**89.7%**	**89.9%**	**87.8%**	**90.1%**	**90.2%**
HHV-based efficiencies	**94.8%**	**91.2%**	**91.2%**	**89.5%**	**91.5%**	**91.5%**

3.2 Refinery overall efficiency based on LP modeling data

The refinery overall efficiencies by PADD region that are estimated based on LP modeling data are illustrated by Figure 1. On average, the overall efficiencies differ among the PADD regions, with those in PADD I being the highest and those in PADD III being the lowest. This is an example of how crude properties and refinery configurations have influenced the energy use and thus the overall efficiency of refineries. We note that PADD I refineries on average process light and sweet crude (i.e., with low sulfur content), and thus require the simplest refinery configuration among all PADD regions, as shown in Table 4. Moreover, the significant variations in the overall refining efficiency between PADDs (except PADD1 with few modeled refineries) are evident. This is primarily because the refineries within each PADD differ significantly in the crude oil properties, the refinery configurations, and the product slates. To characterize this wide variation in refinery overall efficiencies, we developed a probability distribution function (PDF) to quantify the variation of the point-estimation of the U.S. average refining efficiency. The PDF of the LHV-based overall efficiencies follows a Weibull distribution, with a shape parameter of 29.941, a scale parameter of 0.89897, and a location parameter of 0. Furthermore, compared to the efficiency estimates based on EIA statistics, the efficiencies based on LP modeling data are higher in all PADDs except for PADD I.

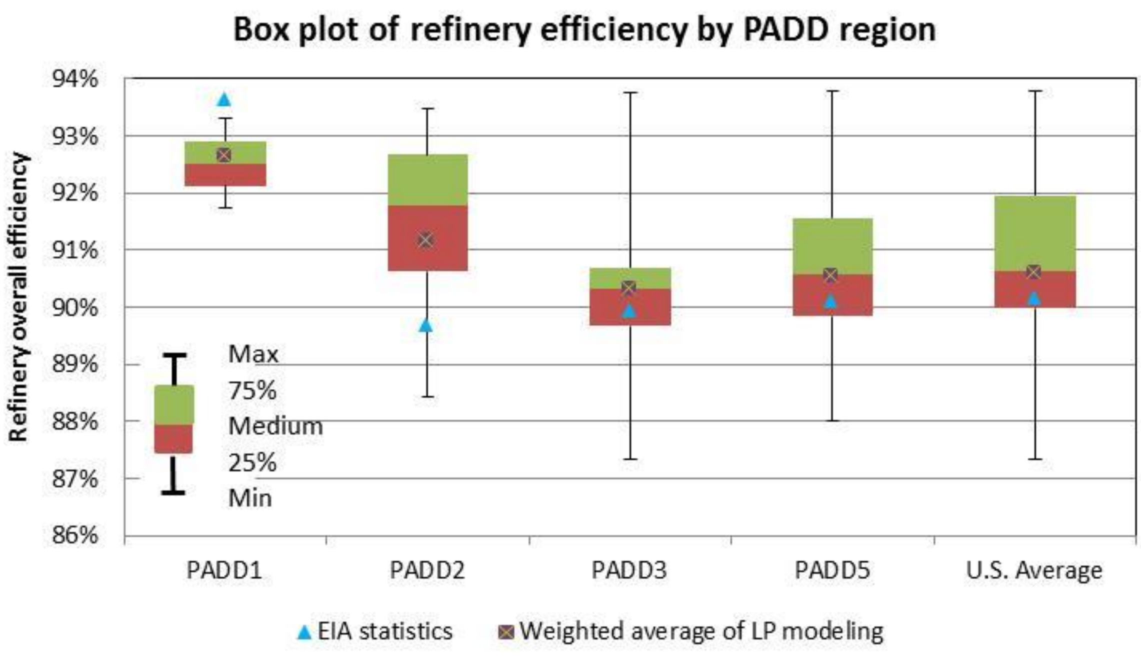

Figure 1. Boxplot of LHV-based refinery overall efficiencies based on LP modeling results by PADD region, in comparison to efficiencies estimates in 2010 based on EIA statistics.

3.3 Correlation between refinery overall efficiency and crude oil properties, refinery complexity index and the product slates

We attempted to construct a statistically acceptable regression model that correlates the refinery overall efficiency with key influencing factors such as crude oil properties, refining products slate, and complexity index of refinery configurations. Of different regression relationships tested, we selected based on regression statistical significance a linear multivariate regression model that employed the API gravity and the sulfur content (wt%) of the crude oil, the heavy products (residual fuel oil, asphalt, coke, slurry oil and reduced crude) yield (mmBtu/mmBtu of total products), and the complexity index of the refinery as the predictors, as shown in Equations (5) and (6) for LHV and HHV basis, respectively.

$$\eta_{LHV} = 0.8759 + 0.002008 \times API - 0.007628 \times S + 0.07874 \times HP - 0.001847 \times CI \qquad (5)$$

$$\eta_{HHV} = 0.8749 + 0.002141 \times API - 0.003388 \times S + 0.07830 \times HP - 0.001676 \times CI \qquad (6)$$

Where:

η_{LHV} and η_{HHV} are the refinery overall efficiency on the LHV or HHV basis;
API is the API gravity of crude oil;
S is the sulfur content (wt%) of crude oil;
HP is the heavy product yield, i.e., the total yield of residual fuel oil, asphalt, coke, slurry oil and reduced crude, in mmBtu of mmBtu of total products;
and CI is the complexity index of the refinery.

The predictors in Equation (5) were all statistically significant at a significant level of 1% ($p<0.01$). Equation (5) shows that overall refining efficiency increases with higher crude API, lower crude sulfur content, higher heavy products yield and lower refinery complexity index. This agrees well with expected cause-effect relationships between refining efficiency and the predictor parameters. We note that the heavy products yield in the regression model along with the complexity index imply the white products yield.

Figure 2 shows that the LHV-based overall efficiencies predicted by the regression model agree well with those calculated based on LP simulations. This is implied by the high R-square value (0.92) of the regression model. Figure 3 illustrates that the typical percentile values, e.g., the minimum, maximum, 25th percentile, median, and 75th percentile values of the predictions by the regression model agree very well with those of the observations reported by EIA. Similarly good prediction of the HHV-based efficiencies is obtained by Equation (6).

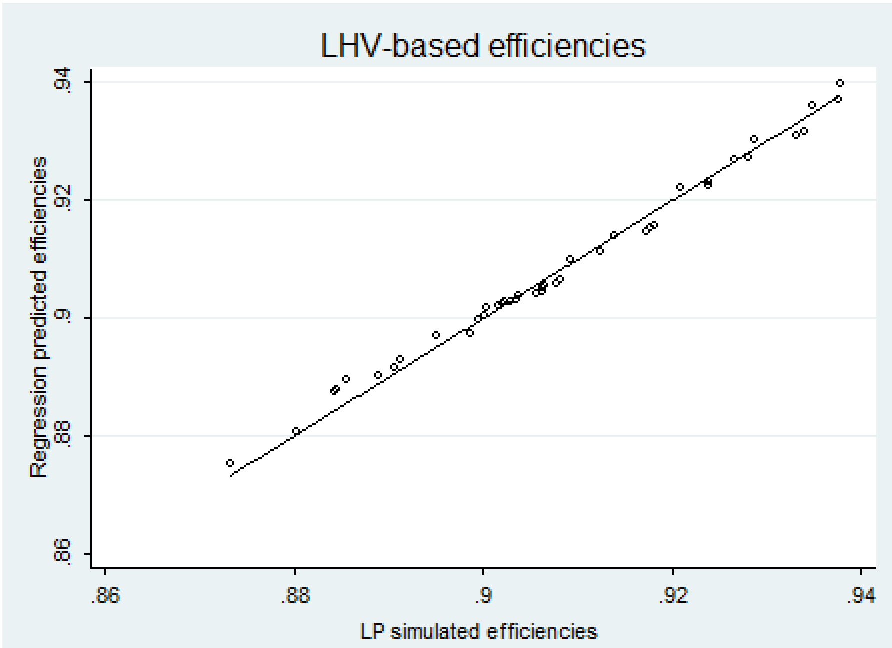

Figure 2. Comparison of the LHV-based overall efficiencies predicted by the regression model to those based on LP simulations.

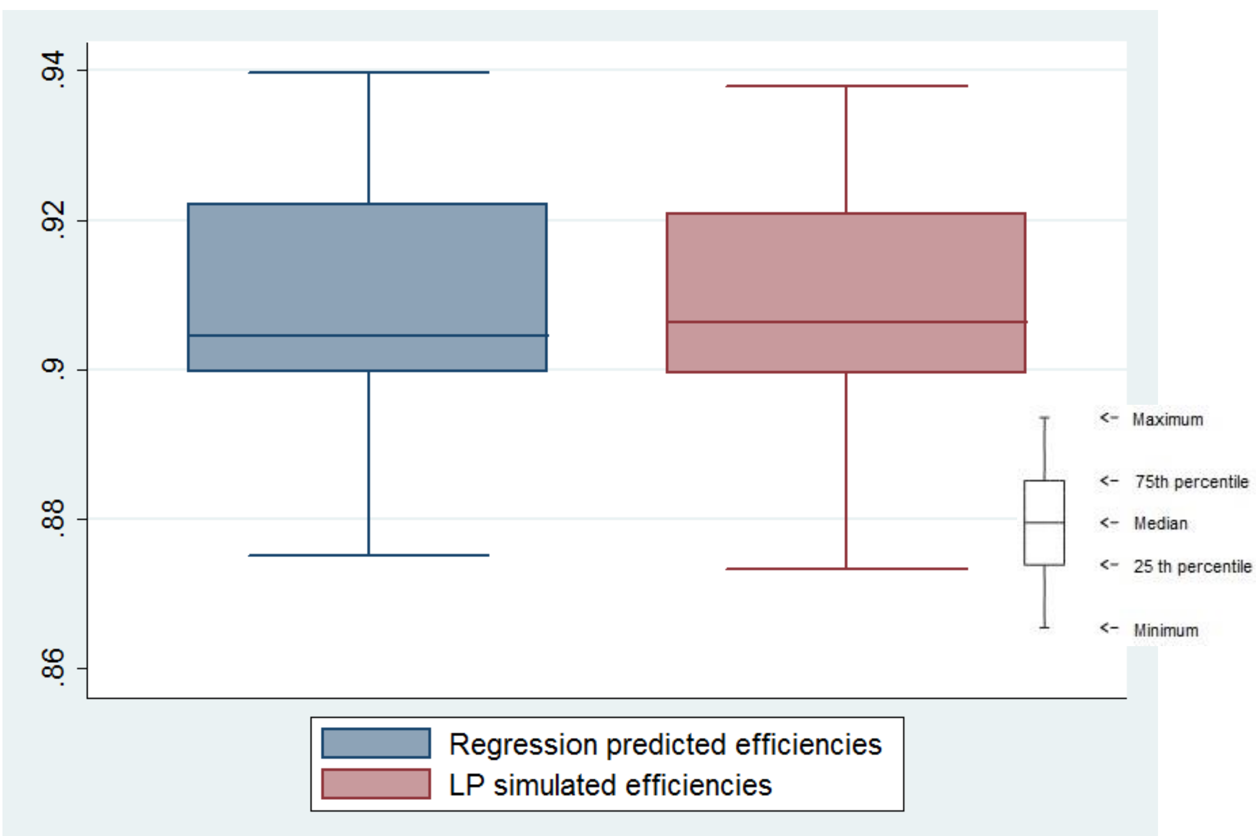

Figure 3. Comparison of the percentile values of the LHV-based overall efficiencies predicted by the regression model to those based on LP simulations

3.4 Projection of refinery overall efficiency

With the linear regression model as depicted in Equation (5), we projected overall refining efficiency from 2011 to 2030, based on the projection of crude oil properties, the heavy products yield and the refinery complexity index.

The API gravity and sulfur content of crudes differ among various crude sources. We projected the API gravity and sulfur content of the U.S. average crudes from 2012 to 2035, as shown in Table A1 of the Appendix. This is done based on the API gravity and sulfur contents of various crude types that are produced domestically in the U.S. (EIA, 2013b) and in many other countries and regions in the world (EIA, 2013f) in 2012, and on the projection of the U.S. crude oil sources in the future (EIA, 2013e). Table A1 shows that there is a decreasing trend in the weighted average API gravity and an increasing trend in the weighted average sulfur content of the crudes in the U.S. from 2012 to 2035. On the other hand, we assume that the heavy products yield and the refinery complexity index will remain the same at the current levels, as shown in Table 5.

The overall refining efficiencies in future years, as predicted by the regression model, were further adjusted to be consistent with the efficiency based on EIA statistics, using Equation (7). This adjustment is intended to make LP simulated efficiencies to be consistent with efficiencies from EIA statistics representing all refineries. With the adjusted projections of the refinery overall efficiencies, the refining product-specific efficiencies are further adjusted using our previously developed methodology (Palou-Rivera et al., 2011). Table 7 shows the adjusted projections of refinery overall efficiencies and the product-specific efficiencies through year 2035.

$$\eta_{adj\_projection,yr} = \frac{\eta_{projection,yr}}{\eta_{\mathrm{Pr}ojection,2010}} \times \eta_{EIA,2010} \tag{7}$$

Where:

$\eta_{adj\_projection,yr}$ is the adjusted projection of the overall efficiency in future year yr;

$\eta_{projection,yr}$ represents the overall efficiency in future year yr projected using Equation (5) or (6);

and $\eta_{\mathrm{Pr}ojection,2010}$ and $\eta_{EIA,2010}$ are the projected overall efficiencies in 2010 using Equation (5) or (6) and the EIA data, respectively.

Table 7. Adjusted projection of refinery overall efficiencies and product-specific efficiencies in the U.S.

	Projected refinery overall efficiency	**Adjusted refinery overall efficiency**	**Gasoline, diesel and LPG**	**Residual oil and naphtha**
2010	91.3%	90.2%	89.3%	95.7%
2011	91.3%	90.2%	89.3%	95.7%
2012	91.3%	90.2%	89.3%	95.7%
2013	91.3%	90.2%	89.3%	95.7%
2014	91.3%	90.2%	89.3%	95.7%
2015	91.3%	90.2%	89.3%	95.7%
2016	91.3%	90.2%	89.3%	95.7%
2017	91.3%	90.2%	89.3%	95.7%
2018	91.3%	90.2%	89.3%	95.7%
2019	91.3%	90.2%	89.3%	95.7%
2020	91.3%	90.2%	89.3%	95.7%
2021	91.3%	90.2%	89.3%	95.7%
2022	91.3%	90.2%	89.3%	95.7%
2023	91.3%	90.1%	89.3%	95.7%
2024	91.3%	90.1%	89.3%	95.7%
2025	91.3%	90.1%	89.3%	95.7%
2026	91.3%	90.1%	89.2%	95.7%
2027	91.3%	90.1%	89.2%	95.7%
2028	91.3%	90.1%	89.2%	95.7%
2029	91.3%	90.1%	89.2%	95.7%
2030	91.3%	90.1%	89.2%	95.7%
2031	91.3%	90.1%	89.2%	95.7%
2032	91.2%	90.1%	89.2%	95.7%
2033	91.2%	90.1%	89.2%	95.7%
2034	91.2%	90.1%	89.2%	95.7%
2035	91.2%	90.1%	89.2%	95.7%

Slight changes in the overall efficiencies are projected due to the variation in crude properties shown in Table A1.

The regression formula should be used with caution to predict the refinery overall efficiency. The changes in crude API, crude sulfur content, refinery heavy product yield and the refinery complexity index should be made within a reasonable range to avoid assuming unrealistic crude inputs, refinery outputs, or refinery configurations for the US refining sector. We provide scenarios for various combinations of these parameters in Table 8.

Table 8. Combination scenarios of the predictors in the regression formula and the estimated efficiencies

	Crude API	Crude S (wt%)	Refinery HP (%)	Refinery CI	Estimated efficiency (%)
A	16	1	22	14	89.2
B	20	3.5	12	13	87.5
C	26	2.5	9	11	89.6
D	28	2	10	12	90.3
E	30	2.5	9	10	90.6
F	34	1	9	9	92.7

For GREET modeling of urban air emissions, we also evaluated the urban shares of refineries on the basis of the locations of refineries and the crude oil process capacities in 2010 (EIA, 2013g), as shown in Table 9.

Table 9. Urban shares of refineries by PADD region in 2010

	PADD 1	PADD 1	PADD 1	PADD 1	PADD 1	U.S. average
Urban share (%)	78.0	48.7	90.4	85.3	78.6	76.4

4 Conclusions

We analyzed the refinery overall efficiency on the basis of the EIA statistics and the LP modeling results provided by Jacobs Consultancy. The refinery overall efficiency estimated based on both datasets showed wide variation between PADD regions and within PADD regions. A further investigation based on the regression analysis showed that the key drivers to the variation in refinery overall efficiencies were variation in crude properties, i.e., API gravity and sulfur content, the heavy products yield, and the refinery complexity index. A statistically significant linear regression model that employs these drivers as predictors and has a strong prediction power was constructed and applied to project the refinery overall efficiencies and the product-specific efficiencies with adjustment. The results are incorporated in GREET1_2013.

Appendix

Table A1. Projected volume-weighted average API gravity and sulfur content (wt%) of crudes to feed the U.S. refineries

	API	S%		API	S%		API	S%
2012	30.62	1.64	**2020**	30.54	1.64	**2028**	30.50	1.68
2013	30.63	1.64	**2021**	30.55	1.64	**2029**	30.48	1.68
2014	30.62	1.64	**2022**	30.54	1.64	**2030**	30.47	1.69
2015	30.61	1.64	**2023**	30.55	1.65	**2031**	30.47	1.69
2016	30.60	1.64	**2024**	30.55	1.65	**2032**	30.46	1.69
2017	30.57	1.63	**2025**	30.56	1.66	**2033**	30.45	1.69

2018	30.57	1.63	**2026**	30.54	1.67	**2034**	30.44	1.69
2019	30.55	1.64	**2027**	30.52	1.68	**2035**	30.43	1.70

References

Argonne National Laboratory (ANL), 2013. Greenhouse gases, Regulated Emissions, and Energy use in Transportation (GREET) model. http://greet.es.anl.gov.

Canadian Association of Petroleum Products (CAPP), 2013. Crude Oil Forecast, Markets and Transportation. http://www.capp.ca/forecast/Pages/default.aspx.

Chemical Economics Handbook (CEH), 2007. Marketing Research Report, Hydrogen.

U.S. Energy Information Administration (EIA), 2012a. Refinery Capacity Report 2012. http://www.eia.gov/petroleum/refinerycapacity/archive/2012/refcap2012.cfm.

EIA, 2012b. Petroleum Supply Annual 2011. http://www.eia.gov/petroleum/supply/annual/volume1.

EIA, 2013a. Number and Capacity of Petroleum Refineries. http://www.eia.gov/dnav/pet/pet_pnp_cap1_dcu_nus_a.htm.

EIA, 2013b. Crude Oil Inputs Qualities. http://www.eia.gov/dnav/pet/pet_pnp_crq_a_EPC0_YCS_pct_m.htm

EIA, 2013c. Downstream Processing of Fresh Feed Input. http://www.eia.gov/dnav/pet/pet_pnp_dwns_dc_nus_mbblpd_m.htm

EIA, 2013d. Refinery & Blender Net Production. http://www.eia.gov/dnav/pet/pet_pnp_refp_dc_r50_mbbl_a.htm.

EIA, 2013e. Annual Energy Outlook. http://www.eia.gov/forecasts/aeo/source_oil_all.cfm.

EIA, 2013f. Company Level Imports. http://www.eia.gov/petroleum/imports/companylevel.

EIA, 2013g. Refinery Capacity Report 2012, Table 6. Refiners' Total Operable Atmospheric Crude Oil Distillation Capacity. http://www.eia.gov/petroleum/refinerycapacity.

Palou-Rivera, I., Han, J. and Wang, M., 2011. Updated Estimation of Energy Efficiencies of U.S. Petroleum Refineries. http://greet.es.anl.gov/publication-petroleum.

The U.S. Environmental Protection Agency (EPA), 2013a. Gasoline. http://www.epa.gov/otaq/fuels/gasolinefuels/index.htm.

The U.S. Environmental Protection Agency (EPA), 2013b. Diesel. http://www.epa.gov/otaq/fuels/dieselfuels/index.htm

Nelson, W.L., 1976. Nelson Complexity Index. Oil & Gas Journal. 13, 20, and 27.

Hirshfeld, D.S. and Kolb, J.A., 2012. Analysis of energy use and CO_2 emissions in the U.S. refining sector, with projections for 2025. Environmental Science & Technology 46, 3697-3704.