

Uncertainty Analysis of Life Cycle Greenhouse Gas Emissions from Petroleum-Based Fuels and Impacts on Low Carbon Fuel Policies

ARANYA VENKATESH,[†]
PAULINA JARAMILLO,[‡]
W. MICHAEL GRIFFIN,^{*,†,§} AND
H. SCOTT MATTHEWS^{†,‡}

*Civil and Environmental Engineering Department,
Department of Engineering and Public Policy, and Tepper
School of Business, Carnegie Mellon University, 5000 Forbes
Avenue, Pittsburgh, Pennsylvania 15213-3890, United States*

*Received July 22, 2010. Revised manuscript received
September 29, 2010. Accepted October 1, 2010.*

The climate change impacts of U.S. petroleum-based fuels consumption have contributed to the development of legislation supporting the introduction of low carbon alternatives, such as biofuels. However, the potential greenhouse gas (GHG) emissions reductions estimated for these policies using life cycle assessment methods are predominantly based on deterministic approaches that do not account for any uncertainty in outcomes. This may lead to unreliable and expensive decision making. In this study, the uncertainty in life cycle GHG emissions associated with petroleum-based fuels consumed in the U.S. is determined using a process-based framework and statistical modeling methods. Probability distributions fitted to available data were used to represent uncertain parameters in the life cycle model. Where data were not readily available, a partial least-squares (PLS) regression model based on existing data was developed. This was used in conjunction with probability mixture models to select appropriate distributions for specific life cycle stages. Finally, a Monte Carlo simulation was performed to generate sample output distributions. As an example of results from using these methods, the uncertainty range in life cycle GHG emissions from gasoline was shown to be 13%—higher than the typical 10% minimum emissions reductions targets specified by low carbon fuel policies.

Introduction

The Annual Energy Review 2008 (1) published by the Energy Information Administration (EIA) reports that the U.S. transportation sector consumes about 30% of all primary energy, of which 95% is petroleum-based. These statistics underlie two significant concerns: (a) national security, given that the U.S. imports 60% of its crude oil, and (b) associated greenhouse gas (GHG) emissions of combusting fossil fuels. These concerns have been reflected in recent legislation such as the Energy Independence and Security Act (EISA) (2) and California's Low Carbon Fuel Standard (LCFS) (3). Applying

environmental life cycle assessment (LCA) as a tool, both policies estimate emissions reductions possible with alternative fuels. The LCFS calls for a minimum 10% reduction in average carbon intensity of these fuels through substitutions in the fuel mix sold. Similar reductions are expected from EISA.

Such policies predominantly implement emissions reduction targets based on previously found single point estimates of GHG emissions from conventional and alternative fuels. This deterministic approach fails to acknowledge the uncertainty and variability (henceforth collectively referred to as uncertainty, as explained by Lloyd and Ries (4)) in the data and methods of LCA. The policies thus overestimate the reliability of the point estimate results, such that consequent decisions made may prove to be costly and misleading (5). Well-known earlier analyses estimating GHG emissions from petroleum-based fuels in the U.S. include the Greenhouse gas, Regulated Emissions and Energy use in Transportation (GREET) model developed by the Argonne National Laboratory (6) and the National Energy Technology Laboratory (NETL) baseline report on emissions from petroleum-based fuels (7). Results from modified versions of the GREET model have been used to support both the LCFS and the EISA. Although the results from these models are consistent, partially due to the similarity in data sources used, these tools are predominantly deterministic in nature—detailed uncertainty modeling of the emissions from petroleum-based fuels has not been carried out previously. Thus these standards do not incorporate the range of uncertainty in estimating baseline emissions for fossil fuels, nor the proposed alternatives.

Methods to deal with uncertainty in LCA quantitatively, as suggested by many previous studies and summarized by Lloyd and Ries (4) and Williams et al. (8) include probabilistic simulation, intervals, scenario modeling, fuzzy data sets, and analytical uncertainty propagation. While dealing with parameter uncertainty, due to incomplete knowledge of the true value of any LCA model input, probability distributions are specified using data or expert judgment, and simulation methods and uncertainty importance analyses are commonly used to establish uncertainty in output (9). Comparisons between conventional and alternative fuels that account for uncertainty are made possible using results from these methods. In LCA, a system could have either a few or a large number of inputs depending on the boundary considered. Using analytical methods could be computationally intensive for systems with a large number of input variables, such as those considered in this study, and hence simulation methods are usually preferred.

In this study, an uncertainty analysis of life cycle GHG emissions from petroleum-based conventional fuels consumed in the United States was developed using statistical modeling methods. Inventory data parameters were represented by probabilistic distributions encompassing both parameter uncertainty (inputs that are truly uncertain) as well as temporal (inputs vary with time) and spatial (inputs vary with region) variability. Predictions from a partial least-squares (PLS) regression model were used when data were unavailable. Stochastic modeling was performed using Monte Carlo simulation, and confidence intervals were derived from resulting sample output distributions, along with an uncertainty importance analysis. A method for comparison of results with similar output distributions for alternative fuels, to be incorporated within a climate policy framework, is proposed.

* Corresponding author phone: (412) 268-2299; fax: (412) 268-7357; e-mail: mwg@andrew.cmu.edu.

[†] Civil and Environmental Engineering Department.

[‡] Department of Engineering and Public Policy.

[§] Tepper School of Business.

Methods

A streamlined process-based life cycle of petroleum-based fuels is considered in this study, which begins with crude oil extraction and ends with combustion. The extraction of domestic and imported crude oil releases GHG emissions due to process fuel consumption and fugitive emissions. The crude oil is then transported to refineries by pipeline and ocean tanker. At refineries, fuel consumption and captive and merchant hydrogen production required for outputs such as gasoline, diesel, residual fuel oil, still gas, and liquefied petroleum gas (LPG) contribute to significant GHG emissions. Next, the refined products are transported from refineries to consumers by barge, pipeline, rail, and truck, and are finally combusted in vehicles and/or other equipment.

The functional unit considered in this study is 1 MJ of petroleum-based fuel. Life cycle GHG emissions factors of all fuels used were obtained from literature or developed within the study. Wherever possible, monthly data for the years between 2005 and 2009 were used as the temporal basis. As suggested by Huijbregts (10), the mean values of inputs and associated uncertainty were estimated to characterize any temporal variability in the life cycle inventory data. When sufficient data were available, fitting data to distribution functions was carried out through the Decision Tools Suite @Risk add-in tool for Excel that reports the top best-fitting distributions using the chi-squared test statistic. A Monte Carlo simulation with 10,000 iterations was performed to estimate the range of resulting emissions from fuels produced. Probability mixture models were used to represent life cycle GHG emissions factors for fuels used within the United States.

Crude Oil Extraction and Transport. Data for GHG emissions from crude oil extraction were obtained from the Ecoinvent database in the SimaPro 7.1 LCA software (11), the NETL Petroleum-Based Fuels Life Cycle Greenhouse Gas Analysis 2005 Baseline Model (12) that contains data from the GaBi LCA software (13), and an article by McCann and Magee (14). The data along with an associated discussion on their applicability is presented in the Supporting Information. Using these data points, the minimum, maximum, and most likely (estimated mode) values were used as parameters of triangular distributions to represent the GHG emissions from crude oil extraction for each region from which the U.S. imports.

Emissions from transport of crude oil were estimated for each of the exporting and domestic regions. Because these emissions were correlated with regional emissions from crude extraction, they were added to the latter for subsequent simulations. Fuel consumption data for transporting crude oil were obtained from the GREET model, which was used in conjunction with appropriate transport distances to obtain total energy use and associated emissions per unit mass of crude oil transported. A probability mixture model, explained in subsequent sections, was used to represent extraction and transport emissions from domestic and imported crude oil consumed in the U.S. by selecting emissions from each of these world regions.

Crude Oil Refining. Petroleum refineries are massively complex process-based systems that synthesize a number of products while utilizing large amounts of energy and hydrogen for hydrotreatment processes. The annual energy use at refineries in the U.S. is given by EIA (15) at the Petroleum Administration Defense District (PADD) level. Quantities of refinery fuel inputs used such as LPG, distillate fuel, petroleum coke, residual fuel oil, and still gas are reported, along with purchased inputs such as coal, natural gas, and electricity. Impacts from producing and consuming captive (on-site) and merchant (off-site) hydrogen by steam methane reforming of natural gas are not included in EIA

data and were analyzed separately. Other contributions to refining emissions such as venting and flaring emissions on-site were not expected to be significant as demonstrated by previous deterministic analyses (7) and were not included but would further contribute to uncertainty.

Energy consumption data at refineries were collected by PADD for the period 1999–2008. Additional impacts from captive hydrogen production include process emissions from the reforming of natural gas and the upstream emissions of natural gas used as feedstock that were estimated by methods described in the Supporting Information.

At the most disaggregated level, the total annual GHG emissions at refineries due to energy and hydrogen use were estimated using available data. However, it is expected that these emissions vary on a monthly basis, due to changes in quality of input crude as well as the demand for products (for example, heating oil demand peaks during the winter months). Hence a multiple linear regression model was developed to predict monthly energy and hydrogen use GHG emissions at the refinery level.

The total annual emissions including fuel and hydrogen use at refineries were normalized by the volume of crude oil input to obtain emissions per barrel of crude oil, for refineries in the five PADDs over a period of 10 years—a total of 50 data points. The refining emissions per barrel of crude oil input to the refineries were expected to vary predominantly with the quality of the crude oil—with heavier, sourer crude oil using more energy. A preliminary analysis of the API gravity of crude oil consumed showed that lower density (lighter) crude is used in refineries in Group 1 (PADDs 1, 2, and 4) compared to the heavier crude in Group 2 (PADDs 3 and 5), as presented in the Supporting Information. Therefore, separate emissions factors were developed for fuels produced in Group 1 and Group 2 using probability distributions and regression results to represent critical input variables.

For the regression analysis, predictor variables other than crude density and sulfur content were included to improve the variance in the response variable (emissions per barrel of crude) explained by the model. It was proposed that differences in the refinery product slate could influence energy use (and therefore emissions) at refineries. Therefore, the predictor variables added include the mass fractions of the refinery products such as gasoline, diesel, residual fuel oil and others—a total of 12 inputs to a multiple linear regression model. However, the predictor variables used for this analysis were found to be correlated indicating that the interpretation of results obtained could be affected by collinearity (16–18). A solution to this problem using a simple least-squares approach could not be easily estimated due to the potential near-singularity in the input correlation matrix leading to large variances in regression coefficients (19, 20).

A number of solutions have been suggested previously in literature: ridge regression, stepwise regression, principal components regression, and partial least-squares regressions, to name a few (21). Partial least-squares (PLS) regression is commonly used when data in regression have a large number of (sometimes correlated) inputs. It produces linear combinations of original inputs to the system, which are subsequently used as new input components to the regression (21). Previous studies suggest the superiority of this technique and hence it was used for this analysis (22–24). Further information on this algorithm is presented in the Supporting Information.

Model selection methods for choosing the optimum number of input components in PLS regression are available, as summarized in Li et al. (25). They propose that the choice of method for model selection depends strongly on the purpose of the model and suggest that the Akaike Information Criterion (AIC) may be the preferred method for models aimed at prediction, and hence it is used in this study. The

model that minimizes the AIC value is selected as best among those considered, i.e., this provides the best-fitting model, with the optimum number of components. Further information on this method and its application is described in the Supporting Information.

PLS regression was used to fit a model describing relationships between the predictor and response variables. The predictor variables used were standardized (z scores) and AIC was used to select the optimum number of components for the model. The regression coefficients from the PLS model were then used in conjunction with U.S. monthly crude quality and refinery output mass fractions estimated from EIA refinery yield data (26) to predict corresponding emissions per barrel of crude oil refined.

Product Transport. This life cycle stage was developed only for gasoline and diesel fuel being major refinery outputs, contributing to nearly 70% of production. These assumptions were not anticipated to have a significant impact on the final results of the other refinery fuels, since this stage was expected to contribute to less than 1% of life cycle emissions, as reported earlier by GREET and NETL models. The transport distances for gasoline and diesel, as well as the percent of fuel transported by barge, pipeline, ocean tanker, rail, and truck were obtained from the GREET model (6). Their energy intensities were primarily obtained from the Transportation Energy Data Book (27) and GREET. The percent of imported finished fuel per unit volume of finished fuel used within the U.S. was obtained from EIA's imports data (28) and assumed to be transported by ocean tanker. The data and methods are presented in the Supporting Information.

Combustion. The average, minimum, and maximum values of emissions from combustion of petroleum fuels were obtained from the IPCC 2006 Good Practices guide (29). A summary of these values, used as the characteristic parameters of triangular distributions, is presented in the Supporting Information. Note that the range of combustion emissions values is considerably smaller than other life cycle stages, as shown in the results section.

Allocation. The crude extraction and transport, and refining emissions were previously developed per unit barrel of crude, which were then allocated to each of the refinery products. Since the product transport and combustion stages were developed separately for each output, no allocation was required for these stages.

The crude oil extraction emissions per barrel were allocated to the refinery products on a mass basis. EIA refinery yields (26) are a measure of the gross production from refineries, rather than the net production volumes reported by EIA. Assuming that refinery yields represent volumetric fractions of the refinery output slate, they were converted to mass fractions using average petroleum product density data from EIA (30), fitted to distributions and used for allocation. The resulting emissions per barrel of crude allocated to each product were then divided by its volumetric fraction, represented by a distribution, to obtain the emissions associated per unit volume of product. Because the mass and volume fractions of products are entirely correlated, simulations were carried out assuming a correlation coefficient of 1 between the two input variables. Crude oil transport emissions were estimated per unit mass of crude oil. Assuming a mass balance at the refinery stage, a unit mass of crude oil was estimated to produce an equivalent mass of petroleum product. Hence the emissions per unit mass of crude were assumed equivalent to emissions per unit mass of product. A distribution was fitted to the range of values obtained for allocated emissions and used to represent GHG emissions from the crude extraction and transport stage.

The refining emissions per barrel of crude obtained from the crude oil refining regression model were then allocated

to the refinery outputs such as diesel and gasoline based on their energy content. The emissions per barrel of crude allocated to each product were then divided by its volumetric fraction to obtain the emissions associated per unit volume of product, as performed for the crude oil extraction phase. A distribution was fitted to the range of values obtained for allocated emissions and used to represent GHG emissions from the crude refining stage.

Probability Mixture Models. Using methods illustrated in the previous sections, GHG emissions factors for extraction and transport were developed for crude oil obtained from different world regions for both refinery groups. Within each refinery group, the probability that the given fuel is obtained from crude belonging to a particular world region is given by the fraction of crude oil obtained from that region. The resulting probability distribution function for crude extraction and transport emissions produced in a refinery group is a convex combination of the probability distribution functions of extraction and transport emissions of crude oil from different world regions. It can therefore be described by a probability mixture model that is the weighted sum of its constituent probability distribution functions, rather than by a weighted distribution of the absolute values.

The fractional contributions of crude from different world regions were estimated using quantities of domestic and imported crude oil consumed between 2005 and 2009 obtained from EIA (31, 32), as explained in the Supporting Information. The emissions estimated for crude oil extraction and transport from the mixture model were added to emissions from other fuel life cycle stages, obtained as described previously, to estimate life cycle emissions factors for fuels produced in each refinery group. The probability that petroleum fuels consumed in the U.S. are obtained from refineries in each group was assumed to be the fraction of crude oil consumed in these groups. Therefore, an additional probability mixture model was used to represent life cycle GHG emissions from fuels consumed in the U.S., henceforth referred to as the base scenario.

Results

The mean values of emissions factors for petroleum-based fuels developed in this study, along with 90% confidence intervals represented by error bars, are presented in Figure 1. The mean values were found to be 89 g CO₂e/MJ of gasoline and 93 g CO₂e/MJ of diesel. Of the total for gasoline, nearly 80% consists of combustion emissions, with contributions of 9% from crude extraction and transport, 11% from refining, and 1% from product transport, with similar contributions for the other fuels. The mean emissions from the lighter refinery fuels, still gas and LPG, and the heavier refinery products, diesel and residual fuel oil, were found to be within 5% of the mean emissions from gasoline, while petroleum coke was found to have significantly higher life cycle emissions due to its higher carbon content. Comparisons with life cycle emissions factors for gasoline and diesel developed by GREET, NETL, those incorporated in the California Low Carbon Fuel Standard (from a California-GREET model, as presented in a California Air Resources Board report (33)), and those incorporated in the EISA (also modified from the GREET model) (34), show similar results. The consistency of these results does not imply the absence of uncertainty in the results, for example, due to variations in the operating conditions such as input crude quality, refinery fuels used, refinery output slate, and time of year. However, the error bars that represent the 90% confidence intervals of simulation outputs indicate that the uncertainties in the results, previously ignored by most LCA studies, are indeed non-negligible.

The cumulative distribution function for life cycle emissions from gasoline consumed in the United States (base

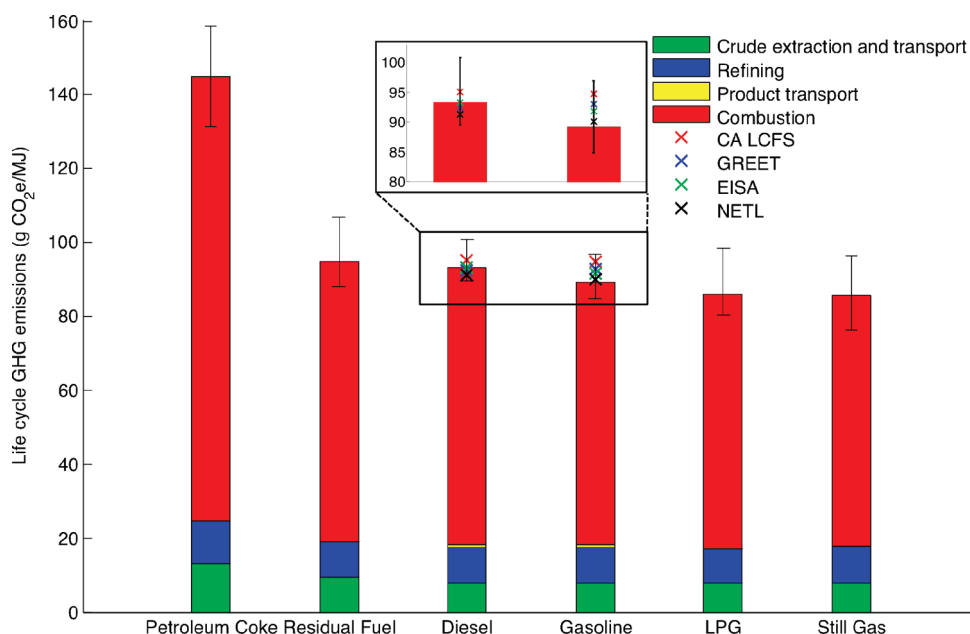


FIGURE 1. Comparison of life cycle GHG emissions factors for petroleum-based fuels consumed within the U.S. obtained from this study and versus existing literature.

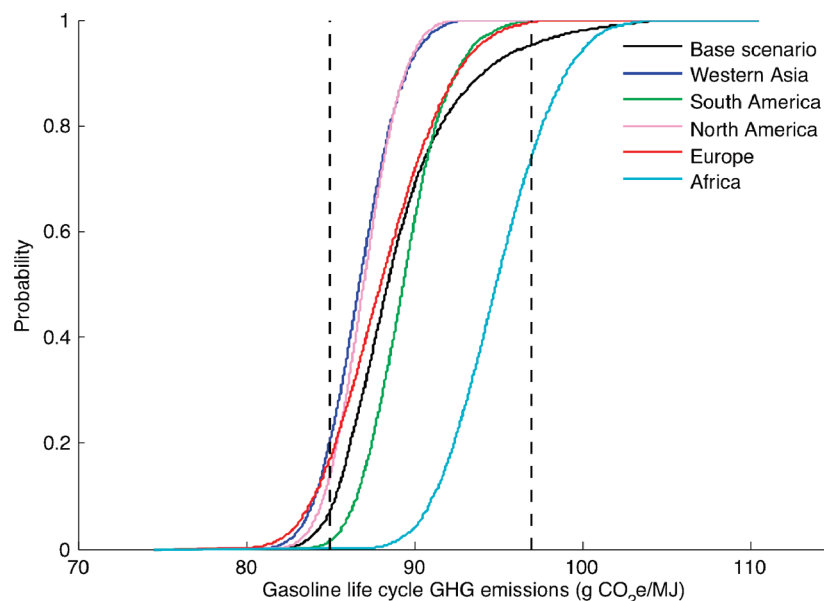


FIGURE 2. Cumulative distribution functions with a comparison of life cycle GHG emissions of gasoline produced from crude oil obtained from different regions with base scenario; 90% confidence interval for base scenario represented by dashed lines.

scenario) is presented in Figure 2, along with the 90% confidence interval represented by dashed lines. The life cycle GHG emissions factors for gasoline produced from crude oil obtained from different world regions are also compared with base scenario results. Significant differences in life cycle emissions of gasoline produced from different sources of crude oil were thus observed, with gasoline from African crude oil having the highest emissions, with mean emissions nearly 7% higher than those of the base scenario.

Simulation results for upstream and life cycle GHG emissions factors of gasoline produced in the two refinery groups are presented in the Supporting Information. A fairly significant overlap of the probability distribution functions for gasoline upstream emissions in the two refinery groups was observed. Similarly, only small differences in mean values and the shape of the probability distribution functions for the life cycle GHG emission factors of gasoline produced in the two refinery groups was observed, indicating that

differences in crude extraction, transport, and refining emissions in different refinery groups do not affect overall results significantly. The probability distribution functions for emissions from life cycle of gasoline and diesel consumed in the U.S. are presented in the Supporting Information.

The 90% confidence interval for gasoline life cycle emissions was found to lie between 85 and 97 g CO₂e/MJ, a range that spans about 13% of the mean value, as shown in Table 1. The sample results were best fitted by a log–logistic distribution function. The distribution types and parameters that were found to best represent the range of life cycle greenhouse gas emissions from the refinery products are presented in the Supporting Information. These uncertainty ranges could be used in other research or incorporated into existing life cycle inventory models to produce more robust analysis tools.

The coefficient of variation (COV), defined as the ratio of the standard deviation of a probability distribution to its

TABLE 1. Summary Statistics for the GHG Emissions from the Life Cycle Stages of Gasoline (The Coefficient of Variation Is the Standard Deviation over the Mean, whereas the Lower and Upper Bounds of the 90% Confidence Interval Are Expressed As a Percentage of the Mean Value)

life cycle stage	mean	standard deviation	COV	90% CI	
	(g CO ₂ e/MJ)	(g CO ₂ e/MJ)		lower bound	upper bound
crude extraction + transport	7.7	3.3	0.43	-36%	93%
refining	9.8	1.3	0.13	-21%	21%
product transport	0.7	0.1	0.08	-6%	11%
combustion	71.0	1.3	0.02	-3%	3%
life cycle	89.3	3.8	0.04	-5%	8%

TABLE 2. Results from Uncertainty Importance Analysis for Gasoline and Diesel

parameter	gasoline		diesel	
	Spearman's rank correlation coefficients	contribution to variance	Spearman's rank correlation coefficients	contribution to variance
combustion	0.42	55%	0.23	23%
refining - group 2	0.26	21%	0.31	41%
weight - crude mixture model	0.18	9%	0.15	10%
refining - group 1	0.15	7%	0.16	11%
weight - refinery mixture model	0.07	2%	0.07	2%

mean, is a normalized indicator of the spread of the distribution. The COV for each of the life cycle stages of the petroleum fuels, based on gasoline simulation results, is also presented in Table 1. It is observed that uncertainty in upstream emissions from gasoline is much higher, when compared to combustion emissions, as indicated by the COV and the 90% confidence intervals for these stages.

An uncertainty importance analysis for the life cycle GHG emissions was performed and the results for gasoline and diesel are shown in Table 2, where the Spearman's rank correlation coefficients between output emissions and top five inputs and their contribution to variance (expressed as the normalized squares of the correlation coefficients) are presented. Although the emissions from combustion contribute to nearly 80% of mean value of life cycle emissions of gasoline, they only contribute to 55% of the variance in results. Refining emissions contribute to nearly 30% of variance in results, as compared to about 11% of the mean. The emissions from the combustion of diesel were reported to have a narrower distribution when compared to gasoline. Although the absolute variance in emissions from refining gasoline and diesel were found to be the same, the total variance in life cycle emissions from diesel was smaller than that of gasoline, and hence the percent contribution of refining was higher. The weights of the mixture model selecting emissions between refinery groups is not observed to be critical, as expected, because impacts from both groups are similar, as discussed.

For the model described in previous sections, the probability distributions chosen to represent the uncertainty in input parameters were mainly selected by data-fitting. However, for crude extraction and combustion emissions, distribution types were selected subjectively (triangular distributions) due to limited readily available data. A comparative analysis using uniform distributions for both parameters was therefore performed to observe the sensitivity of results to this choice. When the crude extraction emissions were assumed to be uniformly distributed, the mean GHG emissions from gasoline increased by only 0.9%, while the 90% confidence interval increased to 17% of the mean value. When both crude extraction and combustion emissions were assumed to be uniformly distributed, the mean GHG emissions increased by 1.4% while the 90% confidence interval remained 17% of the mean value.

Much of the emissions from crude oil extraction in African countries are due to poor practices with high flaring and venting of gases. If the emissions intensity of these extraction processes were reduced to North American levels, a reduction in the life cycle GHG emissions of gasoline consumed in the U.S. could be expected. Simulation results indicate that the average expected life cycle GHG reductions for gasoline are 1.6%. If the emissions intensity of extraction processes in North America were reduced to West Asian levels, a reduction of about 1% in the life cycle emissions of gasoline could be expected. Although this does not seem significant compared to the mean reduction targets, they do offer an alternate, more certain, and potentially less expensive method of helping achieve the emissions reduction targets, along with those commonly considered in policy (such as the introduction of biofuels).

Discussion

Given that the uncertainty in life cycle GHG emissions for gasoline, as estimated by this study, spans 13% of its mean value, it is interesting to note that this range is higher than the emissions reductions target of 10% proposed by the California Low Carbon Fuel Standard (3). Put another way, it appears that if a fixed emissions reductions target were applied, using the highest value of emissions of gasoline on the right tail of the distribution as the baseline, existing gasoline at the left tail of the distribution would already meet reduction targets. The success of the policy would then depend on which point along the distribution the chosen deterministic target lies. Hence it seems infeasible to design climate policy by assigning reduction targets using single point estimates as baselines, without incorporating an understanding of the uncertainty that generated the point estimates, and expect any noticeable change from existing emissions levels.

Previous analyses using deterministic emissions estimates to design policy targets dismiss its inherent uncertainty as either negligible or as difficult to deal with and hence largely tend to ignore its implications, although the data used in this analysis have been readily available. Instead of using a single value for the standards and a somewhat arbitrary percent reduction target, sound decision-making would call for at least an acknowledgment of the range of existing values.

Developing similar probability distributions for alternative fuels allows their comparison with baseline petroleum fuel emissions (35). Given fixed percent reduction targets based on emissions values of petroleum-based fuels, subjectively chosen from the range (for instance, the fifth percentile or the mean value), the probability that the emissions of alternative fuels lie below the reduction target can be estimated and used to determine preferences of fuel type. The percentage reduction targets can subsequently be varied to determine the level of confidence in achievable emissions reductions. Thus both, a target and probable emissions reductions, may be selected analytically. This may be followed by a concerted effort toward its increasing the penetrability of the alternative fuel, as determined within the framework of a climate policy.

Although uncertainty and variability were labeled as uncertainty earlier, their differences and relevance to this analysis are highlighted. For instance, much of the range of emissions values used for crude oil extraction in each world region is uncertain (given that different databases report different values for the same processing stage). On the other hand, for the refining life cycle stage, spatial and temporal variability characterize the range of values obtained, as emissions were considered to change with time period as well as region of processing, depending on factors such as crude type used and product slate mixes. However, policymakers currently have no feasible way of tracking where or when a particular gallon of fuel was refined and from that viewpoint this range could be considered uncertain. Although the methods for dealing with both include probabilistic simulation, the range of emissions obtained as a result of variability could potentially be improved by conducting appropriate piece-wise analyses. The bottom line is that there is both significant uncertainty and variability in estimates such as these and any decision-maker selecting numbers for policy should have an understanding of both.

Finally, as discussed by Lloyd and Ries (4), the variance in results as obtained from most life cycle assessments, including this study, is an underestimate. This is because the range of values used to describe parameter uncertainty or variability is often based on average or best estimates reported in literature and databases that do not necessarily reflect the true range of possible values that are mostly unreported. This is especially observed in the crude oil extraction emissions estimates and average crude quality indicators for refining developed and used within this study. Therefore, the range of life cycle emissions from gasoline and other petroleum-based fuels estimated in this study should be used as an indicator of the narrowest possible range of uncertainty in these impacts and in reality, a wider range may be expected.

Modeling environmental impacts over the life cycle of products and processes has much inherent uncertainty, due to the nature of available data, assumptions, temporal and geographical inconsistencies, to cite a few reasons. The development and application of LCA results at a small scale, such as within organizations or institutions, reduces some uncertainty, primarily due to the availability and control of pertinent data. However, caution must be exercised when using such results at larger scales, for example, in policy design at the state and federal level.

This study proposes a method to deal with this by using statistical methods to characterize the potential range of life cycle GHG emissions of petroleum-based transportation fuels. Regional variations in impacts were considered, although no significant differences for petroleum fuels produced in different refinery groups were found. Comparisons with alternative fuel types may be achieved by indicating how likely it will be for one fuel type to be preferred over another by means of probability values, given fixed emissions reductions targets. Such analyses are essential to

design any effective alternative fuels policy to reduce GHG emissions from transportation.

Acknowledgments

This material is based upon work supported by the U.S. National Science Foundation (grant 0628084), the Energy Foundation, and the Steinbrenner Graduate Fellowship Program. Any opinions, findings, and conclusions expressed in this material are those of the authors and do not necessarily reflect the views of these organizations.

Supporting Information Available

Detailed explanation of methods employed with intermediate and supplementary results, and references. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Literature Cited

- (1) U.S. Department of Energy *Annual Energy Review 2008*; Energy Information Administration: Washington, DC, 2009; <http://www.eia.doe.gov/aer/pdf/aer.pdf>.
- (2) Energy Independence and Security Act of 2007. Public Law 110-140, 2007; http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=110_cong_public_laws&docid=f:publ140.110.pdf.
- (3) Schwarzenegger, A. *Executive Order S-01-07: Low Carbon Fuel Standard*; State of California: Sacramento, CA, 2007.
- (4) Lloyd, S. M.; Ries, R. Characterizing, Propagating, and Analyzing Uncertainty in Life-Cycle Assessment. *J. Ind. Ecol.* **2007**, *11*, 161–179.
- (5) Björklund, A. Survey of Approaches to Improve Reliability in LCA. *Int. J. Life Cycle Assess.* **2002**, *7*, 64–72.
- (6) Wang, M. *The Greenhouse Gases, Regulated Emissions, and Energy Use in Transportation (GREET) Model*; Argonne National Laboratory: Argonne, IL, 2009; <http://greet.es.anl.gov/main>.
- (7) Gerdes, K.; Skone, T. J. Development of Baseline Data and Analysis of Life Cycle Greenhouse Gas Emissions of Petroleum-Based Fuels; National Energy Technology Laboratory: Pittsburgh, PA, 2008.
- (8) Williams, E. D.; Weber, C. L.; Hawkins, T. R. Hybrid Framework for Managing Uncertainty in Life Cycle Inventories. *J. Ind. Ecol.* **2009**, *13*, 928–944.
- (9) Huijbregts, M. A.; Gilijamse, W.; Ragas, A. M.; Reijnders, L. Evaluating Uncertainty in Environmental Life-Cycle Assessment. A Case Study Comparing Two Insulation Options for a Dutch One-Family Dwelling. *Environ. Sci. Technol.* **2003**, *37*, 2600–2608.
- (10) Huijbregts, M. Application of Uncertainty and Variability in LCA. *Int. J. Life Cycle Assess.* **1998**, *3*, 273–280.
- (11) PRé Consultants SimaPro Version 7.1, Amersfoort, The Netherlands, 2007; <http://www.pre.nl/>.
- (12) U.S. Department of Energy. *NETL Petroleum-Based Fuels Life Cycle Greenhouse Gas Analysis 2005 Baseline Model*; National Energy Technology Laboratory, 2009; www.netl.doe.gov/energy-analyses.
- (13) PE International GmbH. GaBi 4.0 Life Cycle Assessment Software - Professional Database; Germany, 2007; <http://www.pe-international.com/index.php?id=152&L=6&redirect=1>.
- (14) McCann, T.; Magee, P. Crude Oil Greenhouse Gas Life Cycle Analysis Helps Assign Values for CO₂ Emissions Trading. *Oil Gas J.* **1999**, *97*, 38–44.
- (15) U.S. Department of Energy. *Petroleum Navigator: Fuel Consumed at Refineries*; Energy Information Administration, 2010; http://tonto.eia.doe.gov/dnav/pet/pet_pnp_capfuel_dcu_nus_a.htm.
- (16) Graham, M. H. Confronting Multicollinearity in Ecological Multiple Regression. *Ecology* **2003**, *84*, 2809–2815.
- (17) Blalock, H. M. Correlated Independent Variables: The Problem of Multicollinearity. *Social Forces* **1963**, *42*, 233–237.
- (18) Mason, C. H.; Perreault, W. D. Collinearity, Power, and Interpretation of Multiple Regression Analysis. *J. Marketing Res.* **1991**, *28*, 268–280.
- (19) Farrar, D. E.; Glauber, R. R. Multicollinearity in Regression Analysis: The Problem Revisited. *Rev. Econ. Stat.* **1967**, *49*, 92–107.
- (20) Stewart, G. W. Collinearity and Least Squares Regression. *Stat. Sci.* **1987**, *2*, 68–84.

- (21) Hastie, T. ; Tibshirani, R. ; Friedman, J. H. *The Elements of Statistical Learning*; Springer: New York, 2001; p 560.
- (22) Wold, S.; Ruhe, A.; Wold, H.; Dunn, W. J., III. The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses. *SIAM J. Sci. Stat. Comput.* **1984**, 5, 735–743.
- (23) Wold, S. Discussion: PLS in Chemical Practice. *Technometrics* **1993**, 35, 136–139.
- (24) Garthwaite, P. H. An Interpretation of Partial Least Squares. *J. Am. Stat. Assoc.* **1994**, 89, 122–127.
- (25) Li, B.; Morris, J.; Martin, E. B. Model Selection for Partial Least Squares Regression. *Chemom. Intell. Lab. Syst.* **2002**, 64, 79–89.
- (26) U.S. Department of Energy. *Petroleum Navigator: Refinery Yield*; Energy Information Administration, 2010; http://tonto.eia.doe.gov/dnav/pet/pet_pnp_pct_dc_nus_pct_m.htm.
- (27) Davis, S. C. ; Diegel, S. W. ; Boundy, R. G. *Transportation Energy Data Book*: ed. 28; Oak Ridge National Laboratory: Oak Ridge, TN, 2009.
- (28) U.S. Department of Energy. *Petroleum Navigator: U.S. Imports by Country of Origin*; Energy Information Administration, 2010; http://tonto.eia.doe.gov/dnav/pet/pet_move_impqus_a2_nus_ep00_im0_mbbbl_m.htm.
- (29) Intergovernmental Panel on Climate Change. 2006 IPCC Guidelines for National Greenhouse Gas Inventories; Japan, 2006; http://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/0_Overview/V0_0_Cover.pdf.
- (30) U.S. Department of Energy. International Energy Annual, 2006; <http://www.eia.doe.gov/emeu/iea/contents.html>.
- (31) U.S. Department of Energy. *Petroleum Navigator: PADD District Imports by Country of Origin*; Energy Information Administration, 2010; http://tonto.eia.doe.gov/dnav/pet/pet_move_impqus_a2_r10_ep00_ip0_mbbbl_m.htm.
- (32) U.S. Department of Energy. *Petroleum Navigator: Refinery Net Input*; Energy Information Administration, 2010; http://tonto.eia.doe.gov/dnav/pet/pet_pnp_inpt2_dc_nus_mbbbl_m.htm.
- (33) California Environmental Protection Agency Air Resources Board. Detailed CA-GREET Pathway for California Reformulated Gasoline Blendstock for Oxygenate Blending (CARBOB) from Average Crude Refined in California, 2009; http://www.arb.ca.gov/fuels/lcfs/022709lcfs_carbob.pdf.
- (34) U.S. Environmental Protection Agency. *Renewable Fuel Standard Program (RFS2) Regulatory Impact Analysis*; Washington, DC, 2010.
- (35) Mullins, K.; Griffin, W. M.; Matthews, H. S. Policy Impacts of Uncertainty in Modeling Life-Cycle Greenhouse Gas Emissions of Biofuels. Submitted Environ. Sci. Technol. 2010.

ES102498A