A MINLP MODEL FOR GLOBAL OPTIMIZATION OF PLUG-IN HYBRID VEHICLE DESIGN AND ALLOCATION TO MINIMIZE LIFE CYCLE GREENHOUSE GAS EMISSIONS

Ching-Shin Norman Shiau
Research Assistant
Mechanical Engineering
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213
cshiau@cmu.edu

Jeremy J. Michalek*
Associate Professor
Mechanical Engineering
Engineering and Public Policy
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213
jmichalek@cmu.edu
* Corresponding Author

ABSTRACT
Plug-in hybrid electric vehicles (PHEVs) have potential to reduce greenhouse gas (GHG) emissions in the U.S. light-duty vehicle fleet. GHG emissions from PHEVs and other vehicles depend on both vehicle design and driver behavior. We pose a twice-differentiable, factorable mixed-integer nonlinear programming model utilizing vehicle physics simulation, battery degradation data, and U.S. driving data to determine optimal vehicle design and allocation for minimizing lifecycle greenhouse gas (GHG) emissions. The resulting nonconvex optimization problem is solved using a convexification-based branch-and-reduce algorithm, which achieves global solutions.

In contrast, a randomized multistart approach with local search algorithms finds global solutions in 59% of trials for the two-vehicle case and 18% of trials for the three-vehicle case. Results indicate that minimum GHG emissions is achieved with a mix of PHEVs sized for around 35 miles of electric travel. Larger battery packs allow longer travel on electric power, but additional battery production and weight result in higher GHG emissions, unless significant grid decarbonization is achieved. PHEVs offer a nearly 50% reduction in life cycle GHG emissions relative to equivalent conventional vehicles and about 5% improvement over ordinary hybrid electric vehicles. Optimal allocation of different vehicles to different drivers turns out to be of second order importance for minimizing net life cycle GHGs.

1. INTRODUCTION
Plug-in hybrid electric vehicle (PHEV) is a promising technology for addressing the issues of foreign oil dependency and global warming within the U.S. transportation sector [2].

PHEVs are similar to ordinary hybrid electric vehicles (HEVs), except the PHEV carries a larger battery pack and offers plug-charging capability [3]. PHEVs use large battery packs to store energy from the electricity grid and propel the vehicle partly on electricity instead of gasoline. Under the average mix of electricity sources in the U.S., vehicles can be driven with lower operation cost and fewer greenhouse gas (GHG) emissions per mile when powered by electricity rather than by gasoline [4]. PHEVs have the potential to displace a large portion of gasoline consumed by the transportation, since approximately 60% of daily U.S. passenger vehicle trips are less than 30 miles [5].

We focus our design study on PHEVs with an all-electric control strategy, which disables engine operation in charge-depleting mode (CD mode) and draws propulsion energy entirely from the battery until it reaches a target state of charge (SOC) [3]. Figure 1 shows the battery energy status in PHEV operation. The distance that a PHEV can travel on electricity alone with a fully charged battery is called its all-electric range (AER). Battery swing is the window between maximum SOC and target SOC, which is determined by the control strategy, and we base our swing definition on percent of cell energy used, rather than percent of SOC. Once the driving distance reaches the AER and the battery is depleted to the target SOC, the PHEV switches to operate in charge-sustaining mode (CS mode), and the gasoline engine provides energy to propel the vehicle and maintain battery charge near the target SOC. In CS mode, the PHEV operates similar to an ordinary HEV.

A blended-strategy PHEV uses a mix of the electric motor and gasoline engine to power the vehicle in CD-mode, while an all-electric PHEV uses only electricity. We confine our scope to all-electric strategy for simplicity, since blended-strategy operation characteristics are sensitive to control parameters.

1 AER is defined as energy-equivalent electric distance for blended-mode PHEVs, but we consider only all-electric PHEVs in this study [6].
and can perform poorly on constrained problems [9]. In contrast, gradient-based global solvers, such as BARON, can produce global solutions for significantly larger problems, but they require that objective and constraint functions be twice differentiable, factorable algebraic functions so that valid convex underestimation functions can be automatically constructed in nodes of the branch and bound tree [10].

We formulate our problem as a twice-differentiable, factorable MINLP that can be solved using BARON to ensure global solutions while managing both continuous and discrete variables. In Section 2, we first develop the formulation with specific models for the objective and constraint functions by specifying the distribution of miles driven per day, vehicle performance models, and the objective and constraint formulations. We then reformulate the model as a factorable, algebraic nonconvex MINLP that can be solved globally. In Section 3, we report solutions for minimum GHG emissions and compare the solution performance to the multi-start method with a local NLP solver. We then conclude in Section 4.

2. MODEL

To optimize a single vehicle for minimum GHG emissions over the population of drivers, we minimize the integral of $f(x,s)$, the GHG emissions per day for vehicle design $x$ when driven $s$ miles per day, times the probability density function of daily driving distance $f_{s}(s)$ over the population of drivers.

$$
\begin{align*}
\text{minimize} & \quad \int_{0}^{\infty} f_{s}(x,s) f_{s}(s) \, ds \\
\text{subject to} & \quad g(x) \leq 0; \quad h(x) = 0
\end{align*}
$$

(1)

where $g(x)$ is a vector of inequality constraints and $h(x)$ is a vector of equality constraints ensuring a feasible vehicle design. We assume each vehicle is charged once per day, so that $s$ indicates the distance traveled between charges.

To extend this model to the case where different drivers are assigned different vehicles based on the distance driven per day, we incorporate a new decision variable $s$ that defines the cutoff point such that drivers who travel less than $s$ per day are assigned the vehicle defined by $x$, and drivers who travel more than $s$ per day are assigned the vehicle defined by $x_{s+1}$. Extending this idea to multiple segments, the formulation for vehicle design and ordered allocation is given by

$$
\begin{align*}
\text{minimize} & \quad \sum_{i=1}^{n} \left( \int_{s_{i-1}}^{s_{i}} f_{s}(x,s) f_{s}(s) \, ds \right) \\
\text{subject to} & \quad g(x) \leq 0; \quad h(x) = 0; \quad \forall i \in \{1, \ldots, n\} \\
& \quad s_{i} \geq s_{i-1}; \quad \forall i \in \{1, \ldots, n\}
\end{align*}
$$

(2)

where $s_{0} = 0; \quad s_{n} = \infty$

Taking a two-vehicle-segment case as an example, the vehicle 1 segment contains vehicles that travel between $[0, s_{1}]$ miles per day; the vehicle 2 segment contains vehicles that travel between $[s_{1}, \infty]$ miles per day; and the optimal value of $s_{1}$ is determined together with the vectors of vehicle design variables $x_{1}$ and $x_{2}$ for vehicle 1 and vehicle 2.
2.1 Distribution of Vehicle Miles Travelled per Day

We use the 2009 National Household Transportation Survey (NHTS) data [5] to estimate the distribution of distance driven per day over the population of drivers. The survey collected data by interviewing 136,140 households across the U.S. on the mode of transportation, duration, distance and purpose of the trips taken on the survey day. We fit the weighted driving data\(^5\) using the exponential distribution. The probability density function below represents the probability density function for vehicle miles traveled by drivers on the day surveyed:

\[
f(s) = \lambda e^{-\lambda s}; \quad s \geq 0
\]

The coefficient \(\lambda\) is 0.0296 estimated using the maximum likelihood method. Because we lack multiple days of data for each vehicle, we assume that a vehicle that travels \(s\) miles per day on the NHTS survey day will travel \(s\) miles every day. This assumption will produce optimistic results on the benefits of optimal allocation, since distance traveled varies over time for individual vehicles in practice.

2.2 Vehicle Performance Models

The vehicle performance evaluations are carried out using the Powertrain System Analysis Toolkit (PSAT) vehicle physical simulator developed by Argonne National Laboratory [11]. In our study the body, powertrain and vehicle parameters for all PHEV and HEV simulations are based on the 2004 Toyota Prius model that uses the split powertrain system with an Atkinson engine, a permanent magnet motor, and a nickel-metal hydride (NiMH) battery pack. To account for structural weight needed to carry heavy battery packs, we include an additional 1 kg of structural weight per 1 kg of battery and motor weight [12]. We created a comparable conventional vehicle (CV) model using a conventional powertrain and 4-cylinder engine to account for larger engine torque and power requirements, and the parameters that define the frontal area, drag coefficient and base weight are adjusted to match the Prius for fair comparison.

For the PHEV design, the Prius engine size is scaled by the peak power output from the base engine (57 kW) using a linear scaling algorithm. Similarly, the motor is scaled from the base motor (52 kW) linearly. Both the engine and motor weights are also scaled proportionally to the peak power. We use the Saft Li-ion battery module in the PSAT package for the PHEV energy storage device. Each cell in the module weighs 0.378 kg, with a modified specific energy of 100 Wh/kg and has a battery cell energy capacity of 21.6 Wh with a nominal output voltage of 3.6 volts. The weight of each 3-cell module is 1.42 kg after accounting for a packaging factor of 1.25. The battery size and capacity are scaled by specifying the number of cells in the battery pack. We assume an 800W base electrical hotel load on the PHEV, the HEV and the CV. To estimate the performance of a PHEV, we use the federal standard Urban Dynamometer Driving Schedule (UDDS) driving cycle [13] to calculate simulated electrical efficiency (miles/kWh) in CD-mode for PHEVs, and gasoline efficiency (mpg) in CS-mode for PHEVs as well as for HEVs and CVs.\(^6\) We also perform a simulated performance test to calculate the time required to accelerate the vehicle from 0 to 60 miles per hour (mph) in the CD mode and in the CS mode.

Because the GHG emissions per mile associated with HEVs and CVs are independent of the number of miles driven per day, we take the HEV and CV to have fixed designs. The HEV is identical to the Prius model, which has a configuration of peak engine power 57 kW, motor power 52 kW, NiMH battery size 168 cells (1.3 kWh), fuel efficiency 60.1 miles per gallon, and 0-60 mph acceleration time 11.0 seconds. Similarly, our CV has an engine size 126 kW and fuel efficiency 29.5 miles per gallon, and 0-60 mph acceleration time 11.0 seconds. For the PHEVs, the design variables \(x\) consist of the engine scaling factor \(x_1\), motor scaling factor \(x_2\), battery pack scaling factor \(x_3\), and battery energy swing \(x_4\). We created a set of polynomial meta-model fits as functions of \(x\) for the PHEV using discrete simulation data points: (1) CD-mode electricity efficiency \(\eta_1\) (mile per kWh); (2) CS-mode gasoline efficiency \(\eta_2\) (mile per gallon); (3) CD-mode 0-60 mph acceleration time \(t_{CD}\) (second); (4) CS-mode 0-60 mph acceleration time \(t_{CS}\) (second); (5) CD-mode battery energy processed (charging and discharging) per mile \(\mu_{CD}\) (kWh/mile); (6) CS-mode battery energy processed per mile \(\mu_{CS}\) (kWh/mile); and (7) final SOC after completing multiple US06 aggressive driving cycles in CS mode \(u_{CS}\) (starting at the target SOC). Metamodels of \(\eta_1\) and \(\eta_2\) are used to calculate energy consumption; \(t_{CD}\) and \(t_{CS}\) are used to ensure comparison of equivalent-performance vehicles; \(\mu_{CD}\) and \(\mu_{CS}\) are used to calculate battery degradation, and \(u_{CS}\) is used to ensure the engine is capable of providing average power needs in CS mode. We evaluated these output values using PSAT over a grid of values for the inputs \(x_1 = \{30, 45, 60\}/57, x_2 = \{50, 70, 90, 110\}/52, x_1 = \{200, 400, 600, 800, 1000\}/1000,\) and multivariate polynomial functions were fit to the data using least squares. The general form of the cubic fitting function \(f_{m3}(x)\) is defined as (the subscript 3 indicates the PHEV case, which will be discussed later).

\[
f_{m3}(x) = a_{m3}x_1^3 + a_{m3}x_2^3 + a_{m3}x_3^3 + a_{m3}x_4^3 + a_{m3}x_1x_2^2 + a_{m3}x_1x_3^2 + a_{m3}x_1x_4^2 + a_{m3}x_2x_3^2 + a_{m3}x_2x_4^2 + a_{m3}x_3x_4^2 + a_{m3}x_1^2x_2 + a_{m3}x_1^2x_3 + a_{m3}x_1^2x_4 + a_{m3}x_2^2x_3 + a_{m3}x_2^2x_4 + a_{m3}x_3^2x_4 + a_{m3}x_1x_2x_3 + a_{m3}x_1x_2x_4 + a_{m3}x_1x_3x_4 + a_{m3}x_2x_3x_4 + a_{m3}x_1^3 + a_{m3}x_2^3 + a_{m3}x_3^3 + a_{m3}x_4^3 + a_{m3}x_1x_2 + a_{m3}x_1x_3 + a_{m3}x_1x_4 + a_{m3}x_2x_3 + a_{m3}x_2x_4 + a_{m3}x_3x_4 + a_{m3}x_1^2 + a_{m3}x_2^2 + a_{m3}x_3^2 + a_{m3}x_4^2 + a_{m3}x_1 + a_{m3}x_2 + a_{m3}x_3 + a_{m3}x_4 + a_{m3}
\]

where the \(a_{m3}\) terms are the coefficients for function \(m\). The polynomial fitting coefficients for \(\eta_1, \eta_2, t_{CD}, t_{CS}, \mu_{CD}, \mu_{CS}\) and \(u_{CS}\) are listed in Table A1 in Appendix.\(^7\) The maximum

\(^5\) We excluded data entries of publication transportation and also excluded drivers who traveled zero miles or more than 200 miles. We fit the distribution to the reported distance traveled on the survey day, and we assume that (1) the survey data calibrated with weightings are representative of the national population, and (2) the distance driven on the survey day is the same distance driven every day for that vehicle.

\(^6\) Examination of alternative driving cycles and the correlation between driving cycle and driving distance is left for future work.

\(^7\) We truncated the acceleration data points greater than 13.0 seconds to improve the metamodel fit, and fit \(\mu_{CS}\), \(\mu_{CS}\) and \(u_{CS}\) using quadratic terms to avoid over-fitting.
2.3 Electric Travel and Battery Degradation

To calculate the objective function of GHG emissions, we first define the distance driven on electric power \( s_E \) and the distance driven on gasoline \( s_G \) as a function of the vehicle’s AER \( s_{AER} \) and the total distance driven per day \( s \). Assuming one charge per day, \( s_E \) and \( s_G \) are given by

\[
s_E(x, s) = \begin{cases} s & \text{if } s \leq s_{AER} \\ s_{AER}(x) & \text{if } s > s_{AER} \end{cases}
\]

\[
s_G(x, s) = \begin{cases} 0 & \text{if } s \leq s_{AER} \\ s - s_{AER}(x) & \text{if } s > s_{AER} \end{cases}
\]

(5)

For PHEVs, we assume that the battery is fully charged once each day. For HEVs and CVs, there is no electrical travel; thus HEV and CV can be seen as special cases with \( s_{AER} = 0 \), so that \( s_E = 0 \) and \( s_G = s \). Assuming constant efficiency \( \eta_B \) (mile per kWh) in CD-mode, the AER of a PHEV can be calculated from the energy capacity per battery cell \( \kappa = 0.0216 \) kWh/cell, the (scaled) number of cells \( x_s \), and the battery swing \( x_g \):

\[
s_{AER}(x) = \kappa(1000x_s)x_g\eta_B
\]

(6)

We use the Peterson model [14] for estimating battery degradation and replacement. The model was constructed by cycling modern A123 LiFePO4 cells under representative driving cycles (non-constant C-rate) and measuring capacity fade as a function of energy processed, including intermediate charging and discharging over the driving cycle. Results show relative energy capacity fade as a linear function of normalized energy processed while driving and while charging. The daily energy processed while driving \( w_{DRV} \) and charging \( w_{CHG} \) a PHEV can be expressed as (unit in kWh):

\[
w_{DRV}(x, s) = \mu_{CD} s_E + \mu_{CS} s_G
\]

\[
w_{CHG}(x, s) = \frac{s_E}{\eta_B}\eta_B
\]

(7)

where \( \mu_{CD} \) and \( \mu_{CS} \) are energy processed per mile (kWh/mile) in CD and CS mode, respectively, and \( \eta_B \) is battery charging efficiency 95% [14]. We assume that energy processed for daily charging is equal to net energy consumed in electrical travel per day. The relative energy capacity decrease can be calculated by the energy processed in driving and charging per cycle per cell per original cell energy capacity:

\[
r_p(x, s) = \alpha_{DRV} w_{DRV} + \alpha_{CHG} w_{CHG}
\]

\[
\frac{1000x_s}{\kappa}
\]

(8)

where \( \alpha_{DRV} = 3.46 \times 10^{-5} \) and \( \alpha_{CHG} = 1.72 \times 10^{-5} \) are the coefficients for relative energy capacity fade. These coefficients are derived from the same data set described in [14]. If the battery end-of-life is defined as the point when the drop in relative energy capacity is \( r_{EOL} \), the battery life \( \theta_{BAT} \), measured in days (or, equivalently, cycles), can be calculated as

\[
\theta_{BAT} = \frac{r_{EOL}}{r_p} \left( \mu_{CD} s_E + \mu_{CS} s_G \right) + \left( \eta_B \eta_B \right)
\]

(9)

The \( r_{EOL} \) criterion is defined at 20% [14].

2.4 Objective and Constraint Functions

The operating (use phase) GHG emissions \( v_{OP} \) represents the average GHG emissions in kg CO2 equivalent (kg-CO2-eq) per day associated with the lifecycle of gasoline and electricity used to propel the vehicle:

\[
v_{OP}(x, s) = \frac{s_E}{\eta_B}\eta_B + \frac{s_G}{\eta_G}(x)\eta_G
\]

(10)

where \( \eta_B = 88\% \) for battery charging efficiency [15], \( v_E = 0.752 \) kg-CO2-eq per kWh for electricity emissions, and \( v_G = 11.34 \) kg-CO2-eq per gallon for gasoline lifecycle emissions. Total lifecycle GHG emissions further includes the GHGs associated with production of the vehicle and battery. The average total lifecycle GHG emissions per day \( f_v(x, s) \) is

\[
f_v(x, s) = \frac{s_E}{\eta_B}\eta_B + \frac{s_G}{\eta_G}(x)\eta_G
\]

\[
+ \frac{v_{VEH}}{\theta_{VEH}(s)} + \left( \frac{1000x_s\kappa\eta_B}{\theta_{BRPL}(s)} \right)
\]

(11)

where \( \theta_{VEH} = s_{LIFE}/s \) is the vehicle life in days, \( s_{LIFE} = 150,000 \) miles [10] is the vehicle life in miles, \( \theta_{BRPL} \) is the battery replacement effective life (defined below), \( v_{BAT} = 120 \) kg-CO2-eq per kWh for Li-ion battery and 230 kg-CO2-eq per kWh for NiMH battery is the lifecycle GHG emissions associated with battery production, \( v_{VEH} = 8.500 \) kg-CO2-eq per vehicle is the lifecycle GHG emissions associated with vehicle production (excluding emissions from battery production) [4].

To compare comparable vehicles, we require that all vehicles meet a minimum acceleration constraint of 0-60 mph in less than 11 seconds. Because we have limited our scope to all-electric PHEVs, we require the acceleration constraint to be satisfied both in CD mode, using electric power alone, and in CS mode, where the gasoline engine is also used. The resulting constraints are \( \ell_{CD}(x) \leq 11s \) and \( \ell_{CS}(x) \leq 11s \). Additionally, we require the gasoline engine to be large enough to provide average power for the vehicle in CS mode under an aggressive US06 driving cycle while maintaining the target SOC level in the battery. The resulting constraint is \( u_{CS}(x) \geq 32\% \). Finally, we impose simple bounds on the decision variables: \( 30/57 \leq x_s \leq 60/57, 50/52 \leq x_s \leq 110/52, 200/1000 \leq x_s \leq 1000/1000, 0 \leq x_s \leq 4 \).

---

* The regression in [14] focused on finding the degradation from energy arbitrage, but in this paper the data was assigned to categories to enable predictions about degradation due to driving and recharging.

---

Footnotes:

8. The lifecycle GHG emissions of electricity is estimated using the US national average electricity emission factor 0.069 kg-CO2-eq/kWh [16] with 9% transmission loss [17].

10. We assume that all vehicles must be replaced every 150,000 miles, which represents the U.S. average vehicle life [18]. This assumption may be unrealistic for vehicles driven very short or very long daily distances because other time-based factors also play a role in vehicle deterioration. However, these factors are only significant for regions of the objective function’s integrand that are relatively insignificant to the integrated objective function, and they do not provide a significant source of error.
\[ \leq 0.8 \text{ to avoid metamodel extrapolation. Any active simple bounds would imply modeling limitation rather than a physical optimum.} \]

### 2.5 MINLP Reformulation

The resulting model formulation (Eq. (2)) involves integration, discrete decisions (vehicle type), and piecewise-smooth functions with derivative discontinuities due to AER and battery life. To solve the problem globally, we pose a factorable, algebraic nonconvex MINLP reformulation that can be solved using the BARON convexification-based branch-and-reduce algorithm [19]. First, the exponential distribution form for the NHTS data fit allows the integral in Eq. (2) to be simplified in terms of two algebraic formulae: the cumulative density function \( F_s \):\[
F_s(s) = \int_0^s f(s) ds = \int_0^s e^{-\lambda s} ds
\]
and the partial expected function \( F_E \):\[
F_E(s) = \int_0^s s \lambda e^{-\lambda s} ds = \left[ -e^{-\lambda s} s \right]_0^s + \int_0^s e^{-\lambda s} ds = \frac{1}{\lambda} - e^{-\lambda s} \left( s + \frac{1}{\lambda} \right)
\]
Thus the problem reduces to an algebraic formulation with discrete vehicle-type decisions and piecewise-smooth functions. We next introduce four sets of binary variables to convert the problem to a two-differentiable MINLP. The first binary variable set \( t_{ij} \) identifies the vehicle type \( l \in \{1,2,\ldots,L\} \) for each segment \( i \), where \( t_{ij} \in \{0,1\} \) \( \forall i,j \) and \( \sum t_{ij} = 1 \) \( \forall i \). Here we consider three vehicle types \( l \in \{1,2,3\} \) for CV, HEV and PHEV, respectively, and write the objective function as a binary-weighted function \( \sum f(t)(f(x_i,s)) \).

The second binary variable set \( z_{ij} \in \{0,1\} \) \( \forall i,j \) handles one of the derivative discontinuities by identifying in which of three regions \( j \in \{1,2,3\} \) on the \( s \)-axis each segment \( i \) is located, relative to \( s_{AER} \) \( (\Sigma z_{ij} = 1) \):

1. In region 1, \( s_{AER} \leq s_i \Rightarrow (z_{i1})(s_{AER}(x_i) - s_{i1}) \leq 0 \);
2. In region 2, \( s_{i1} \leq s_{AER} \leq s_i \Rightarrow (z_{i2})(s_{i1} - s_{AER}(x_i)) \leq 0 \) \( \text{and} \ (z_{i3})(s_{AER}(x_i) - s_{i3}) \leq 0 \);
3. In region 3, \( s_{AER} \geq s_i \Rightarrow (z_{i3})(s_i - s_{AER}(x_i)) \leq 0 \)

The population-weighted operation GHG emissions are

\[
F_{AER} = \sum_{j=1}^{3} \sum_{i} z_{ij} f(x_i,s) ds
\]

\[
= \begin{cases} 
\sum_{i} \left( s_{i1}(x_i,s) \frac{V_E}{\eta_G} + \sum_{j} \frac{V_E}{\eta_G} s_{ij} \right) f(s) ds & \text{if } s_{AER} \leq s_{i1} \\
\sum_{i} \left( \frac{V_E}{\eta_G} s_{i1} + \frac{V_E}{\eta_G} s_{i2} \right) e^{-\lambda s_{i2}} ds & \text{if } s_{i1} \leq s_{AER} \leq s_i \\
\sum_{i} \left( \frac{V_E}{\eta_G} s_{i2} e^{-\lambda s_{i2}} ds + \sum_{j} \left( \frac{V_E}{\eta_G} s_{ij} \right) e^{-\lambda s_{i2}} ds \right) & \text{if } s_{AER} \leq s_i \\
\end{cases}
\]

Applying Eq. (12) and (13), the above integral can be expressed as the sum of three analytical functions:

\[
F_{AER} = \sum_{j=1}^{3} \sum_{i} z_{ij} f(x_i,s) ds
\]

\[
= \begin{cases} 
\sum_{i} \left( s_{i1}(x_i,s) \frac{V_E}{\eta_G} + \sum_{j} \frac{V_E}{\eta_G} s_{ij} \right) f(s) ds & \text{if } s_{AER} \leq s_{i1} \\
\sum_{i} \left( \frac{V_E}{\eta_G} s_{i1} + \frac{V_E}{\eta_G} s_{i2} \right) e^{-\lambda s_{i2}} ds & \text{if } s_{i1} \leq s_{AER} \leq s_i \\
\sum_{i} \left( \frac{V_E}{\eta_G} s_{i2} e^{-\lambda s_{i2}} ds + \sum_{j} \left( \frac{V_E}{\eta_G} s_{ij} \right) e^{-\lambda s_{i2}} ds \right) & \text{if } s_{AER} \leq s_i \\
\end{cases}
\]

\[
F_{AER} = \sum_{j=1}^{3} \sum_{i} z_{ij} f(x_i,s) ds
\]

\[
= \begin{cases} 
\sum_{i} \left( s_{i1}(x_i,s) \frac{V_E}{\eta_G} + \sum_{j} \frac{V_E}{\eta_G} s_{ij} \right) f(s) ds & \text{if } s_{AER} \leq s_{i1} \\
\sum_{i} \left( \frac{V_E}{\eta_G} s_{i1} + \frac{V_E}{\eta_G} s_{i2} \right) e^{-\lambda s_{i2}} ds & \text{if } s_{i1} \leq s_{AER} \leq s_i \\
\sum_{i} \left( \frac{V_E}{\eta_G} s_{i2} e^{-\lambda s_{i2}} ds + \sum_{j} \left( \frac{V_E}{\eta_G} s_{ij} \right) e^{-\lambda s_{i2}} ds \right) & \text{if } s_{AER} \leq s_i \\
\end{cases}
\]

\[
F_{AER} = \sum_{j=1}^{3} \sum_{i} z_{ij} f(x_i,s) ds
\]

\[
= \begin{cases} 
\sum_{i} \left( s_{i1}(x_i,s) \frac{V_E}{\eta_G} + \sum_{j} \frac{V_E}{\eta_G} s_{ij} \right) f(s) ds & \text{if } s_{AER} \leq s_{i1} \\
\sum_{i} \left( \frac{V_E}{\eta_G} s_{i1} + \frac{V_E}{\eta_G} s_{i2} \right) e^{-\lambda s_{i2}} ds & \text{if } s_{i1} \leq s_{AER} \leq s_i \\
\sum_{i} \left( \frac{V_E}{\eta_G} s_{i2} e^{-\lambda s_{i2}} ds + \sum_{j} \left( \frac{V_E}{\eta_G} s_{ij} \right) e^{-\lambda s_{i2}} ds \right) & \text{if } s_{AER} \leq s_i \\
\end{cases}
\]
Figure 3. Four conditions for the battery and vehicle VMT curves

Condition (a) occurs when the following inequality is valid:

\[ s_{BAT}^0 \leq s_{LIFE} \] (17)

For condition (b), an analytical expression for \( s_T \) is available by solving \( s_{LIFE} = s_{BAT}(x, s_T) \):

\[ s_T(x) = s_{LIFE} s_{AER} \frac{\alpha_{DRV} (\mu_{CD} - \mu_{CS}) + \alpha_{CHO} (\eta_{EB})^{-1}}{\eta_{EL} k(1000x)} - \alpha_{DRV} \mu_{CS} s_{LIFE} \] (18)

Condition (c) and (d) occur when the following inequality is valid:

\[ s_{BAT}^0 \geq s_{LIFE} \] (19)

The battery replacement effective life \( \theta_{BRPL} \) under the buy-leasce scenario is determined by \( \min(\theta_{VEH}, \theta_{BAT}) \). Therefore, three discrete cases are identified:

1. In case 1 (\( \varphi = 1 \)), \( s_{BAT}^0 \leq s_{LIFE} \) \( \theta_{BAT} < \theta_{VEH} \) \( \forall s \Rightarrow (q_1)(\theta_{BAT}(s = \infty)) - \theta_{VEH} \leq s_{BRPL} = \theta_{BAT} \):
2. (case 2 (\( \varphi = 2 \)), \( \theta_{BAT} \) intersects \( \theta_{VEH} \) at a point \( s_T \Rightarrow (q_2)(\theta_{VEH} - \theta_{BAT}(s = \infty)) \leq 0 \) and \( (q_2)(\theta_{BAT}(s = 0)) - \theta_{VEH} \leq 0 \), so \( \theta_{BRPL} = \theta_{BAT} \) for \( s \leq s_T \), \( \theta_{BRPL} = \theta_{VEH} \) for \( s \geq s_T \); and
3. (case 3 (\( \varphi = 3 \)), \( s_{BAT}^0 \geq s_{LIFE} \) \( \theta_{BAT} \geq \theta_{VEH} \) \( \forall s \Rightarrow (q_3)(\theta_{VEH} - \theta_{BAT}(s = 0)) \leq 0 \) and \( \theta_{BRPL} = \theta_{VEH} \).

The fourth binary variable set \( \gamma_{ik} \) identifies in which region \( s_T \) lies when \( q_2 = 1 \) (\( \varphi = 2 \)). The three conditions \( k \in \{1, 2, 3\} \) on the binary variable set \( \gamma_{ik} \) are:

1. In region 1 \( k = 1 \), \( s_T(x) \leq s_{i-1} \Rightarrow (q_2)(\gamma_{ik})(s_T(x) - s_{i-1}) < 0 \);
2. in region 2 \( k = 2 \), \( s_{i-1} \leq s_T \leq s_i \Rightarrow (q_2)(\gamma_{ik})(s_i - s_T(x)) \leq 0 \) and \( (q_2)(\gamma_{ik})(s_T(x) - s_i) \leq 0 \); and
3. in region 3 \( k = 3 \), \( s_T(x) \geq s_i \Rightarrow (q_2)(\gamma_{ik})(s_i - s_T(x)) \leq 0 \).

The combinations of \( j, k \) and \( o \) result in 27 cases. For each segment \( i \), for each of the cases \( j \in \{1, 2, 3\} \) \( k \in \{1, 2, 3\} \) \( o \in \{1, 2, 3\} \), the integral in Eq. (2) reduces to a twice-differentiable closed form factorable algebraic expression \( F_{ij}(x, s, s_{i-1}, s_t) \). Table 1 presents the summary of the discrete cases with corresponding \( \theta_{BRPL} \) and the components in the total cost function. Among the \( o = 2 \) cases, there are three infeasible cases because the value of \( s_T \) should be greater than \( s_{AER} \) when an \( s_T \) point exists. The analytical expressions of battery GHG emission function \( F_{BG} \) for all discrete cases are listed in follows:

Case (1a): \( \theta_{BRPL} = \theta_{BAT} \) and \( s_{AER} \leq s_{i-1} \)

\[ F_{BG1a} = \frac{v_{BAT}}{t_{EOL}} \left[ \alpha_{DRV} w_{CS} (F_E(s) - F_E(s_{i-1})) + s_{AER} (\alpha_{DRV} w_{CD} + \alpha_{CHO} (\eta_e\eta_{EB})^{-1} - \alpha_{DRV} w_{CS}) \right] (20) \]

Case (1b): \( \theta_{BRPL} = \theta_{BAT} \) and \( s_{i-1} \leq s_{AER} \leq s_i \)

\[ F_{BG1b} = \frac{v_{BAT}}{t_{EOL}} \left[ (\alpha_{DRV} \mu_{CD} + \alpha_{CHO} (\eta_e\eta_{EB})^{-1})(F_E(s) - F_E(s_{i-1})) + \alpha_{DRV} \mu_{CS} (F_E(s) - F_E(s_{AER})) + s_{AER} (\alpha_{DRV} \mu_{CD} + \alpha_{CHO} (\eta_e\eta_{EB})^{-1} - \alpha_{DRV} \mu_{CS}) \right] (21) \]

Case (1c): \( \theta_{BRPL} = \theta_{BAT} \) and \( s_{i} \leq s_{AER} \)

\[ F_{BG1c} = \frac{v_{BAT}}{t_{EOL}} \left[ (\alpha_{DRV} \mu_{CD} + \alpha_{CHO} (\eta_e\eta_{EB})^{-1})(F_E(s) - F_E(s_{AER}) + \alpha_{DRV} \mu_{CS} (F_E(s) - F_E(s_{i-1})) + s_{AER} (\alpha_{DRV} \mu_{CD} + \alpha_{CHO} (\eta_e\eta_{EB})^{-1} - \alpha_{DRV} \mu_{CS}) \right] (22) \]

Case (2a): \( \theta_{BRPL} = \{ \theta_{BAT}, \theta_{VEH} \} \) and \( s_{AER} \leq s_{i-1} \)

\[ F_{BG2a} = \frac{v_{BAT}}{t_{EOL}} \left[ \alpha_{DRV} \mu_{CS} (F_E(s) - F_E(s_{i-1})) + s_{AER} (\alpha_{DRV} \mu_{CD} + \alpha_{CHO} (\eta_e\eta_{EB})^{-1} - \alpha_{DRV} \mu_{CS}) \right] \]

\[ + \frac{1000v_{BAT} \eta_{EL}}{\eta_{CD} \eta_{EL} \eta_{EB} \mu_{CS} \mu_{CD} \mu_{CS} \mu_{CD}} (s_{AER} F_E(s) - F_E(s_{AER})) \] (23)

Case (2b): \( \theta_{BRPL} = \{ \theta_{BAT}, \theta_{VEH} \} \) and \( s_{i-1} \leq s_{AER} \leq s_i \)

\[ F_{BG2b} = \frac{v_{BAT}}{t_{EOL}} \left[ (\alpha_{DRV} \mu_{CD} + \alpha_{CHO} (\eta_e\eta_{EB})^{-1})(F_E(s) - F_E(s_{i-1})) + \alpha_{DRV} \mu_{CS} (F_E(s) - F_E(s_{AER})) + s_{AER} (\alpha_{DRV} \mu_{CD} + \alpha_{CHO} (\eta_e\eta_{EB})^{-1} - \alpha_{DRV} \mu_{CS}) \right] \]

\[ + \frac{1000v_{BAT} \eta_{EL}}{\eta_{CD} \eta_{EL} \eta_{EB} \mu_{CS} \mu_{CD} \mu_{CS} \mu_{CD}} (s_{AER} F_E(s) - F_E(s_{AER})) \] (24)

Case (2c): \( \theta_{BRPL} = \theta_{VEH} \)

\[ F_{BG2c} = \frac{1000v_{BAT} \eta_{EL}}{\eta_{CD} \eta_{EL} \eta_{EB} \mu_{CS} \mu_{CD} \mu_{CS} \mu_{CD}} (s_{AER} F_E(s) - F_E(s_{AER})) \] (25)
The optimal vehicle type, $s_{ij}$, is designed and allocated to drivers based on VMT to minimize life cycle GHG emissions; (2) Each driver has a constant daily driving distance over the vehicle life; (3) The US NHTS weighted driving data are described using the exponential distribution function; (4) Vehicle performances are measured using EPA UDDS driving cycle simulation, and PHEV is assumed one full charge per day; (5) Lifecycle GHG emissions assume the average US grid mix.

3. RESULTS AND DISCUSSION

We consider three driver segment scenarios, $n=1, 2$ and 3, and solve the MINLP model (Eq. (26)) using GAMS/BARON solver to obtain global solutions. The optimal vehicle type, design and allocation ranges for each case are summarized in Table 2. The first two data columns show the performance values of CV and HEV. We plot the functional values at the optimal solution $x^*$ in Figure 4. The plots in the upper row show the lifecycle GHG emissions per mile $f_0(x^*, s)/s$, and the plots in the lower row show the population-weighted lifecycle GHG emissions per day $f_0(x, s)/f_{0s}(s)$. The area under the population-weighted curve is the net lifecycle GHG emissions per person per day in the United States.
For the single-segment case, we found that a PHEV36 has the lowest lifecycle GHG emissions. GHG emissions from the HEV scenario are about 44% lower than the CV scenario, and GHGs from the PHEVs scenarios are 5-6% lower than HEVs. For the two-segment case, the optimal solution is to allocate a PHEV40 to drivers who can charge every 87 miles or less (92% of drivers and 74% of VMT per day) and allocate a smaller-range PHEV25 to drivers who charge less frequently. This optimal allocation of two vehicles reduces daily GHG emissions by only an additional 0.1% compared to allocating all drivers a PHEV36. The solution of three-segment case similarly produces very slight additional GHG reduction, implying that a single segment is able to provide a practical solution for our evaluation of PHEV environmental performances. A significant reduction in GHG emissions is achieved by allocating PHEVs to drivers rather than HEVs or CVs, and there is only a marginal additional gain from optimal allocation in the two- and three-segment cases.

In the plots of the two-segment case (Figure 4(b) and (e)), there are two intersection points between the two PHEV GHG curves, and the optimal single cutoff point is located at the first intersection. Although the PHEV40 GHG curve surpasses the PHEV25 after 87 miles, the difference between two is almost indistinguishable, and the portion of the population driving greater than 87 miles/day is small. Assigning all drivers high-AER PHEVs can significantly reduce petroleum consumption, but this is not necessarily the best solution for minimizing GHGs because reducing the number of unnecessary batteries in these vehicles reduces the emissions associated with battery production as well as the emissions associated with reduced vehicle efficiency caused by carrying heavy batteries. While the largest group of vehicles travel short distances each day, the majority of the GHG emissions are produced by those vehicles that travel between about 25-45 miles/day.

We further tested the single segment case with low carbon electricity scenario (high portion of renewable energy sources) 218 kg-CO2-eq/kWh in the grid mix [4]. The optimal solution shows an optimal large-capacity PHEV87 (upper bound) is best to reduce the average GHG emissions to 4.53 kg-CO2-eq per person per day, and the reduction percentages from CV and HEV are 69% and 45%, respectively. This result implies grid decarbonization is needed in order for large capacity PHEVs to have significantly superior GHG performance than HEVs. This finding is consistent with prior studies [4, 12].

To compare the solution performance of the randomized multi-start method with global solutions, we use 1000 random starting points uniformly distributed in the design variable domain with the Matlab SQP NLP solver *fmincon* to find the minimum GHG solutions. The integrals in Eq. (1) and Eq. (2) are approximated by trapezoidal numerical integration, and the vehicle types for the two-segment and three-segment cases are preselected as PHEV-PHEV and PHEV-PHEV-PHEV, respectively. The solution quality is evaluated using relative error $|F-F^*|/F^*$, where $F$ is the optimal objective value found by local solver with each multi-start and $F^*$ is the global solution. The results are presented in Figure 5. All multi-start points reached feasible optimal solutions. For the single-segment model, the multi-start method performs very well, and all multi-start solutions reached the global optimum. For the two-segment case, 59% of the multi-start solutions reach the global optimum (within 10⁻⁶). For the three-segment case, the ratio of finding the global solution decreases to 18%. Using different NLP solvers may affect the percentages, but the results show that when the numbers of driver segments and design variables increase, the probability of random starting points reaching global solution decreases significantly. In this case, local minima are all within 1% of the global minimum.

<table>
<thead>
<tr>
<th>Driver segment scenario</th>
<th>Single segment</th>
<th>Two segments</th>
<th>Three segments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allocation (miles)</td>
<td>20-200</td>
<td>0-200</td>
<td>0-200</td>
</tr>
<tr>
<td>AER (miles)</td>
<td>36</td>
<td>40</td>
<td>25</td>
</tr>
<tr>
<td>Engine power (kW)</td>
<td>126</td>
<td>57</td>
<td>46</td>
</tr>
<tr>
<td>Motor power (kW)</td>
<td>52</td>
<td>70</td>
<td>71</td>
</tr>
<tr>
<td>Battery cells</td>
<td>168</td>
<td>396</td>
<td>435</td>
</tr>
<tr>
<td>Battery design swing</td>
<td>0.8†</td>
<td>0.8†</td>
<td>0.8†</td>
</tr>
<tr>
<td>Battery capacity (kWh)</td>
<td>1.3</td>
<td>8.6</td>
<td>9.4</td>
</tr>
<tr>
<td>CD-mode eff. (miles/kWh)</td>
<td>5.30</td>
<td>5.29</td>
<td>5.35</td>
</tr>
<tr>
<td>CS-mode eff. (mpg)</td>
<td>60.1</td>
<td>60.2</td>
<td>60.0</td>
</tr>
<tr>
<td>CD-mode acceleration (sec)</td>
<td>11.0</td>
<td>11.0</td>
<td>11.0</td>
</tr>
<tr>
<td>CS-mode acceleration (sec)</td>
<td>9.3</td>
<td>9.1</td>
<td>10.3</td>
</tr>
<tr>
<td>Final SOC</td>
<td>0.32</td>
<td>0.32</td>
<td>0.32</td>
</tr>
<tr>
<td>GHG emissions</td>
<td>14.6</td>
<td>8.2</td>
<td>7.78</td>
</tr>
<tr>
<td>(kg-eq-CO₂ per person-day)</td>
<td>-43.8%</td>
<td>-46.7%</td>
<td>-46.8%</td>
</tr>
</tbody>
</table>

 Variable limited by model boundary
Three segments algebraic nonconvex MINLP that can be daily solved globally using convexification with a branch-and-reduce algorithm implemented in GAMS/BARON. We find that minimum life cycle GHG emissions can be achieved by assigning a medium-range PHEV36 to drivers who have daily travel distance up to 200 miles. Results indicate that moving drivers from conventional vehicles to HEVs or PHEVs implies significant reductions in life cycle GHGs, but optimal allocation of vehicles to drivers is of second order importance.

While larger battery packs may be better for reducing petroleum consumption, larger packs do not necessarily result in lower GHGs, and the best solution is a combination of mid-sized packs that reduce unnecessary GHGs associated with battery production and reduced efficiency due to weight. Grid decarbonization makes larger battery packs more competitive for GHG reduction.

Adoption of PHEVs will depend critically on cost. We examine cost and petroleum consumption objectives in a companion paper and examine sensitivity of minimum cost solutions to variation in parameters such as battery prices, fuel and electricity prices, and carbon allowance prices [20]. The global MINLP framework presented here provides confidence in comparing solutions across sensitivity scenarios.

ACKNOWLEDGMENTS

The authors would like to thank Prof. Jay Whitacre, Prof. Chris Hendrickson, Scott B. Peterson, and the members of the Design Decisions Laboratory, the Carnegie Mellon Green Design Institute and Electrified Transportation Group for their feedback and help with model formulation. This research was supported in part by the National Science Foundation’s CAREER Award #0747911, a grant from the National Science Foundation program for Material Use, Science, Engineering and Society (MUSES); Award #0628084, and grants from Ford Motor Company and Toyota Motor Corporation. C.-S.S. acknowledges support from the Liang Ji-Dian Fellowship.

REFERENCES


APPENDIX

Table A1. Polynomial coefficients of the PHEV performance meta-model

<table>
<thead>
<tr>
<th>$f_{m3}$</th>
<th>$\eta_R$</th>
<th>$\eta_G$</th>
<th>$\deltaCD$</th>
<th>$\deltaCS$</th>
<th>$\muCD$</th>
<th>$\muCS$</th>
<th>$\muCS$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m$</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>$a_{m1}$</td>
<td>0.008</td>
<td>2.214</td>
<td>1.457</td>
<td>3.334</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m2}$</td>
<td>0.154</td>
<td>1.087</td>
<td>-5.496</td>
<td>-2.266</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m3}$</td>
<td>0.353</td>
<td>5.578</td>
<td>-28.46</td>
<td>-20.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m4}$</td>
<td>-0.005</td>
<td>-0.815</td>
<td>0.913</td>
<td>0.414</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m5}$</td>
<td>-0.005</td>
<td>0.510</td>
<td>-0.881</td>
<td>-3.524</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m6}$</td>
<td>-0.025</td>
<td>1.562</td>
<td>-1.050</td>
<td>-0.286</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m7}$</td>
<td>0.000</td>
<td>2.212</td>
<td>-0.308</td>
<td>-10.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m8}$</td>
<td>-0.057</td>
<td>-0.613</td>
<td>2.044</td>
<td>1.951</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m9}$</td>
<td>-0.043</td>
<td>0.254</td>
<td>15.61</td>
<td>10.31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m10}$</td>
<td>-0.016</td>
<td>-0.159</td>
<td>0.336</td>
<td>5.808</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m11}$</td>
<td>-0.001</td>
<td>-8.906</td>
<td>-4.634</td>
<td>-6.932</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m12}$</td>
<td>-0.805</td>
<td>-6.095</td>
<td>31.48</td>
<td>15.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m13}$</td>
<td>-0.656</td>
<td>-15.21</td>
<td>34.02</td>
<td>39.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m14}$</td>
<td>0.057</td>
<td>0.089</td>
<td>1.153</td>
<td>7.901</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m15}$</td>
<td>0.080</td>
<td>-3.274</td>
<td>1.169</td>
<td>6.582</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m16}$</td>
<td>0.342</td>
<td>2.498</td>
<td>-32.06</td>
<td>-30.12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m17}$</td>
<td>-0.191</td>
<td>2.622</td>
<td>3.405</td>
<td>-6.734</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m18}$</td>
<td>1.189</td>
<td>9.285</td>
<td>-54.47</td>
<td>-26.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m19}$</td>
<td>-0.347</td>
<td>5.837</td>
<td>9.570</td>
<td>-4.098</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_{m20}$</td>
<td>4.960</td>
<td>57.68</td>
<td>44.23</td>
<td>32.10</td>
<td>2.196</td>
<td>1.441</td>
<td>0.140</td>
</tr>
</tbody>
</table>

The terms are fit with quadratic form.


