2023 IEEE/ASME International Conference on
Advanced Intelligent Mechatronics (AIM)

A Grey-Box Surrogate Vehicle Energy Consumption Model Capable of Real-Time Updating

Lingyun Hua[1], Jian Tang[1], Hussein Dourra[2], and Guoming Zhu[1]

***Abstract*—Vehicle energy consumption model, as a function of its operational environment, plays a significant role in real-time optimization of vehicle route and speed for a given pair of origin and destination with a desired arrival time for minimal energy consumption. In this paper, a Grey-Box vehicle energy consumption model is developed based on a high-fidelity vehicle dynamic model with environmental influence based on the Kriging modeling method, which includes rolling resistance, aerodynamics, gravity and energy consumption of air conditioning and heater (HVAC), along with environmental conditions such as temperature, wind speed, etc. The data-driven model, trained based on Gaussian process assumption, ensures the accuracy of the resulting model with a modeling error below 2.5%. The real-time model updating is based on Recursive Least-Squares (RLS) optimization using current driving data so that the model used for route and speed optimization represents the current vehicle status. The proposed Grey-Box model is validated in Computer-In-the-Loop (CIL) simulations using SUMO and MATLAB with less than 2% error of energy consumption, which is a significant improvement over the vehicle dynamic model with up to 35% error in certain cases. A case study also indicates energy consumption reduction for vehicle route-speed optimization.**

***Index Terms*—Energy consumption model, Grey-Box model, Kriging model, Recursive Least-Squares optimization**

I. Introduction

Vehicle driving range and energy consumption attract more and more attention in automotive industry, leading to the development of eco-route planning that requires a vehicle energy model [1] to drive on the route with speed below the speed limit [2]. Thus, simultaneously optimizing vehicle route and speed for a given pair of origin and destination with a desired travel time limit [3] based on the target vehicle dynamic model is developed. However, optimization with Genetic algorithm [3] is complicated and slow, making it impossible for real-time application. In addition, the driving environment usually plays a significant role in route planning [5]. These are the main motivations for developing a Surrogate vehicle energy consumption model with environmental factor inputs to optimize route and speed in real-time for reducing energy consumption and extending drive range.

There are multiple methods to estimate vehicle energy consumption. Dynamic vehicle model has been widely used in this research for decades, however, it provides poor energy estimation accuracy due to the idealization of vehicle dynamics [6]. A well composed vehicle dynamic model was proposed in [7], where the performance of motor, transmission and other powertrain components are considered and high estimation accuracy of transient energy consumption is achieved. A dynamic model with the transmission shift is adopted in [4] using a fixed and short distance energy consumption to simplify the model. A small distance model leads to high computational load and makes it not suitable for real-time application. In addition, the energy consumption is also affected by wind speed (aerodynamics) [8], road conditions [9], [14], air conditioning and heater (HVAC), etc. A complex regression model, composed of several high order polynomial functions for energy consumption, emissions, etc. [25], was designed to optimize a route that results in error due to overfitting and is only capable of initial optimization. In summary, the developed vehicle dynamic model is relatively accurate for determining energy consumption, but its nonlinearity and differential terms in the model make it difficult for calculating energy consumption in real-time. Note that for the route-speed optimization, the differential term associated with acceleration and deceleration is highly dependent on changing traffic and cannot be predicted. As a result, including this term in the model will not improve estimation accuracy during the optimization process and it shall be considered as statistical information used in the modeling process. In addition, these models mentioned above were created based on vehicle physics with fixed model coefficients for vehicle mass, rolling and aerodynamic resistance, etc. But these parameters change due to vehicle aging and wearing (e.g., tire wearing), which could affect the model accuracy [10] and route selection. Therefore, the ability to update the energy consumption model in real-time is very important.

To the best knowledge of authors, there is no existing work using Grey-Box approach for energy modeling. This paper proposes to use the Grey-Box approach to model the steady-state vehicle energy consumption with two key mentioned features: easy to evaluate and real-time updating. The data-driven speed-based Grey-Box model [11] is designed in polynomial format of Kriging model [12] that is easy to evaluate for fast optimization and the model coefficients are formulated as functions of vehicle physical and driving environmental parameters. The model coefficients are trained using input-output data sets with Gaussian Process (GP) optimization and Recursive Least Squares (RLS). The inputs are vehicle speed, ambient temperature, vehicle mass, road grade and wind speed.

[1]Hua, Tang, and Zhu are with Department of Mechanical Engineering, Michigan State University, East Lansing, MI, 48824, USA (emails: hualingy, tangji11, and zhug @msu.edu, respectively)

[2]Dourra is with Magna International, Troy, MI, 48098, USA. (email: Hussein.Dourra@magna.com)

The data-driven model can be updated in real-time to adapt to current vehicle status such as vehicle aging and tire wear. These are the main contributions of this paper.

This paper is organized as follows. Section II describes the physical model of vehicle energy consumption for a unit distance with environmental effects, and Section III introduces the architecture of the Grey-Box Kriging model. Section IV describes the off-line model learning process using MATLAB DACE toolbox with the help of RLS optimization and associated online updating methodologies based on RLS. SUMO-based simulation results of the model training and updating are presented in Section V. Section VI adds some conclusions.

II. VEHICLE MODELING

A. Vehicle Energy Consumption Model

Wide range of vehicle dynamic models are used for control, optimization, and diagnostics, but this paper proposes a simplified steady-state model (with zero acceleration), considering the effects of environmental conditions (wind speed, temperature, etc.), in equation (1). As introduced in [13], the rolling resistance coefficient is a 2^{nd} order function of vehicle speed, and based on [14], it is a function of environmental temperature as well. Thus, the rolling resistance force F_r is formulated in equation (2). Considering the gusting wind has an inevitable influence on vehicle energy consumption [8] and assuming that the vehicle front and rear areas are identical, the aero-dynamic resistance force F_a can be formulated as a function of relative vehicle speed v_r in equation (3). Since this model is intended to reflect steady-state unit distance energy consumption, the acceleration is assumed to be zero. Considering effect of gravity F_g, the vehicle longitudinal dynamic model is formulated below in equation (1).

$$F_p = F_r + F_g + F_a \tag{1}$$

$$F_r = mg(c_1 + c_2 v^2)(1 - c_3 T)\cos(\theta_r) \tag{2}$$

$$F_g = mg\sin(\theta_r),\ F_a = 0.5C_D\rho A v_r^2,\ v_r = v + w \tag{3}$$

where F_p is the total propulsion force of vehicle with constant speed; m, v, g, and T are the vehicle mass, speed, gravity and air temperature, respectively; $c_1 = 0.015, c_2 = 0.0095$, and $c_3 = 0.001$ are calibrated vehicle constants and its operational environment; θ_r is the road grade; C_D, ρ, and A are the vehicle air resistance coefficient, air density, and air resistance frontal surface area, respectively; w is the wind speed, which is positive when the direction of wind is opposite to that of vehicle travel and negative otherwise. Then, the power of propulsion force can be calculated by $P_p = F_p v$.

The energy consumption of vehicle subsystems (such as air conditioning (AC) and heater) is also taken into account. Assuming that the power consumption of mentioned devices P_d is only affected by the temperature. Power usage of these subsystems is assumed to increase when the temperature is low due to required heating and when the environmental temperature is high since AC needs to be turned on to maintain comfort. Based on the energy consumption curve of the Nissan Leaf HVAC system [16] a representative trajectory of P_d denoted as 'True data' is generated in Fig. 1 which is calculated from these equations in Table I for three temperature zones.

TABLE I
POWER CONSUMPTION ON VEHICLE HVAC DEVICES

Temperature	Function
$[-20, 5]$°C	$P_d = 1.8 - 0.063T - 0.0015T^2$
$[5, 35]$°C	$P_d = 2 - 0.13T + 0.004T^2$
$[35, 50]$°C	$P_d = -3.8 + 0.27T - 0.0027T^2$

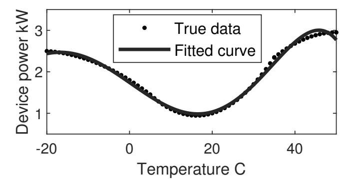


Fig. 1. Power usage on devices

The vehicle total energy consumption is evaluated for traveling a given unit distance $d_u = 500m$ in this study. The travel time t is calculated based on the assumption of constant speed v. As a result, the unit distance energy consumption E_u can be calculated by

$$E_u = (P_d + P_p)t, \quad t = d_u/v \tag{4}$$

Thus, a unit distance vehicle energy consumption model can be developed with varying environments and reformulated in equation (5). For a given environmental condition, the proposed gray-box model is developed in following Sections.

$$E_u = [mg(f cos\theta_r + sin\theta_r) + 0.5C_D\rho A v_r^2 + P_d/v]d_u \tag{5}$$

where $f = (c_1 + c_2 v^2)(1 - c_3 T)$.

B. Traffic Model in SUMO

In this paper, the vehicle and traffic models are modeled and simulated in SUMO [23]. It is well-known that SUMO is a platform for simulating the Global Positioning System (GPS) and Vehicle-To-Everything (V2X) with both macro (e.g. vehicle route and speed limits, etc.) and microscopic traffic (e.g., acceleration and deceleration, etc.) environment. It is capable of studying autonomous route selection and vehicle speed controls. Although the microscopic vehicle control is not covered in this paper, SUMO with its own longitudinal speed control algorithm [23] is adopted to perform vehicle acceleration or brake control on the route since traffic lights and speed limit change affect energy consumption. As a result, speed-based vehicle energy consumption can be obtained.

In addition to microscopic speed control, the vehicle route is available in SUMO and can be controlled by MATLAB through SUMO Traffic Control Interface (TraCI) that allows MATLAB to retrieve the map information (e.g., speed limit, road length, etc.) and send control commands back to SUMO to change the vehicle status and its trip route. The mentioned features can be used to validate the proposed Grey-Box model and study the benefits of the proposed model.

TABLE II
TERM DEFINITIONS OF ENERGY CONSUMPTION MODEL E_u

Order of v	**Term parameter**
v^{-1}	P_d
v^0	$mg(c_1 - c_1c_3T + \sin\theta_r) + 0.5C_D\rho Aw^2$
v	$C_D\rho Aw$
v^2	$mg(c_2 - c_2c_3T) + 0.5C_D\rho A$

III. GREY-BOX KRIGING MODEL

A Kriging (Gaussian process) model [12] is a surrogate model composed of a polynomial function with extrinsic noise. In this paper, the Kriging model is extended with intrinsic noise since the energy consumption model derived from vehicle dynamic model (5) may contain measurement noises and other unknown factors. Thus, the following form of stochastic Kriging model is proposed.

$$y(x) = f^T(x)\beta + z(x) + \epsilon(x) \tag{6}$$

where x denotes the model input (i.e., vehicle speed v); $f(x)$ is a vector function of vehicle speed; β is a parameter vector (constant) to be determined during the model fitting process; $z(x)$ is the extrinsic noise that is a realization of a zero-mean stochastic process with its variance τ^2 [15]; $\epsilon(x)$ is the intrinsic noise due to measurement, assumed to be Gaussian with variance γ^2. Note that the Grey-Box model is used to replace the environmental vehicle energy consumption model (see Section II). It is assumed that intrinsic (measurement) noise is small and ignored in the model but still considered in training process. Also, under the real-world driving condition, there are many factors not considered in the proposed model (such as road surface effect on energy consumption), and as a result, the extrinsic term is kept in the proposed Grey-Box model; see equation (7). The stochastic Kriging model below is generated by solving the parameter vector β and associated variance τ^2; see Section IV.B for details.

$$y(x) = f^T(x)\beta + z(x) \tag{7}$$

Expanding equation (5) introduced in Section II to equation (8) and uniting the like terms, it can be seen that the unit distance energy consumption model is a function of vehicle speed v (see terms in Table II), where $\cos\theta_r \approx 1$ since the road grade θ_r is relatively small. Accordingly, for the scenario that relevant speed $v_r = v + w$ is positive, polynomial function (9) can be formulated due to the orders of vehicle speed v in Table II, where vehicle speed v is represented by x and β denotes the model coefficients. In equation (9), the inverse order term '$\beta_{-1}x^{-1}$' is obtained from the energy usage, P_d/v, of HVAC devices (e.g., AC, heater); the constant term contains the energy consumption not related to vehicle speed (such as vehicle gravity and part of wind effect); the 1^{st} order term is from the expansion of air drag force F_a in equation (3) and the 2^{nd} order term contains forces of air and rolling resistance.

$$\begin{aligned} E_u =& [mg(f cos\theta_r + sin\theta_r) + 0.5C_D\rho Av_r^2 + P_d/v]d_u \\ =& [mg(c_2 - c_2c_3T)v^2 + mg(c_1 - c_1c_3T + \sin\theta_r) \\ &+ 0.5C_D\rho A(v^2 + w^2 + 2wv) + P_dv^{-1}]d_u \end{aligned} \tag{8}$$

$$f^T(x)\beta = \beta_{-1}x^{-1} + \beta_0 + \beta_1x + \beta_2x^2 \tag{9}$$

However, there is a scenario in that wind speed is fast enough to push the car when vehicle speed is low, that is, relative speed $v_r < 0$. With the increasing vehicle speed, the relative speed will become non-negative $v_r \geq 0$. Thus, assuming the vehicle rear areas is the same as the front area A for simplicity, the quadratic term should be a function of relative speed v_r and two Surrogate models, distinguished by the sign of v_r, are needed. Note that different rear area will be used in future. Since other terms in equation (9) are functions of vehicle speed, for the purpose of evaluating the model in full vehicle speed range, it is favorable to use $v_r = 0$ for switching between the two Surrogate models instead of using wind speed w. Furthermore, with a fixed wind speed, the switching sign of v_r in air resistance force makes the model non-analytic and cannot be represented by a single 2^{nd} order polynomial; see the curve of air drag force in Fig. 2 a). Therefore, to get rid of the switching sign of relevant speed v_r, a higher order polynomial function is proposed to fit the curve of air drag force for a given range of wind speed w. To evaluate the feasibility, the curve of air drag force is fitted using the linear Least-Squares method as follows.

For a wind speed range of $w \in [-20, -10]$ m/s, a 3^{rd} order polynomial function of relative speed $v_r = v + w$ is adopted, and the resulting model is simple and accurate. As shown in Fig. 2 a), the fitted curve provides a fairly good approximation of the original function. Therefore, for a given wind speed range of $w \in [-20, -10]$ m/s, the selected 3^{rd} order polynomial is applicable and the modified Kriging model (9) can be extended to (10) as follows.

$$\begin{aligned} y(x) =& \beta_{-1}x^{-1} + \beta_0 + \beta_1x + \beta_2x^2 + \beta_3(x+w)^3 + z(x) \\ =& \beta^T f(x, w) + z(x) \end{aligned} \tag{10}$$

where $f(x, w) = [x^{-1}, 1, x, x^2, (x+w)^3]^T$, w is known wind speed, and $\beta = [\beta_{-1}, \beta_0, \beta_1, \beta_2, \beta_3]^T$.

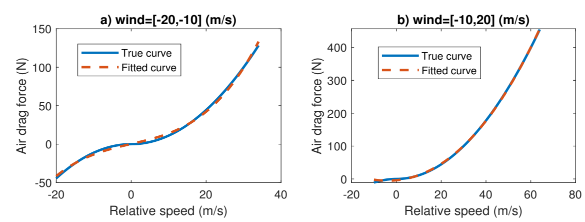


Fig. 2. True and fitted air drag force curves with varying wind speed

In addition, the power consumption P_d on HVAC devices described in Section II consists of three segments. To analyze P_d continuously and maintain the accuracy, the trajectory of P_d can be approximated with a 5^{th} order polynomial of temperature $T \in [-20, 50]$ °C. The 5^{th} order polynomial is selected to well fit the trajectory using Least-Squares algorithm with 95% confidence. The fitted polynomial is in equation (11) and plotted in Fig. 1 noted with 'Fitted curve'.

$$\begin{aligned} P_d =& 1.7 - 0.07T - 1\times10^{-4}T^2 + 9.6\times10^{-5}T^3 \\ &+ 2.6\times10^{-7}T^4 - 2.8\times10^{-8}T^5 \end{aligned} \tag{11}$$

As a result, the first Grey-Box model for a given wind speed range of $w \in [-20, -10]$ m/s is obtained in equation (12). By replacing the parameters of model terms in Table II with functions of environment inputs: temperature T, vehicle mass m, road grade θ_r and wind speed w. The coefficients are constants represented by β to be trained.

In equation (12), vehicle speed v is denoted by x. These sub-parameters of β are from the vehicle dynamics and are illustrated in Table III. Note that β_{-1i} are obtained based on equation (11) for energy consumption on vehicle HVAC devices. Since rolling resistance coefficient f in equation (5) contains constant and a 2^{nd} order term of vehicle speed, both β_{0j} and β_{2j} are allocated in equations of β_0 and β_2 in equation (12). The gravity force is not affected by speed and to reduce the possible model fitting error, β_{03} is designed as part of β_0. The other sub-parameters are for air drag force. β_{06} and β_{12} are from the air drag force. β_{04}, β_{05}, β_{11}, and β_{13} are added for improving the model accuracy. β_{31} is the parameter for the 3^{rd} term of relative speed v_r.

TABLE III
SUB-PARAMETERS IN THE GREY-BOX MODEL

Energy usage	Related sub-parameters
Devices	β_{-1i} $(i = 1, 2..., 6)$
Rolling resistance	β_{0j}, β_{2j} $(j = 1, 2)$
Gravity	β_{03}
Air drag	β_{0k} $(k = 4, 5, 6)$, β_{1p} $(p = 1, 2, 3)$, β_{2q} $(q = 3, 4, 5), \beta_{31}$

For driving under the wind speed range of $w \in [-10, 20]$ m/s, the air drag force in Fig. 2 b) is fitted by a 2^{nd} order polynomial. Comparison in Fig. 2 b) shows that the fitted curve matches the true one well. Thus, the 2^{nd} order polynomial is good enough to represent the air resistance force in the Grey-Box model and (12) is still applicable by simply assigning $\beta_3 = 0$, leading to the second surrogate model. Therefore, the Grey-Box Kriging model of energy consumption for a unit driving distance is obtained in two regions for wind speed ranges of $w \in [-20, \ -10]$ m/s and $w \in [-10, \ 20]$ m/s.

IV. MODEL TRAINING

The overall model training process is shown in Fig. 3, where β_3 and the rest of βs are trained separately, since the vehicle cross-sectional area is fixed in lifetime so that the value of β_3 will not be affected by environmental factors. To obtain β_3, a data set of air drag force with vehicle and wind speed in S1 (see Fig. 3) is needed. In this paper, the training data is generated using the high-fidelity model in equation (5). To make this data practical, normal distribution measurement noise is added with zero mean and given variance. Recursive Least Squares (RLS) [18] method is applied to find optimal fit of β_3 in S2 (see Fig. 3). If there are multiple sets of data available, the mean value of fitted β_3 will be adopted. Then, with the fitted β_3, another set of observed data is required to train the β_{-1}, β_0, β_1, and β_2. The data set in S3 (see Fig. 3) is obtained using the same method as that for S1 but is used for the vehicle energy consumption as a function of vehicle speed and other environmental factors (such as temperature, mass, etc.). Using Kriging model training in S4 (see Fig. 3), parameter set of β_{-1}, β_0, β_1, and β_2 can be calculated based on a different combination of environment factors. And the sub-parameters (in Table III), representing relationships among βs and environment factors in equation (12), are solved by the RLS method. Note that the fitting for β_3 is only for the wind speed range of $w \in [-20, \ -10] m/s$, and otherwise $\beta_3 = 0$.

Since the Surrogate models for the two wind speed ranges are trained separately, large error may occur when the wind speed change causes switching between two models. To reduce the model switching error, the following procedure is taken in S3. Considering the Grey-Box model for wind speed $w \in [-10, \ 20]$ is designed heavily based on physics with higher fidelity than the one for the wind speed range of $w \in [-20, \ -10]$, the former one is adopted as the baseline model. Then a hysteresis is designed within wind speed $w \in [-12, \ 10]$ for the model with a wind speed range of $w \in [-20, \ -10]$. In the hysteresis wind speed range, the mean values of energy usage from observed data and the energy calculated from baseline model with identical inputs are used as training data. Then, the trained model will converge to the baseline model in the hysteresis range and the error is eliminated.

Considering that vehicle aging may reduce the accuracy of the energy consumption model for specific driving environments, a model updating method using current driving data is developed. The model updating process is similar to the structure in Fig. 3, but only goes through steps S3, S4, and S5 since β_3 is assumed to be fixed in the training process, and a forgetting factor is added to the RLS method in order to compensate for slow vehicle aging. The detailed RLS with forgetting factor and Kriging model training method is introduced in the following section.

A. Recursive Least-Squares (RLS)

To explain the used RLS method, β_1 calculation is used as an example. The regression function is $Y = \phi^T \theta$, where Y is the value of β_1; regressor $\phi = [1, \ w, \ w^2]^T$ and parameter vector $\theta = [\beta_{11}, \ \beta_{12}, \ \beta_{13}]^T$ to be estimated. The objective of RLS is to find parameter θ that minimizes the loss function,

$$L = \frac{1}{2} \sum_{i=1}^{n} \lambda^{n-i} [Y(i) - \phi^T(i) \hat{\theta}(i-1)]^2 \tag{13}$$

where λ is the forgetting factor and $\lambda \in (0, \ 1]$ which defines how fast the old data is diminished; the iteration step $i \in [2, \ n]$, where n is the number of environment factors. Then by solving equation (13), the recursive updating equation (14) can be obtained. The initial P is a diagonal unit matrix multiplied by a calibrated factor of 10^{-10}. The detailed derivation can be found in [19].

$$\hat{\theta}(i) = \hat{\theta}(i-1) + K(i)[Y(i) - \phi(i)\hat{\theta}(i-1)] \tag{14}$$

$$K(i) = P(i-1)\phi(i)(\lambda + \phi^T(i)P(i-1)\phi(i))^{-1} \tag{15}$$

$$P(i) = P(i-1)[I - \phi^T(i)K(i)]/\lambda \tag{16}$$

The estimated β_1 is updated at each iteration by equation (14). The forgetting factor $\lambda = 1$ when fitting β_3 in S3

$$
\begin{aligned}
y(x) = \beta_{-1}x^{-1} + \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3(x+w)^3 + z(x) = \beta^T f(x) + z(x) \\
\begin{cases}
\beta_{-1} = \beta_{-11} + \beta_{-12}T + \beta_{-13}T^2 + \beta_{-14}T^3 + \beta_{-15}T^4 + \beta_{-16}T^5, \\
\beta_0 = m(\beta_{01} + \beta_{02}T + \beta_{03}\sin\theta_r) + \beta_{04} + \beta_{05}w + \beta_{06}w^2, \\
\beta_1 = \beta_{11} + \beta_{12}w + \beta_{13}w^2, \quad \beta_2 = m(\beta_{21} + \beta_{22}T) + \beta_{23} + \beta_{24}w + \beta_{25}w^2, \quad \beta_3 = \beta_{31}
\end{cases}
\end{aligned}
\tag{12}
$$

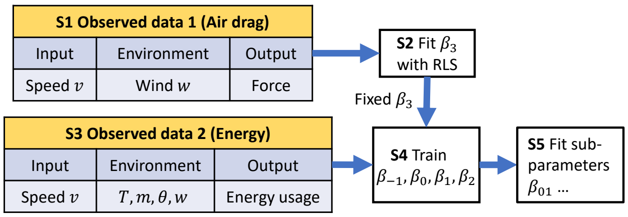


Fig. 3. Grey-Box model training process

and initial fitting of sub-parameters β_{01}, β_{02}, *etc.* in S5 of Fig. 3; and $\lambda = 0.9999$ is used for fitting sub-parameters in the model update process that forgets old data at a very slow rate. Besides, a saturation, which limits changes on current sub-parameters within $\pm 10\%$, is designed to keep these fitted sub-parameters within the normal range.

B. Kriging Model Training

As introduced in Section III, in order to handle the assumed Gaussian noise, the model training method of stochastic Kriging model is adopted in S4 of Fig. 3. The training process is developed based on methods introduced in [21] by modifying the Design and Analysis of Computer Experiments (DACE) MATLAB toolbox and SK extension package by Nielsen's group [20], [22]. Through the training process, parameter sets β_i ($i = -1,\ 0,\ 1,\ 2$) are obtained under corresponding environmental factors. The detailed solution of a single set of β_i is introduced as follows.

Recall the applied Kriging model in equation (7), where $y(x)$ is the observed output, the input is vehicle speed $x = v$, and vector $f(x)$ is known. Since the β_3 is fitted in advance, the 3^{rd} order term is removed from the output, the new observed output $y_n(x)$ is calculated by

$$
y_n(x) = y(x) - \beta_3(x+w)^3 \tag{17}
$$

where $f(x) = [x^{-1}, 1, x, x^2]^T$ and $\beta = [\beta_{-1}, \beta_0, \beta_1, \beta_2]^T$.

Extrinsic noise $z(x)$ of the Kriging model is assumed to be a Gaussian process with zero mean and certain covariance. Note that the covariance can be calculated by correlation functions based on the distances between all pairs of inputs x^j and x^k ($j, k \in [1, m]$), where m is the dimension of observed sample with a single environment factor set. The chosen squared exponential correlation function $Corr(j,k)$ is

$$
Corr(j,k) = \exp\left(-\sigma \left|x^j - x^k\right|^2\right) \tag{18}
$$

where σ is correlation parameter scaled by distance between points determined during fitting observed data with Gaussian process. Then, for stochastic Kriging model, the covariance matrix $R(j,k)$ contains extrinsic noise and independent intrinsic noise, a constant diagonal matrix, can be calculated by

$$
R(j,k) = \tau^2 Corr(j,k) + \gamma^2 I \tag{19}
$$

where τ is the variance of Gaussian process to be determined. Now, the three unknown parameters τ, σ, and β can be determined by maximizing the likelihood function (20) with the observed data; see [15], [17] for details.

$$
\begin{aligned}
\ln L(\tau, \sigma, \beta) &= -\frac{1}{2}\left[n\ln(2\pi\tau^2) + \ln(\det(R)) + \varepsilon\right] \\
\varepsilon &= (Y_n - F\beta)^T R^{-1}(Y_n - F\beta)/\tau^2
\end{aligned}
\tag{20}
$$

where $Y_n = [y_n(x^1), ..., y_n(x^m)]$ is the vector of m observed outputs and $F = [f^T(x^1), ..., f^T(x^m)]^T$ is a vector of polynomials with sample input under a single environment factor set. Once the likelihood function is maximized, the τ, σ, and β are obtained and the Kriging model is established. The trained β with corresponding environment factor set will be used to fit sub-parameters with RLS; see the sub-section above for details. Note that when the trained Kriging model (7) is used for energy estimation, the extrinsic noise $z(x)$ is set to zero since it is only considered during the training process of β.

V. SIMULATION VALIDATION

To validate the developed model training method, a Grey-Box model is trained using observed data and compared to the high-fidelity model (5). Case study co-simulations were conducted in MATLAB and SUMO for validating the Grey-Box model. Considering the real-world driving condition, the environmental factors are chosen from the ranges in Table IV. The test vehicle parameters are provided in Section II

TABLE IV
ENVIRONMENT FACTORS

Factor	Temperature(°C)	Mass(kg)	Grade(°)	Wind(m/s)
Range	$[-20, 50]$	$[1000, 1200]$	$[-4, 4]$	$[-20, 20]$

A. Model Training Validation

For this study, the Grey-Box model is trained using the data from the high-fidelity vehicle energy model (5) subject to environmental changes with added measurement noise (zero mean) with variance equal to 1% of maximum output; see Section IV. The peak of the noise is about $\pm 2\%$ of the model maximum output. The test input (vehicle speed) is generated by Latin hypercube method.

In this simulation, the Grey-box model is trained using the data from co-simulation mentioned above. To validate the accuracy of trained model in general case, the trained model is tested with a new set of random inputs different from the data used for model training. As a result, the trained model is validated by comparing the model outputs with new inputs and the corresponding calculation results of vehicle dynamic model. Besides, since the wind speed are separated into two ranges, two special scenarios in Table V are used for the co-simulations. The associated two-scenario results of driving

$500m$ are presented in Fig. 4. The drive distance could be different. Since the vehicle speed is constant and the noise of observed data is zero mean, the unit energy consumption can be rescaled to any distance. It can be seen from Fig. 4 that the energy consumption is high at low speed due to extended drive time over a unit distance, which leads to increased HVAC usage. Fig. 4 shows that the trained models match the true data well for both scenarios. And in Fig. 5(a), the maximum error of all cases is below 2.5% which is acceptable. Note that the error is not evaluated for these scenarios with close to zero energy consumption since those are considered to be singular cases.

Another validation study is conducted with experimental data, where the environmental factors are with varying road grade, fixed temperature at 20°C and zero wind speed. The Grey-Box model is trained with the experimental data and is validated with a different set of experimental data. The test result is shown in Fig. 5(b). It can be seen that these blue dots are around the 45° line, indicating that the Grey-Box model is able to provide good energy estimation of actual vehicle at steady-state. The estimation error is mainly caused by not operating the vehicle at steady-state and sensor noise.

TABLE V
ENVIRONMENT FACTORS

No.	T **(°C)**	$m(kg)$	$\theta_r(°)$	$w(m/s)$
a	0.3	1035	0.57	-14.3
b	41.3	1184	2.69	3.0

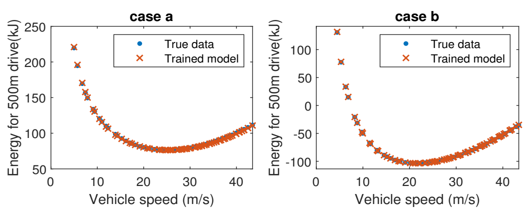
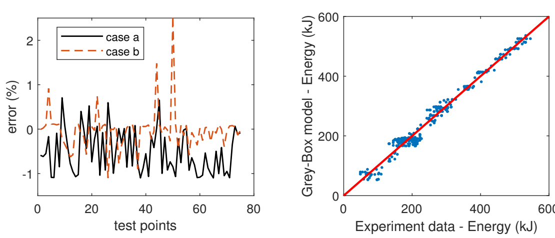


Fig. 4. Observed data vs. trained model in two scenarios

For validating the developed model updating scheme, a co-simulation scenario is designed under temperature $T = 15$°C, vehicle mass $1130kg$, road grade 0° and wind speed $5m/s$. This co-simulation studies the performance of model updating method when a series of new data points are obtained under current driving conditions. And in order to show the changes in observed data and vehicle performance (energy consumption), the deviation between initial and new observed data is set to be 1% of the energy consumption. The result of model update is shown in Fig. 6. It can be seen from the plot, these circled points (previously and newly observed data) depart from each other and the more crowded points are the newly observed data. In the zoomed part of Fig. 6, the updated model data points, marked by crossings, are located close to the new observed points and less on the previous data ones. Therefore, it is verified that the model update method is capable of updating the Grey-Box model to match the new vehicle driving data so that the updated Kriging model will be more accurate under current scenario than the old one. Especially, this updating process is able to compensate for vehicle aging. Meanwhile, it takes less than 20 seconds to complete the model updating process using a Windows computer with an i7-11700F CPU. Thus, the capability of real-time model updating is confirmed since the model update is relatively fast.



(a) Error with high fidelity model (b) Error with experimental data

Fig. 5. Accuracy validation of model training

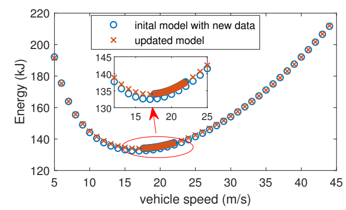


Fig. 6. Comparison of observed data and trained model in random situation

B. Grey-Box Model Evaluation

To evaluate the necessity of developing the proposed model, three co-simulations are carried out under fixed environments with one varying environmental factor for the vehicle driving $500m$ at a speed of $13m/s$. The simulation results are shown in Tables VI and VII. These comparisons are conveyed among three models, where the designed Grey-Box model is noted as the 'Grey-Box', the vehicle energy consumption model, called the 'Baseline model,' with environmental factors provides the true energy consumption in equation (5) and the commonly used basic energy consumption model derived from the vehicle dynamic model in [6] is denoted as the 'simple model' as shown in Tables VI and VII. Note that the 'simple model' uses the same notations as equation (5), but the energy consumption on vehicle subsystems P_d and wind speed w is not included with a fixed rolling resistance coefficient, since this model is for vehicles not capable of obtaining wind speed information, and the Grey-Box model does use this information. In addition, vehicle mass and road grade are not compared since all vehicle models consider these parameters.

$$E_u = [mg(f_r cos\theta_r + sin\theta_r) + 0.5C_D\rho Av^2]d_u \tag{21}$$

In Table VI, the temperature is set to be 15°C that is warm without heater on, and -5°C with heater on. The other factors are fixed as shown. It can be seen from the baseline results that vehicle consumes 35% more energy at -5°C, indicates that the temperature does have a significant effect on energy consumption. Note that turning HVAC system on for a certain period will significantly increase the total energy consumption, especially under cruising conditions when the steady-state operation energy consumption is low and energy consumption also goes up at low temperatures due to increased rolling resistance. The simulation results from Grey-Box model have an error of less than 2% which is much more accurate than the simple model with an error of around 35%.

Similarly, in Table VII, wind speed leads to a significant change in energy consumption of baseline model with down-wind of $-12m/s$, which uses 19.9kJ less energy than head-wind at $5m/s$. The Grey-Box model still offers high accuracy. However, the simple model has about 20% error, leading to poor estimations of energy usage and drive range. As a result, the driving environment significantly affects energy consumption and should be considered in the vehicle model.

TABLE VI
MODEL ENERGY CONSUMPTION UNDER TWO TEMPERATURES
$(m = 1000kg, \theta = 0°, w = 0m/s, distance = 500m)$

Temperature	Grey-Box	Baseline model	Simple model
15 °C	131.4 kJ	131.2 kJ	110.1 kJ
-5 °C	180.2 kJ	177.7 kJ	110.1 kJ

TABLE VII
MODEL ENERGY CONSUMPTION UNDER TWO WIND LEVELS
$(T = 15°C, m = 1000kg, \theta = 0°, distance = 500m)$

Wind speed	Grey-Box	Baseline model	Simple model
-12 m/s	124.6 kJ	122.0 kJ	108.4 kJ
5 m/s	141.7 kJ	141.9 kJ	108.4 kJ

To highlight the benefit of the proposed Grey-Box Kriging model, a special driving scenario is set up, where temperature is $T = 15°C$, vehicle mass $1200kg$, road grade 0.5°and wind speed $-15m/s$. And the road is configured with a length of $100m$, speed upper limit of $22.22m/s(50mph)$ and lower limit of $13.33m/s(30mph)$ (60% of the upper limit). In this test, three models: simple vehicle model with fixed parameters (accurate vehicle mass and road grade but fixed predefined rolling and aero-dynamic resistance coefficient), proposed Grey-Box model and baseline model (vehicle dynamic model with environment factors) are compared in Fig. 7.

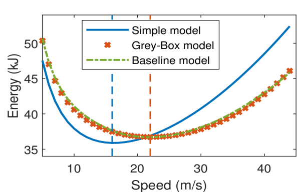


Fig. 7. Comparison of Grey-Box and fixed dynamic models

These three models result in different model curves as a function of vehicle speed, indicating that the performance of vehicle energy consumption models varies based on the given scenario. Comparing the energy curves of baseline and Grey-Box models, the energy curve of baseline model matches that of the Grey-Box model well, indicating that the Grey-Box model can represent the actual vehicle energy consumption. Note that the vehicle energy consumption model with environmental factors is considered to be the true model. By checking the minimum energy consumption points of the simple model with fixed parameters and the Grey-Box model, it can be seen that the optimal speed based on the simple model with fixed parameters is 16.2 m/s, while the optimal speed based on the Grey-Box model is 22 m/s that is higher than that from the simple model by considering low air drag force due to the negative wind speed (not considered in the simple model). This illustrates that the optimal speed for minimizing energy consumption varies under different vehicle operational environments and Fig. 8 further confirms it. Note that the proposed model is convex in speed.

In the simulation study of Fig. 8, the energy consumption is calculated by the developed Grey-Box model for driving $100m$. The environment factors are fixed with vehicle mass $m = 1200kg$, road grade $\theta_r = 0$, wind speed $w = 5m/s$, but the temperature varies in a range of $T \in [5, 35]$. Then, with vehicle speed from 5 to $44(m/s)$, the response surface of energy consumption is plotted in Fig. 8, which shows that the optimal speed (minimal energy consumption) changes as a function of environmental temperature.

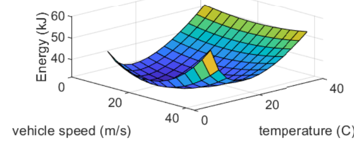


Fig. 8. Energy consumption with varying speed and temperature

These simulation results demonstrate that the developed Grey-Box model is able to reflect vehicle performance (energy consumption) with changing driving conditions and the proposed model is ready for minimizing energy consumption by selecting optimal vehicle route and speed. Since the proposed model is very simple, it is expected that real-time optimization is feasible. Note that the similar optimization process developed in [3] cannot be used for online applications due to the utilization of baseline vehicle dynamic model for optimization.

C. Case Study

A study of route planning problem is carried out by MATLAB and SUMO co-simulation to evaluate the importance of environmental effects to route and speed optimization based on the proposed Grey-Box model. The map in Fig. 9 is located in Graz, Austria, where traffic lights, vehicle acceleration and deceleration are active in this simulation with default SUMO settings to simulate real-world driving condition. Comprehensive studies will be part of our future work. The temperature is set to $-5°C$, vehicle mass $1000kg$, road grade 0°and wind blowing from north to south at $10m/s$ and assumed to be unchanged during the simulation. The wind speed for each road segment is updated within a limited area centered at the eco-vehicle ($2km$ radius for this study). The wind speed information will be updated as the eco vehicle moves. It takes about 7 seconds to update the route and speed, which is acceptable for optimization before the vehicle approaches the intersection. In this study, the aim is to find the most eco-route using the Dijkstra algorithm [24] based on energy consumption weightings on all streets. These weightings are calculated by the proposed Grey-Box model and the simple model with fixed environment parameters. The comparison is made in terms of the two routes selected by the Dijkstra algorithm based on energy consumption weights from the two models; see routes in Fig. 9, where the red route is from the simple model and the green one is from the Grey-Box model. The co-simulation results are shown in Table VIII.

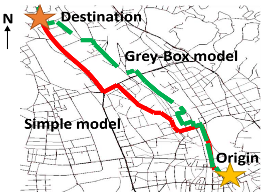


Fig. 9. Optimized routes based on simple and Gray-Box models

TABLE VIII
ROUTE SELECTIONS USING SIMPLE AND GRAY-BOX MODELS
($T = -5$°C, $m = 1000kg$, $\theta = 0°$, $w = 10m/s$, N TO S)

	Time	Distance	Consumption	Improve
Grey-Box	$546s$	$5.23km$	4.31 kJ	5.27 %
Simple model	$497s$	$5.14km$	4.55 kJ	0 %

In Table VIII, the route of Grey-Box model uses 4.31 kJ energy (5.27% less than 4.55 kJ simple). However, the route of Grey-Box model travels $0.9km$ more distance than the simple model and takes extra $49s$. The difference in route selection in this study is attributed to the consideration of environmental temperature and wind, where the temperature is set to -5°C and the heater is assumed to be on. This leads to consuming more energy if the driving duration is long, and as a result, these streets with high-speed limits are chosen. On the other hand, the wind increases energy consumption as vehicle speed goes up, and therefore, low speed streets are favorable under windy conditions. In this study, high gust wind encountered by the vehicle requires more energy to maintain the speed than the HVAC, the route chosen based on the Grey-Box model utilizes these streets with low-speed limits to reduce energy consumption. However, since the simple model does not consider the wind speed, it chooses the fastest route. This study indicates that energy consumption can be reduced by using the proposed Grey-box model for route optimization.

VI. CONCLUSION

In this paper, a Grey-Box model is proposed for estimating vehicle energy consumption of a unit distance as a function of vehicle speed under different environmental conditions, where the Kriging modeling method is used to train the Grey-Box model and recursive Least-Squares method is used for online updating. The simulation results of model training indicated that the developed model is able to match the high-fidelity dynamic model with an error that less than 2.5%. And the model update method, validated by a simulation study, indicates that the model can be updated online to enhance the model accuracy and compensate for vehicle aging. Comparing simple dynamic and the Grey-Box models shows that the proposed model is able to estimate energy consumption accurately for route and speed optimization, and in addition, the case study of route planning based on the proposed Grey-Box model shows an energy consumption reduction of 5.27%. However, the proposed Grey-Box model is developed based on the vehicle dynamic model, and energy consumption estimation could be affected by ignored factors and measurement errors. The future work is to study estimation errors caused by unmodeled factors and measurement errors to further improve the model accuracy for route and speed optimization.

REFERENCES

[1] Ericsson, et al. ”Optimizing route choice for lowest fuel consumption–potential effects of a new driver support tool.” Transportation Research Part C: Emerging Technologies 14.6 (2006): 369-383.

[2] K. Ahn and H. Rakha, “Field evaluation of energy and environmental impacts of driver route choice decisions,” in 2007 IEEE Intelligent Transportation Systems Conference, pp. 730–735, IEEE, 2007.

[3] Liu, Haiou, et al. ”Economic adaptive cruise control for a power split hybrid electric vehicle.” IEEE Transactions on Intelligent Transportation Systems 21.10 (2019): 4161-4170.

[4] Miao C, Liu H, Zhu G G, et al. Connectivity-based optimization of vehicle route and speed for improved fuel economy[J]. Transportation Research Part C: Emerging Technologies, 2018, 91: 353-368.

[5] Krozel, Jimmy, Changkil Lee, and Joseph SB Mitchell. ”Turn-constrained route planning for avoiding hazardous weather.” Air Traffic Control Quarterly 14.2 (2006): 159-182.

[6] Hua, Lingyun, et al. ”Minimal Energy Transient Motion Control of Electrical Connected Vehicles.” IEEE/ASME Transactions on Mechatronics 27.4 (2022): 2116-2124.

[7] Powell, B. K., K. E. Bailey, and S. R. Cikanek. ”Dynamic modeling and control of hybrid electric vehicle powertrain systems.” IEEE Control Systems Magazine 18.5 (1998): 17-33.

[8] Xie, Qingqing, et al. ”Advanced vehicle dynamic model for EV emulation considering environment conditions.” 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering. IEEE, 2017.

[9] Wei, Wenpeng, Hussein Dourra, and Guoming Zhu. ”Vehicle tire traction torque estimation using a dual extended kalman filter.” Journal of Dynamic Systems, Measurement, and Control 144.3 (2022).

[10] Pearce, et al. ”Energy conservation from systematic tire pressure regulation.” Energy Policy 35.4 (2007): 2673-2677.

[11] Kroll, Andreas. ”Grey-box models: Concepts and application.” New frontiers in comp. intelligence and its applications 57 (2000), pp. 42-51.

[12] Xia, Bin, Ziyan Ren, and Chang-Seop Koh. ”Utilizing kriging surrogate models for multi-objective robust optimization of electromagnetic devices.” IEEE transactions on magnetics 50.2 (2014): 693-696.

[13] Pal, Anuj, et al. ”Dynamic system identification for a nonlinear vehicle model using q-markov cover under different operational conditions.” 2021 American Control Conference (ACC). IEEE, 2021.

[14] Ejsmont, Jerzy, et al. ”Influence of temperature on the tyre rolling resistance.” Int. J. Automot. Technol. 19.1 (2018), pp. 45-54.

[15] Tang, Jian, et al. ”Stochastic Bayesian optimization for predicting borderline knock.” Int. J. of Engine Research 24.3 (2023): 793-807.

[16] Yuksel T, Michalek J J. Effects of regional temperature on electric vehicle efficiency, range, and emissions in the United States[J]. Environmental science & technology, 2015, 49(6): 3974-3980.

[17] Pal, Anuj, et al. ”Experimental Case Study of Stochastic Surrogate-Assisted Engine Calibration.” IEEE/ASME Transactions on Mechatronics 27.6 (2022): 4897-4907.

[18] Zhang, Wenlai, and Xingxuan Wang. ”System identification of miso second-order system based on least square method.” 2016 35th Chinese Control Conference (CCC). IEEE, 2016, pp. 9457-9461.

[19] Young, Peter C. Recursive estimation and time-series analysis: an introduction. Springer Science & Business Media, 2012.

[20] Lophaven SN, et al. DACE: a Matlab kriging toolbox. 2002. Vol. 2. IMM, Informatics and Mathematical Modelling, The Technical University of Denmark.

[21] Sobester, András, et al. Engineering design via surrogate modelling: a practical guide. John Wiley & Sons, 2008.

[22] Ankenman B, et al. Stochastic kriging for simulation metamodeling. In 2008 Winter Simulation Conference. New York: IEEE, pp. 362–370.

[23] Lopez, Pablo Alvarez, et al. ”Microscopic traffic simulation using sumo.” 2018 21st international conference on intelligent transportation systems (ITSC). IEEE, 2018.

[24] Javaid, Adeel. ”Understanding Dijkstra's algorithm.” Available at SSRN 2340905 (2013).

[25] Boriboonsomsin K, et al. Eco-routing navigation system based on multi-source historical and real-time traffic information[J]. IEEE Transactions on Intelligent Transportation Systems, 2012, 13(4): 1694-1704.