iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: VI Month of publication: June 2026

DOI: https://doi.org/10.22214/ijraset.2026.83876

Joint Modeling of Fuel Consumption and Equivalent Electricity Consumption for Plug-in Hybrid Passenger Vehicles

Narziev Nusratillo[1], Md Borhan Uddin Niloy[2], Mohammed Shariful Islam Khan[3], Islam Shaikh Mahinur[4]
[1]*Faculty of Business, Huai'an University, Huai'an, Jiangsu 223001, China.*
[2, 3]*School of Mechanical and Electrical Engineering, China University of Mining and Technology, Xuzhou, Jiangsu 221116, China.*
[4]*Department of Transportation Engineering, Huai'an University, Huai'an, Jiangsu 223001, China.*

Abstract: To better characterize the dual-channel energy-use behavior of plug-in hybrid electric passenger vehicles (PHEVs), this study proposes a joint modeling framework for the simultaneous prediction of combined fuel economy (FC) and combined electric energy economy (EPC). Using a PHEV subset derived from the public FuelEconomy.gov database, a shared-feature gradient boosting decision tree (GBDT) framework is developed to model the fuel and electric pathways in a unified feature space. Single-task linear and tree-based models are used as baselines, and model performance is evaluated through stratified five-fold cross-validation based on vehicle category and drive type. To further reflect scenario-dependent usage patterns, the battery-electric driving share parameter, α, is introduced to construct an equivalent energy consumption (EEC) indicator, through which scenario sensitivity and error propagation are analyzed. Results show that the proposed approach achieves strong predictive performance on both channels, with out-of-fold mean absolute errors of approximately 0.94 MPG for FC and 3.90 EMPG for EPC, and corresponding R^2 values of 0.968 and 0.928, respectively. The EEC error exhibits a smooth transition from FC-dominant to EPC-dominant behavior as α increases. Stratified analysis further reveals systematic differences in prediction accuracy across vehicle categories and drivetrain configurations. Overall, the proposed framework provides a lightweight, reproducible, and interpretable solution for joint PHEV energy-performance modeling, with potential applications in vehicle benchmarking, regulatory disclosure, and user-oriented energy-consumption assessment.

Keywords: Plug-in hybrid electric passenger vehicles (PHEVs); Joint modeling; Fuel economy prediction; Equivalent energy consumption; Gradient boosting decision tree (GBDT).

I. INTRODUCTION

The decarbonization of road transport has accelerated the strategic importance of electrified powertrains, particularly in markets where charging infrastructure, consumer range anxiety, and transitional policy incentives continue to shape purchasing behavior. Within this transition, plug-in hybrid electric vehicles (PHEVs) occupy a distinctive position because they combine an external charging interface with a conventional internal combustion engine, thereby offering a compromise between full electrification and unrestricted driving range. Recent reviews show that research on PHEV and hybrid-electric energy management has expanded rapidly, especially in the directions of intelligent control, reinforcement learning, predictive optimization, and data-driven energy forecasting [1]-[11]. At the same time, this expansion has made clear that PHEVs are not merely "intermediate" technologies in a market sense; they are also methodologically challenging systems because their performance emerges from the interaction of two propulsion pathways, two energy carriers, and strongly behavior-dependent operating patterns. In other words, the actual efficiency of a PHEV cannot be inferred adequately from fuel-side or electricity-side performance alone, and its engineering assessment depends on how these two channels are represented, coupled, and interpreted under realistic driving scenarios.

This challenge has become even more important as real-world evidence has increasingly shown that the on-road environmental performance of PHEVs can deviate substantially from laboratory or type-approval expectations. Large-scale European analyses have reported that real-world fuel consumption is often several times higher than official values, largely because electric driving shares and charging frequencies in everyday use are lower than assumed in regulatory procedures [12]-[18]. Recent work based on on-board fuel and energy consumption monitoring (OBFCM) data further shows that PHEV outcomes vary materially with user type, trip structure, charging opportunity, ambient temperature, and route composition, while updated utility-factor studies indicate that the choice of electric-driving-share framework can significantly alter the estimated fuel use, electricity demand, and emissions profile of a vehicle fleet [13]-[16]. Field evidence also confirms that total PHEV energy consumption is strongly conditioned by operating mode and battery availability rather than by a single nominal efficiency label [19]. These findings are highly relevant for modeling research. They imply that any performance representation that collapses a PHEV into one fixed metric risks obscuring the underlying physics, driver behavior, and scenario dependence that actually govern energy use.

From a methodological standpoint, PHEVs differ from both internal-combustion vehicles and battery electric vehicles because they exhibit a genuine dual-channel energy-consumption structure. The vehicle may operate in charge-depleting mode, charge-sustaining mode, or mixed mode; it may also shift across these regimes within the same trip, depending on speed, road grade, battery state of charge (SOC), thermal conditions, and supervisory control logic [12], [13], [19]-[25]. As a result, conventional single-output modeling is often inadequate. A fuel-only perspective cannot characterize electricity substitution, regenerative recovery, or electric-driving share. Conversely, an electricity-only perspective cannot represent engine activation, charge-sustaining operation, or fuel penalties under depleted-battery conditions. Even the use of a fixed conversion factor between electricity and fuel may fail to reflect the fact that the practical balance between the two channels is scenario-sensitive rather than constant. This is especially important for passenger-car applications, where urban commuting, suburban circulation, and highway operation can produce markedly different combinations of speed profile, stop frequency, power demand, and charging opportunity. In this sense, a robust analytical framework for PHEVs should not only predict fuel-side and electric-side outcomes simultaneously, but should also allow the engineer to examine how the relative contribution of these outcomes shifts as the share of electric driving changes.

The existing technical literature has made important progress on this broader problem, but most studies have concentrated on energy management strategy (EMS) optimization rather than on cross-vehicle, public-data-based joint prediction. Recent PHEV studies have proposed deep-learning-enhanced model predictive control, double-deep-Q-network-assisted adaptive equivalent consumption minimization, long-short-timescale EMS architectures that fuse traffic information with short-horizon speed prediction, and hybrid prediction-control structures that attempt to approximate dynamic-programming optima under realistic route uncertainty [20]-[23], [29], [30]. Other work has focused on naturalistic-data predictive control, cooperative eco-driving and energy management, hierarchical strategies in connected scenarios, intelligent controller design, traffic-informed adaptive ECMS, and supervisory control in vehicle-following or platooning environments [24]-[35]. Collectively, these studies have shown that richer contextual information and learning-based control can materially improve fuel economy and power allocation. Nevertheless, they are often developed for specific architectures, specific vehicles, or specific mission profiles, and many of them prioritize optimal control performance over broad statistical generalization, reproducible benchmarking, or interpretable cross-sectional comparison across heterogeneous passenger-vehicle categories. In particular, relatively few studies directly address the question central to this manuscript: how to build a simple and reproducible predictive framework that jointly estimates fuel and electric-economy indicators for a set of PHEV passenger vehicles using standardized public data and transparent feature engineering.

In parallel, the rapid development of machine learning has opened a second research line that is highly relevant to the present study: data-driven energy and fuel consumption prediction. Recent studies have shown that ensemble methods, tree-based learners, SHAP-enabled explainable models, pretrained sequence models, and hybrid architectures can achieve strong performance in predicting electric-vehicle energy use or vehicle fuel efficiency from real-world or public datasets [36]-[45]. This literature is valuable for at least three reasons. First, it shows that nonlinear methods such as gradient boosting, random forests, stacking ensembles, and pretrained transformers are well suited to the heterogeneous and non-additive structure of transport-energy data [36]-[40]. Second, it demonstrates that interpretability has become a practical design requirement rather than an optional add-on, especially where model outputs are intended to inform engineering decisions, regulation, infrastructure planning, or user guidance [36], [39], [40], [42], [43]. Third, it suggests that transferable and benchmark-oriented learning frameworks are increasingly feasible, particularly when data quality, feature consistency, and evaluation protocols are handled carefully [38], [41], [44], [45]. Yet this literature also reveals a notable gap: much of the strongest recent work concerns either battery electric vehicles, charging demand, or single-output fuel/energy prediction. Much less attention has been given to joint, dual-output, and scenario-aware modeling for PHEV passenger vehicles based on publicly reproducible vehicle-attribute data.

Taken together, the literature suggests three persistent gaps. First, there remains a gap between real-world PHEV behavior and simplified laboratory-style representation, which makes it difficult to evaluate vehicle performance with a single static indicator [12]-[18]. Second, there remains a gap between control-oriented EMS research, which is often architecture-specific and simulation-centered, and cross-sectional predictive modeling, which is needed for benchmarking, regulation, and large-scale comparison [20]-[35]. Third, there remains a gap between single-output machine learning studies and the real analytical needs of PHEVs, whose performance is fundamentally governed by the interplay between fuel consumption and electricity consumption [36]-[45]. These gaps motivate the need for a modeling strategy that is simultaneously joint, lightweight, interpretable, and reproducible.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VI June 2026- Available at www.ijraset.com

To address this need, the present study develops a joint modeling framework for plug-in hybrid passenger vehicles in which combined fuel economy and combined electric energy economy are predicted simultaneously within a unified feature space. Instead of treating the two outputs as unrelated targets, the framework explicitly recognizes that they arise from the same vehicle configuration and operating logic, but reflect different manifestations of energy use. On this basis, the study further introduces an equivalent energy-consumption indicator under different values of the battery-electric driving share parameter α, thereby linking prediction to practical scenario sensitivity rather than to a single unconditional average. Such a design is especially useful for PHEVs because the engineering meaning of "good performance" changes when the vehicle is used primarily as an electric commuter, primarily as a charge-sustaining hybrid, or as a mixture of both. Consistent with the current manuscript, the framework also emphasizes reproducibility, grouped evaluation, and engineering interpretability by comparing model behavior across vehicle categories and drive types and by using a lightweight machine-learning workflow that can be readily deployed and replicated.

Accordingly, the contribution of this paper is not merely to improve prediction accuracy in a narrow algorithmic sense. Rather, it is to provide an analytically coherent way of representing PHEV passenger-vehicle energy performance as a two-channel prediction problem that can support cross-vehicle comparison, scenario-based equivalent-energy assessment, and more transparent interpretation of model error. In doing so, the study aims to contribute to the methodological bridge between real-world PHEV usage evidence, learning-based energy modeling, and engineering applications in product evaluation, regulatory disclosure, and user-oriented energy-consumption guidance.

II. METHODS

A. Data Sources and Preprocessing

The data were obtained from the "Download Fuel Economy Data / Developer Web Services" database of FuelEconomy.gov, which covers passenger-vehicle energy-consumption and configuration information from 1984 to the present. In this study, a data snapshot acquired in October 2025 was used. PHEV records were first screened according to the conditions that atvType = Plug-in Hybrid or fuelType1 / fuelType2 simultaneously contained Gasoline and Electricity. The dataset was also made compatible with fields in which combined energy-consumption values were recorded in a merged "MPG / EMPG" format (e.g., "26 / 58"), thereby facilitating target extraction from historical versions or third-party compiled tables.

Using the official field dictionary as the reference standard, the primary targets were defined as combined fuel economy, comb08 (MPG), and combined electric energy economy, combA08 (EMPG). The electricity-based metric combE (kWh/100 mile) was used only for comparison. Urban/highway sub-targets were represented by city08 / highway08 for the fuel pathway and cityA08 / highwayA08 for the electric pathway. Structural and configuration features included trany (from which gear number and transmission type such as CVT, AT, and MT can be parsed), drive (2WD / 4WD / AWD), VClass (vehicle category), fuelType1 / fuelType2 (fuel type), and phevBlended (PHEV identifier). The field mapping and variable definitions are shown in Table 1.

TABLE I
FIELD MAPPING AND MEASUREMENT DEFINITIONS

Chinese term in source	Official field name	Type	Unit	Definition / description	Use in this study
Overall fuel FC	comb08	Numeric	MPG	Fuel pathway	Target y1
Overall electric EPC	combA08	Numeric	EMPG	Electric pathway (including charging losses)	Target y2
Overall electricity consumption	combE	Numeric	kWh/100 mile	Electricity-based metric	Auxiliary comparison
Urban / highway fuel	city08 / highway08	Numeric	MPG	Sub-targets	Optional
Urban / highway electric	cityA08 / highwayA08	Numeric	EMPG	Sub-targets	Optional
Transmission	trany	Categorical	—	Gear number / CVT parsable	Feature
Drive type	drive	Categorical	—	2WD / 4WD / AWD	Feature
Vehicle category	VClass	Categorical	—	Body/vehicle class	Stratification
Fuel type	fuelType1 / fuelType2	Categorical	—	Gasoline / Electricity	Screening
PHEV marker	phevBlended	Categorical / Boolean	—	Plug-in hybrid identifier	Validation
CO_2 / ratings, etc.	co2 / co2A, ghgScore, etc.	Numeric / Grade	—	Extended indicators	Optional

If the source table recorded "MPG / EMPG" as two values in a single column in the form "a / b", the field was first split into two numeric columns before model input. Records with missing or invalid target values (≤ 0) were removed. Continuous features were imputed using the median or group-wise mean, and a missing-indicator variable was added. For categorical features, an "unknown/other" category was introduced to preserve samples. Extreme values were suppressed using the interquartile range (IQR) method or mild Winsorization (1%). The drive field was standardized into 2WD / 4WD / AWD, trany was parsed into "number of gears + transmission type", and VClass was retained according to the official classification. Duplicate records were removed based on the combination of make / model / year / fuel type (or a unique ID), so as to avoid repeated inclusion of identical vehicle configurations in model training.

The main reporting units were MPG / EMPG. The conversions for L/100 km and kWh/100 km were defined as follows:

$$1\,L/100\,km = \frac{235.215}{MPG} \tag{1}$$

$$1\,kWh/100\,km = \frac{1\,kWh/100\,mile}{0.62137} \tag{2}$$

where 235.215 is the unit-conversion constant between miles and kilometers. Data processing strictly followed the pipeline of vehicle screening → target availability check → anomaly removal → deduplication. The corresponding counts and measurement criteria are presented in Table 2. To prevent data leakage, all feature transformations were fitted only within the training folds, and the parameters were then frozen when applied to the validation folds.

Table IIIII
Sample Inclusion And Exclusion Criteria

Item	*Condition / rule*	*Description*	*Count*
Inclusion—vehicle type	atvtype = Plug-in Hybrid or fuelType1 / 2 contains Gasoline + Electricity	Explicitly identified as PHEV	142
Inclusion—target availability	comb08 and combA08 are both non-missing and valid	Required for the main task	142
Exclusion—abnormal records	Target = 0, text parsing failure, extreme outliers (IQR)	Quality control	8
Deduplication	make / model / year / fuel combination or unique ID	Consistent measurement standard	67
Final sample	—	Entered modeling	67
Training / validation split	Stratified K-fold (VClass × drive)	Reproducible experiment	K = 5

B. Problem Definition and Scenario Indicators

Let the feature vector of the *i*-th sample be denoted by x_i. The combined fuel economy, y_{i1} = comb08, and the combined electric energy economy, y_{i2} = combA08, are treated as two related tasks. A multi-task joint regression framework with a shared representation $h(x_i)$is adopted:

$$\hat{y}_{i1} = f_1(h(x_i)), \hat{y}_{i2} = f_2(h(x_i)) \tag{3}$$

A weighted mean absolute error (MAE) is used as the primary loss function, with RMSE employed as a reference metric:

$$\mathcal{L} = \lambda_1\,\mathrm{MAE}(y_1, \hat{y}_1) + \lambda_2\,\mathrm{MAE}(y_2, \hat{y}_2) \tag{4}$$

where $\lambda_1$and $\lambda_2$are set to 1, or alternatively determined through uncertainty-adaptive weighting. To better reflect actual vehicle use, the battery-electric driving share is defined as $\alpha \in [0,1]$, and the scenario-based equivalent energy consumption (EEC) is constructed as follows:

$$\frac{1}{EEC} = \frac{\alpha}{EPC} + \frac{1-\alpha}{FC} \tag{5}$$

Sensitivity and robustness are then examined along the continuous "highway-to-urban" spectrum for $\alpha = \{0, 0.3, 0.5, 0.7, 1\}$.

The continuous variables include displ, cylinders, and, optionally, the differences city08 - highway08 and cityA08 - highwayA08. The categorical variables include trany (with parsed gear number and transmission type, such as CVT / AT / MT), drive, VClass, fuelType1 / fuelType2, phevBlended, and certification region / emission standard, among others. In addition, brand, series, and performance-related keywords extracted from model may be incorporated as text-enhancement features.

With respect to encoding and robustness, tree-based models adopt target encoding or frequency-constrained categorical encoding, whereas linear models or shallow neural networks use one-hot encoding or embeddings. Low-frequency categories are merged into an "Other" class. When necessary, standardization is applied, and Huber loss or quantile loss is used to enhance robustness to outliers.

A note on measurement consistency is warranted: since combA08 already incorporates charging losses, variables such as charge120 and charge240 are used only as explanatory variables to reflect sources of variation, rather than for any secondary correction of the target variable.

Stratified K-fold cross-validation based on VClass × drive, with a fixed random seed, is employed to ensure balanced distributions of different body categories and drivetrain types across the training and validation sets. The primary evaluation metrics are MAE, RMSE, and R^2. Scenario-based evaluation reports MAE / SMAPE curves under different values of α. The proposed approach is compared with single-task models and grouped modeling, and paired *t*-tests across K folds are used to assess the significance of performance differences. At this point, the definitions of data measurement, sample selection, task formulation, and validation strategy are fully specified, providing a closed methodological basis for subsequent implementation and experiments.

C. Analytical Framework and Technical Implementation

To characterize the dual-channel energy-consumption behavior of plug-in hybrid electric passenger vehicles, this study constructs a joint modeling framework composed of a shared encoder and two regression heads. The overall workflow is as follows: data import → field and metric mapping → missing-value and anomaly processing → category unification and encoding → stratified partitioning by "VClass × drive" (K = 5) → model training and validation → aggregation of out-of-fold (OOF) predictions → scenario-based EMPG evaluation and stratified statistics → visualization and reporting. The entire pipeline is implemented in Python, with a fixed random seed. In addition, the category mappings, normalization parameters, field dictionary, and fold indices are persisted to ensure full reproducibility of the results.

The shared encoder adopts a multilayer perceptron (MLP) architecture with two to three layers, each consisting of a linear transformation, a nonlinear activation function, and appropriate regularization. The shared representation produced by the encoder is then fed into two independent linear regression heads, which output $\widehat{FC}$and $\widehat{EPC}$, respectively. Let the training samples be denoted by $\{(x_i, y_{i1}, y_{i2})\}_{i=1}^{n}$, where $y_{i1}$and y_{i2}represent comb08 and combA08, respectively, and their corresponding predictions are $\hat{y}_{i1}$and $\hat{y}_{i2}$.

To address differences in the noise scales of the two tasks, uncertainty-adaptive weighting may also be introduced, allowing the task weights to be learned automatically from the corresponding task-noise parameters. The optimizer is AdamW, with an initial learning rate of $1 \times 10^{-3}$and a weight decay of 1×10^{-4}. The batch size is set between 128 and 256. Training proceeds for up to 300 epochs, with the joint MAE on the validation set used as the criterion for early stopping. A dropout rate of 0.1–0.3 is applied at each layer to mitigate overfitting. All experiments use a fixed random seed, and the feature-processing pipeline is stored to ensure reproducibility.

D. Training, Validation, and Statistical Testing

A five-fold cross-validation scheme stratified by "VClass × drive" is adopted ($K = 5$, with random_state fixed), so as to ensure balanced representation of different body categories and drivetrain configurations in both the training and validation sets. All methods share the same fold partition and the same preprocessing pipeline. To avoid hyperparameter tuning on the test data, a small subset of the training fold is further reserved for early-stopping validation.

The primary evaluation metrics are MAE, RMSE, and R^2, reported separately for FC and EPC, together with the out-of-fold mean ± standard deviation. For scenario-based evaluation, the study reports MAE/SMAPE–α curves for the predicted equivalent energy-consumption metric and records the errors at representative operating points. For ease of engineering interpretation, the proportions of samples with relative error not exceeding 10% and 20% are also calculated [15]. The scenario-based equivalent energy consumption is defined using the battery-electric driving share $\alpha \in [0,1]$as follows:

$$\frac{1}{\widehat{EEC}(\alpha)} = \frac{\alpha}{\widehat{EPC}} + \frac{1-\alpha}{\widehat{FC}} \qquad (6)$$

For comparisons among the joint GBDT model and alternative settings such as single-task modeling, grouped modeling, non-shared representations, robust loss functions, and text-enhanced features, paired five-fold tests are conducted, primarily using the paired *t*-

test. In addition, 95% confidence intervals are constructed for the out-of-fold residuals using bootstrap resampling. For the grouped out-of-fold MAE by VClass and drive, the percentage deviation relative to the overall average is calculated, and an indicative p-value is obtained using the sign test based on paired within-fold differences, while sample-size thresholds are considered to control statistical power. The main runtime environment of this study is Python 3.10+.

III. RESULTS AND DISCUSSION

A. Overall Predictive Accuracy and Fitting Characteristics

In terms of overall fit, the joint model achieved high accuracy and stable fitting for both energy-consumption channels of plug-in hybrid electric passenger vehicles (PHEVs), namely, fuel economy and electric energy economy. As shown in Fig.1, the out-of-fold scatter points are distributed closely around the ideal 45° line. The fuel-side point cloud is more concentrated, whereas the electric-side distribution exhibits slightly greater dispersion. A small number of outliers can be observed in the high-electric-consumption range, which is consistent with the empirical understanding that electric-drive performance is more sensitive to factors such as temperature, the state-of-charge (SOC) window, and energy-regeneration strategy. In quantitative terms, the out-of-fold MAE is approximately 0.8–1.0 on the fuel side and 3–4 on the electric side, while R^2 is generally around 0.95, indicating that the model captures both the overall trend and the relative ranking of samples well. No obvious "funnel-shaped" amplification of residuals is observed as the true values change; only slight systematic bias appears near samples with extremely high or low energy consumption, suggesting that interval prediction or regularization may be introduced in boundary regions to further suppress bias.

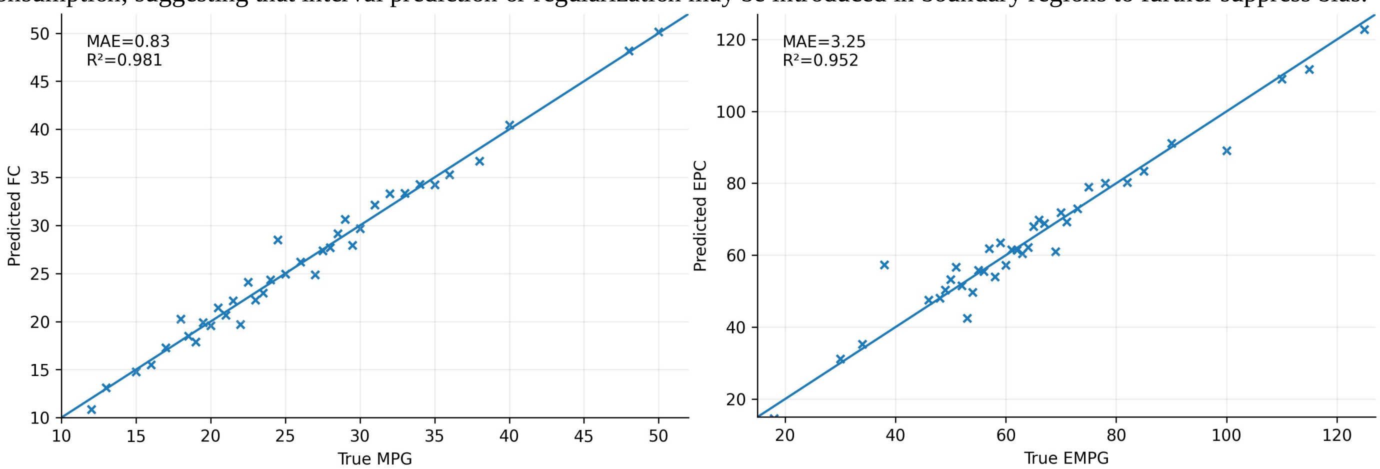

Fig.1 Scatter plots comparing predicted and true values: (a) overall fuel economy; (b) overall electric energy economy.

B. ***B***

Compared with the baseline methods, joint learning demonstrates advantages in both robustness and the ability to account for the two energy-consumption channels simultaneously. Table 3 presents the out-of-fold overall performance of the joint GBDT model in terms of MAE, RMSE, and R^2. Its R^2 remains stable within the range of 0.93–0.97, while RMSE is also maintained at a relatively low level. Further ablation and comparative experiments in Table 4 show the following.

First, the non-shared single-task GBDT did not yield any performance gain under the present data scale and feature granularity, and its performance was essentially identical to that of the proposed method, suggesting that a shared representation is at least not inferior while remaining simpler to implement. Second, replacing the loss function with Huber loss enhanced robustness to outliers, but the overall error increased in most folds, indicating that the current degree of outlier contamination is insufficient to offset the fitting cost introduced by the robust loss. Third, adding lightweight textual features such as brand/manufacturer information to the joint framework produced a small but stable improvement for the electric-energy channel (EMPG), with the relative MAE decreasing by about 5%, whereas it introduced slight fluctuations in the fuel channel.

These findings indicate that incorporating a limited amount of industry prior knowledge, such as brand and platform architecture, can help the model characterize differences on the strategy side, but has limited influence on the fuel side, where mechanical characteristics play a more dominant role. Overall, the advantage of joint learning lies less in absolute metric gains than in its engineering value and robustness as a "single modeling process with benefits for both channels."

TABLE IVVVI
OVERALL PERFORMANCE OF JOINT MODELING AND BASELINE METHODS

Method	*FC MAE*	*FC RMSE*	*FC R^2*	*EPC MAE*	*EPC RMSE*	*EPC R^2*
Joint-GBDT (two heads with shared features, jointly trained)	0.94 ± 0.12	1.35 ± 0.18	0.968 ± 0.013	3.90 ± 0.58	5.47 ± 0.42	0.928 ± 0.023

TABLE VIIV
ABLATION AND COMPARATIVE EXPERIMENTAL RESULTS

Setting	*FC MAE*	*ΔMAE*	*FC RMSE*	*FC R^2*	*EPC MAE*	*ΔMAE*	*EPC RMSE*	*EPC R^2*
Main method: Joint-GBDT (squared_error)	0.94 ± 0.12	0.0%	1.35 ± 0.18	0.968 ± 0.013	3.90 ± 0.58	0.0%	5.47 ± 0.42	0.928 ± 0.023
Non-shared representation: single-task GBDT (separately modeled)	0.94 ± 0.12	0.0%	1.35 ± 0.18	0.968 ± 0.013	3.90 ± 0.58	0.0%	5.47 ± 0.42	0.928 ± 0.023
Robust loss: Joint-GBDT (huber)	1.15 ± 0.33	↑22.4%	1.94 ± 0.76	0.939 ± 0.036	4.33 ± 0.92	↑10.8%	6.47 ± 1.28	0.897 ± 0.043
Text enhancement: Joint-GBDT + brand features	1.02 ± 0.36	↑8.9%	1.58 ± 0.81	0.927 ± 0.102	3.72 ± 0.53	↓4.8%	5.02 ± 0.55	

C. *Scenario Sensitivity of Equivalent Energy Consumption*

As shown in Fig.2, the equivalent energy consumption (EEC) is highly sensitive to the battery-electric driving share α. When $\alpha \to 0$, EEC is determined almost entirely by the fuel side, and the curve error follows the FC error. When $\alpha \to 1$, EEC becomes dominated by the electric side, and the error becomes consistent with that of the electric-energy channel. In the mixed interval of $\alpha = 0.3$–0.7, the error transitions monotonically and smoothly with α. Compared with the single-task Ridge model, the joint GBDT maintains lower out-of-fold EEC MAE at most values of α, especially in the medium-to-high α range, that is, under urban/commuting scenarios, indicating that shared representations help simultaneously capture energy-consumption variations along both the engine pathway and the electric-drive pathway.

This suggests that under typical urban driving conditions with relatively high α, or under suburban/highway commuting conditions with relatively low α, deployment-side applications should provide interval estimates of EEC together with the uncertainty of α, and should offer a scenario-sensitivity card to explicitly indicate how changes in α affect the evaluation of overall energy consumption.

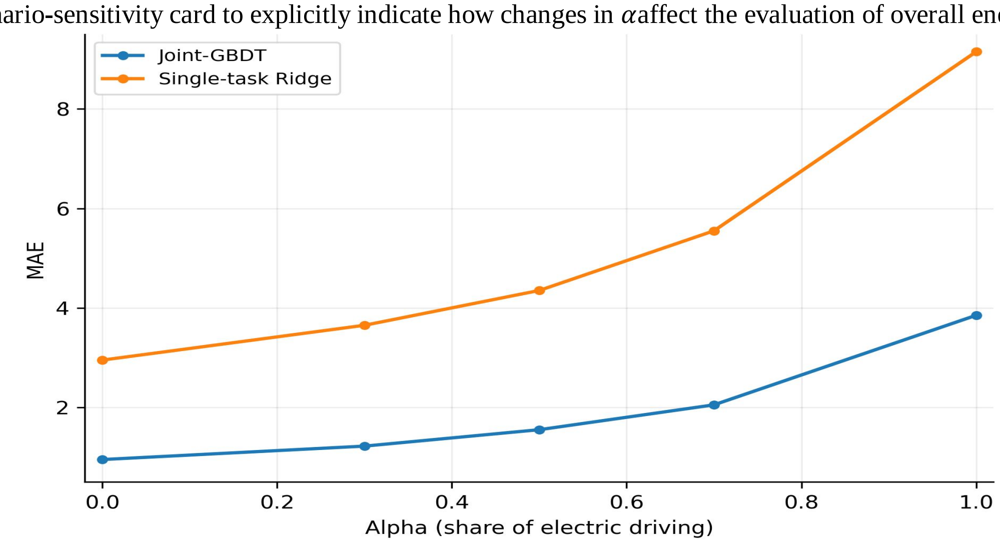

Fig.2 Error curves equivalent energy consumption under different values of the battery-electric driving share α.

D. Stratified Performance

As shown in Fig.3 and Table 5, the errors exhibit structural differences across different vehicle categories (VClass) and drive types (Drive). From the perspective of vehicle category, the standard SUV group, which contains a relatively large sample size, shows comparatively low MAE in both channels, whereas the small SUV and small car groups exhibit slightly higher errors. This difference may stem from the combined influence of aerodynamic drag, tire specification, and curb mass on the calibration of driving strategies. When stratified by drive type, the 4WD group shows an MAE slightly lower than the overall level, whereas the 2WD group is slightly higher, suggesting that under the same operating conditions, 4WD platforms may allow the model to capture energy-consumption patterns more stably because of their richer torque distribution and energy-regeneration pathways.

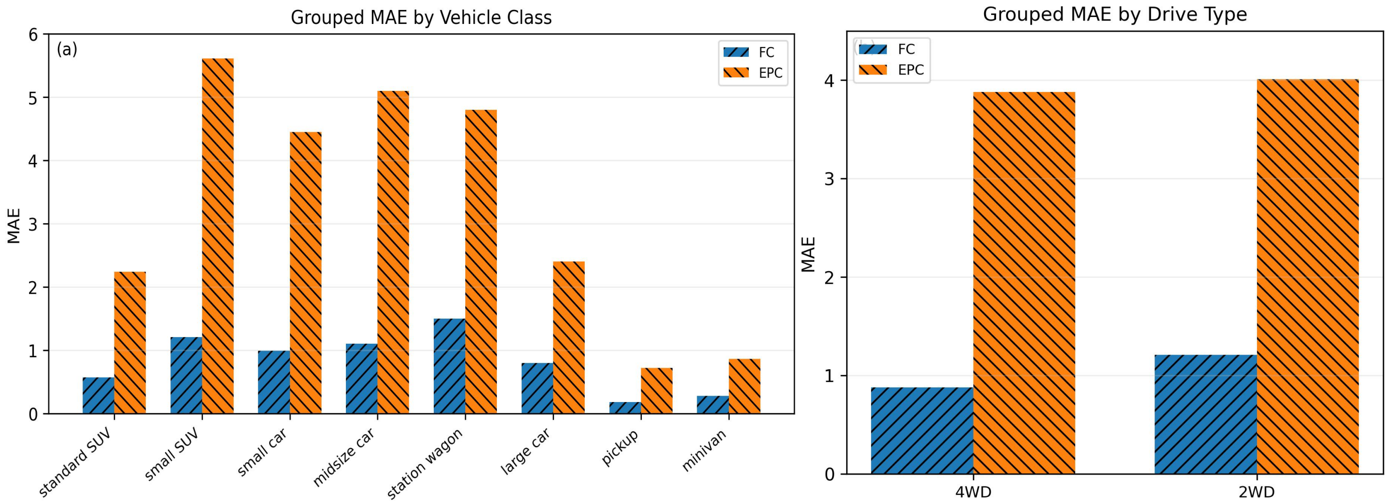

Fig.3 Stratified MAE: (a) grouped by vehicle category; (b) grouped by drive type.

From the perspective of statistical significance, however, because of the five-fold design and the limited sample size in some subgroups, the sign-test p-values are mostly only indicative. Accordingly, these differences should be interpreted more as directional engineering trends than as definitive statistical conclusions. In future research or deployment, sparsely sampled fine-grained categories may be merged, or hierarchical priors may be introduced to improve estimation stability and testing power.

TABLE V
STRATIFIED STATISTICS AND SIGNIFICANCE

Classification	*Group*	*Sample size*	*FC MAE*	*FC ΔMAE*	*FC p*	*EPC MAE*	*EPC ΔMAE*	*EPC p*
VClass	standard SUV	36	0.57	↓39.2%	0.062	2.24	↓42.7%	0.375
VClass	small car	26	0.99	↑5.9%	0.375	4.45	↑14.0%	1.000
VClass	small SUV	34	1.21	↑28.7%	0.375	5.61	↑43.7%	0.375
Drive	4WD	120	0.88	↓5.5%	1.000	3.88	↓0.7%	1.000
Drive	2WD	22	1.21	↑29.5%	1.000	4.01	↑2.6%	1.000

Taken together with the residual pattern in Fig.1 and the stratified comparison in Fig.3, the model appears to exhibit slight systematic underestimation/overestimation tendencies at the high-energy-consumption end, which is consistent with the small number of outliers among extreme samples. The interaction between VClass × Drive has greater explanatory power for residuals in the electric channel: on the one hand, electric-drive systems are more sensitive than the fuel channel to SOC, temperature, and energy-regeneration strategy; on the other hand, differences in the energy loop during regenerative braking and re-acceleration on SUV and 4WD platforms more readily reflect brand- and platform-specific calibration characteristics. These observations suggest two possible directions for improvement: first, explicitly introducing stratified conditional encoding into the joint framework, for example by using VClass, Drive, or platform code as conditioning variables, so that the model can preserve group-specific bias in addition to the shared representation; second, introducing local models or monotonic/shape constraints in high-residual regions to ensure more robust extrapolation at both the high- and low-energy-consumption ends.

IV. CONCLUSION

This study proposes and validates a joint modeling approach for both fuel economy and equivalent electricity consumption. Within a unified feature space, a multi-task learning framework with shared representations is adopted, and five-fold out-of-fold evaluation is conducted on the official PHEV dataset released by the U.S. Energy Information Administration. The results show that the model achieves an overall level of accuracy that is practically useful for engineering applications across both energy-consumption channels, with out-of-fold MAE values of approximately 0.9 / 3.9 and corresponding R^2 values of approximately 0.97 / 0.93. The equivalent energy-consumption indicator, EMPG, exhibits clear scenario sensitivity to the battery-electric driving share, α, and the prediction error transitions smoothly across the interval from FC-dominant to EPC-dominant behavior.

Stratified analysis further reveals directional differences among different vehicle categories and drive types in terms of error structure and sensitivity. This provides empirical evidence for the subsequent introduction of conditional modeling and targeted post-processing, and indicates that, compared with simple grouped modeling or single-task modeling, the proposed joint modeling approach achieves a more appropriate balance among accuracy, stability, and interpretability. Although the model still has limitations in out-of-domain generalization and in sparsely sampled subgroups, the proposed framework has clear application potential in regulatory disclosure, product development, and user-oriented energy-consumption guidance, especially through the incorporation of additional key technical and usage features, the explicit modeling of hierarchical conditions such as vehicle category and drive type, and the introduction of uncertainty evaluation.

Future work will focus on recalibration based on domestic catalog and real-world test data, as well as fusion modeling using multi-source information, including on-board operational data, road conditions, and climatic factors, in order to further improve the model's robustness and interpretability under real driving conditions and across different climate zones.

Authorship contribution: Narziev Nusratillo: Conceptualization, Methodology, Investigation, and Writing – Review & Editing. Md Borhan Uddin Niloy: Data Curation, Software, Formal Analysis, Visualization, and Writing – Review & Editing. Mohammed Shariful Islam Khan: Methodology, Validation, Resources, Investigation, and Writing – Review & Editing. Islam Shaikh Mahinur: Conceptualization, Methodology, Software, Formal Analysis, Data Curation, Visualization, Project Administration, Writing – Original Draft, and Writing – Review & Editing. All authors have read and agreed to the final version of the manuscript.

Data availability: The original data used in this study are publicly available from FuelEconomy.gov. The cleaned and processed data generated during the current study, together with the modeling-related materials, are available from the corresponding author on reasonable request.

REFERENCES

[1] A. Recalde, R. Cajo, W. Velasquez, and M. S. Alvarez-Alvarado, "Machine Learning and Optimization in Energy Management Systems for Plug-In Hybrid Electric Vehicles: A Comprehensive Review," Energies, vol. 17, no. 13, Art. no. 3059, 2024.

[2] J. J. Jui, M. A. Ahmad, M. M. I. Molla, and M. I. M. Rashid, "Optimal energy management strategies for hybrid electric vehicles: A recent survey of machine learning approaches," Journal of Engineering Research, 2024, doi: 10.1016/j.jer.2024.01.016.

[3] M. N. Boukoberine, M. F. Zia, T. Berghout, and M. Benbouzid, "Reinforcement learning-based energy management for hybrid electric vehicles: A comprehensive up-to-date review on methods, challenges, and research gaps," Energy and AI, Art. no. 100514, 2025.

[4] H. He, X. Meng, Y. Wang, A. Khajepour, X. An, R. Wang, and F. Sun, "Deep reinforcement learning based energy management strategies for electrified vehicles: Recent advances and perspectives," Renewable and Sustainable Energy Reviews, vol. 192, Art. no. 114248, 2024.

[5] Y. Li et al., "Deep reinforcement learning for intelligent energy management systems of hybrid-electric powertrains: Recent advances, open issues, and prospects," IEEE Transactions on Transportation Electrification, 2024.

[6] X. Zhang, Z. Zhang, Y. Liu, Z. Xu, and X. Qu, "A review of machine learning approaches for electric vehicle energy consumption modelling in urban transportation," Renewable Energy, vol. 234, Art. no. 121243, 2024.

[7] E. M. Szumska, Ł. Pawlik, D. Frej, and J. Ł. Wilk-Jakubowski, "Machine Learning Applications in Energy Consumption Forecasting and Management for Electric Vehicles: A Systematic Review," Energies, vol. 18, no. 20, Art. no. 5420, 2025.

[8] A. H. Ganesh and B. Xu, "A review of reinforcement learning based energy management systems for electrified powertrains: Progress, challenge, and potential solution," Renewable and Sustainable Energy Reviews, vol. 154, Art. no. 111833, 2022.

[9] C. Yang et al., "Efficient energy management strategy for hybrid electric vehicles/plug-in hybrid electric vehicles: Review and recent advances under intelligent transportation system," IET Intelligent Transport Systems, 2020.

[10] J. P. Torreglosa et al., "Analyzing the improvements of energy management systems for hybrid electric vehicles using a systematic literature review: How far are these controls from rule-based controls used in commercial vehicles?" Applied Sciences, 2020.

[11] Z. Wang et al., "Co-Optimization of eco-driving and energy management for connected HEV/PHEVs near signalized intersections: A review," Applied Sciences, 2023.

[12] P. Plötz, S. Link, H. Ringelschwendner, M. Keller, C. Moll, G. Bieker, J. Dornoff, and P. Mock, "Real-world usage of plug-in hybrid vehicles in Europe: A 2022 update on fuel consumption, electric driving, and CO2 emissions," ICCT/Fraunhofer ISI White Paper, 2022.

[13] A. Tansini et al., "Learning from the real-world: Insights on light-vehicle efficiency and CO2 emissions from long-term on-board fuel and energy consumption data collection," JRC Publications Repository, 2025.

[14] M. Mądziel et al., "Real-World CO2 Emissions of Plug-In Hybrid Vehicles," Energies, vol. 19, no. 5, Art. no. 1165, 2026.
[15] B. S. Guzmán, E. Asikainen, M. Elser, G. Fontaras, and coauthors, "Utility factor frameworks for plug-in hybrid electric vehicles: A comparative assessment," Transportation Research Part D: Transport and Environment, vol. 150, Art. no. 105098, 2026.
[16] P. Plötz, "Regulate reality in vehicle emission policy," Communications Earth & Environment, 2025.
[17] International Council on Clean Transportation, "Life-cycle greenhouse gas emissions from passenger cars in the European Union: A 2025 update and key factors to consider," 2025.
[18] P. Plötz et al., "Real-world Fuel Consumption and Potential Future Regulation of Plug-In Hybrid Electric Vehicles in Europe—An Empirical Analysis of about one Million Vehicles," Ariadne Analysis, 2025.
[19] J. Mamala et al., "Factors Influencing Energy Consumption in Plug-In Hybrid Vehicles: A Field Study," SAE Technical Paper 2025-01-0414, 2025.
[20] X. Sun et al., "An energy management strategy for plug-in hybrid electric vehicles based on deep learning and improved model predictive control," Energy, vol. 269, 2023.
[21] D. Shi, H. Xu, S. Wang, J. Hu, L. Chen, and C. Yin, "Deep reinforcement learning based adaptive energy management for plug-in hybrid electric vehicle with double deep Q-network," Energy, vol. 305, Art. no. 132402, 2024.
[22] D. Chen, T. Chen, Z. Li, Z. Liu, C. Sun, and H. Zhao, "Energy management strategy for plug-in hybrid electric vehicles based on vehicle speed prediction and limited traffic information," Energy, 2025.
[23] P. Zhu, J. Hu, Z. Zhu, F. Xiao, J. Li, and H. Peng, "An efficient energy management method for plug-in hybrid electric vehicles based on multi-source and multi-feature velocity prediction and improved extreme learning machine," Applied Energy, 2025, doi: 10.1016/j.apenergy.2024.125096.
[24] X. Tang et al., "Naturalistic data-driven predictive energy management for plug-in hybrid electric vehicles," IEEE Transactions on Transportation Electrification, vol. 7, pp. 497-508, 2020.
[25] Y. Wang, Y. Wu, Y. Tang, Q. Li, and H. He, "Cooperative energy management and eco-driving of plug-in hybrid electric vehicle via multi-agent reinforcement learning," Applied Energy, vol. 332, Art. no. 120563, 2023.
[26] W. Cui, N. Cui, T. Li, Z. Cui, Y. Du, and C. Zhang, "An efficient multi-objective hierarchical energy management strategy for plug-in hybrid electric vehicle in connected scenario," Energy, vol. 257, Art. no. 124690, 2022.
[27] S. Abd-Elhaleem, W. Shoeib, and A. A. Sobaih, "A new power management strategy for plug-in hybrid electric vehicles based on an intelligent controller integrated with CIGPSO algorithm," Energy, vol. 265, Art. no. 126153, 2023.
[28] Z. Lei et al., "An adaptive equivalent consumption minimization strategy for plug-in hybrid electric vehicles based on traffic information," Energy, 2020.
[29] S. Wang, K. Zhang, D. Shi, M. Li, and C. Yin, "Research on economical shifting strategy for multi-gear and multi-mode parallel plug-in HEV based on DIRECT algorithm," Energy, vol. 286, Art. no. 129574, 2024.
[30] S. Wang, K. Yao, W. Guo, X. Xu, Y. Liu, P. Qian, J. Zhao, and P. Dong, "Approximate optimal energy management strategy for multi-speed series-parallel PHEV integrating global prediction and real-time control," Energy, vol. 335, Art. no. 138164, 2025.
[31] J. Li, X. Wu, J. Fan, Y. Liu, and M. Xu, "Overcoming driving challenges in complex urban traffic: A multi-objective eco-driving strategy via safety model based reinforcement learning," Energy, vol. 284, Art. no. 128517, 2023.
[32] J. Li, X. Wu, M. Xu, and Y. Liu, "Multi-objective eco-driving strategy for connected and automated electric vehicles considering complex urban traffic influence factors," IEEE Transactions on Transportation Electrification, 2024, doi: 10.1109/TTE.2023.3349025.
[33] Y. Liu et al., "Hierarchical speed planning and energy management for autonomous plug-in hybrid electric vehicle in vehicle-following environment," Energy, 2022.
[34] W. Tang et al., "Hierarchical energy management strategy based on adaptive dynamic programming for hybrid electric vehicles in car-following scenarios," Energy, 2023.
[35] J. W. Bae et al., "Gaussian process approximate dynamic programming for energy-optimal supervisory control of parallel hybrid electric vehicles," IEEE Transactions on Vehicular Technology, 2022.
[36] S. Yoo, J. Shin, and S.-H. Choi, "Machine learning vehicle fuel efficiency prediction," Scientific Reports, vol. 15, Art. no. 14815, 2025.
[37] S. S. Haghshenas, S. S. Haghshenas, V. Astarita, and G. Guido, "Fuel consumption and CO2 emissions prediction in road transport using a hybrid deep learning approach," Transportation Engineering, Art. no. 100364, 2025.
[38] H. Huang, H. He, Y. Wang, Z. Zhang, and T. Wang, "Energy consumption prediction of electric vehicles for data-scarce scenarios using pre-trained model," Transportation Research Part D: Transport and Environment, vol. 146, Art. no. 104830, 2025.
[39] I. Hussain, K. B. Ching, C. Uttraphan, K. G. Tay, A. Noor, and S. A. Memon, "Optimizing electric vehicle energy consumption prediction through machine learning and ensemble approaches," Scientific Reports, vol. 15, Art. no. 29065, 2025.
[40] M. Z. Ghawy et al., "An Explainable Ensemble Machine Learning Method for Electric Vehicles Energy Consumption Rate Estimation," Concurrency and Computation: Practice and Experience, 2026, doi: 10.1002/cpe.70571.
[41] Y. Wang, J. Wu, H. He, Z. Wei, and F. Sun, "Data-driven energy management for electric vehicles using offline reinforcement learning," Nature Communications, vol. 16, Art. no. 2835, 2025.
[42] T. Zhang et al., "Predicting EV Charging Demand in Renewable-Energy-Powered Grids Using Explainable Machine Learning," Sustainability, vol. 17, no. 9, Art. no. 4158, 2025.
[43] Z. Jia, Z. Cao, J. Yin, L. Wu, N. Wei, Y. Zhang, Z. Jiang, Q. Zhang, and H. Mao, "A transferable vehicle energy consumption evaluation framework for quantifying vehicle electrification energy benefits," npj Sustainable Mobility and Transport, vol. 2, Art. no. 21, 2025.
[44] Y. Zhao, Z. Wang, Z.-J. M. Shen, and F. Sun, "Assessment of battery utilization and energy consumption in the large-scale development of urban electric vehicles," Proceedings of the National Academy of Sciences, vol. 118, e2017318118, 2021.
[45] G. Pozzato et al., "Analysis and key findings from real-world electric vehicle field data," Joule, vol. 7, pp. 2035-2053, 2023.