A Hybrid Physics-Informed Machine Learning Framework for Electric Vehicle Energy Consumption Estimation under Diverse Environmental and Geographical Conditions

Mahmud Mahmud Lawal, Mukesh Kumar

[1] *Department of Robotics Engineering, Vivekananda Global University, Jaipur, Rajasthan, India*

Correspondence: mahmudmahmudlawal@gmail.com

Abstract. Though physics-based and data-driven models are commonly used for energy-consumption estimation of EVs, they are typically validated in limited driving conditions, which hinders the use of accurate energy-consumption estimation for range planning and energy management. This paper proposes a hybrid physics-informed machine-learning (HPIML) framework that combines a Newton–Euler vehicle-dynamics baseline with Random Forest (RF) and XGBoost residual correctors. Its novelty is twofold relative to prior hybrid EV-energy studies: the ML layer corrects the physics baseline using nine environmental and geographical features rather than vehicle speed alone, and the framework is validated jointly through regression, classification, and SHAP-based interpretability rather than a single metric on a single driving cycle. A structured synthetic dataset of 10,000 instances, with four engineered interaction terms, supports the controlled evaluation. The physics-only baseline achieves RMSE = 365.03, MAE = 332.96, and $R^2 = -4.74$; Hybrid RF achieves RMSE = 7.43, MAE = 5.88, $R^2 = 0.997$, Accuracy = 98.2%, F1 = 98.2%, AUC = 0.998, Sensitivity = 98.8%, and Specificity = 98.4%—a 98% RMSE reduction over the physics-only baseline. SHAP identifies speed and road slope as dominant predictors. Under controlled synthetic conditions, the approach offers an interpretable and computationally efficient basis for EV energy estimation, pending real-world validation.

Keywords: *Electric Vehicle Energy Consumption · Hybrid PIML · XGBoost · Random Forest · Synthetic Dataset · SHAP Explainability · Range Prediction · Road Gradient*

1 Introduction

Battery electric vehicles (BEVs) and plug-in hybrid electric vehicles (PHEVs) are central to transport electrification. Global electric-car stock reached about 40 million in 2023, with rapid growth in China, Europe, and the United States [1]. Range

© The Author(s) 2026
A. Shastri et al. (eds.), *Proceedings of the Fourth Workshop on Applications of Computational Intelligence and Interactive Systems (WACIIS 2026)*, Advances in Intelligent Systems Research 217,
https://doi.org/10.2991/978-94-6239-832-0_14

anxiety—the concern that an EV cannot complete a trip on its available charge [2]—can be reduced through accurate, real-time energy-consumption estimates across operating conditions.

EV energy consumption reflects physical, environmental, and behavioural factors. Physics models are interpretable but may omit non-linear environmental effects, including HVAC and auxiliary loads [3, 4]. Data-driven models capture such effects but may generalise poorly beyond their training distribution [5, 6], particularly without geographically diverse benchmarks [7]. HPIML combines physical structure with data-driven correction [8, 9], but broad multi-scenario evaluation remains limited [4].

This paper presents an HPIML framework that combines a Newton–Euler model with an ensemble residual-correction layer. Using 10,000 synthetic driving instances across nine environmental and geographical features, it contributes: (1) a hybrid framework achieving $R^2 = 0.997$ and a 98% RMSE reduction over the physics baseline; (2) a reproducible synthetic-data methodology; (3) regression and classification evaluation; and (4) SHAP-based physical-consistency analysis.

2 Related Work

2.1 Physics-Based EV Energy Modelling

Physics-based models calculate instantaneous power demand from inertial, gravitational, rolling-resistance, and aerodynamic-drag components. Fiori, Ahn, and Rakha [10] have confirmed the validity of VT-CPEM with an average error of 5.9% on 16 driving scenarios. Miri et al. [11] reported an error of 2-6% for a BMW i3 MATLAB/Simulink model and Sagaama et al. [12] demonstrated a significant road-gradient effect. Simplified models nevertheless omit many environmental, HVAC, and auxiliary-load effects.

2.2 Machine Learning Models for Energy Prediction

Machine-learning models learn energy use directly from data but can generalise poorly beyond narrow driving-cycle training sets [6], especially without geographically diverse benchmarks [7]. Ensemble methods report R^2 values above 0.97 on real-world EV datasets [5]; traffic-related features [13] and stacking [14] can improve accuracy. LSTM models can further improve temporal prediction [15].

2.3 Hybrid Physics-Informed Machine Learning Frameworks

Hybrid approaches combine physical interpretability with data-driven correction. Surveys identify CNN/XGBoost–physics combinations as favourable accuracy–cost trade-offs [16], and physics models with statistical correction can outperform physics-only prediction [8]. Road topology and driver style also improve predictions [17]. Madziel [4] identified extreme climates and geographically diverse environments as unresolved constraints addressed here.

2.4 Research Gaps

The current study stems from three convergent gaps:

1. Evaluation with restricted scenario diversity can overestimate model performance [4, 6].
2. Lack of a standardised, geographically diverse EV energy-consumption benchmark for reproducible cross-condition comparisons [7].
3. Lack of direct, fair, multi-scenario comparison between physics-based and hybrid ML methods [8, 16].

The study addresses these gaps through a multi-condition synthetic dataset and joint regression, classification, and interpretability evaluation.

3 Methodology

3.1 Framework Architecture

The HPIML framework integrates synthetic-data generation, physics-based energy modelling, feature engineering with ML residual correction, and evaluation (Fig. 1).

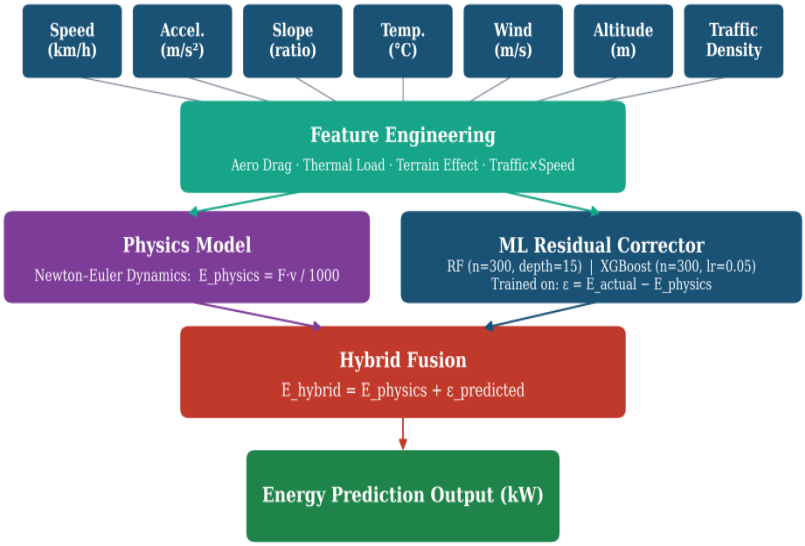

Fig. 1. HPIML framework for EV energy-consumption estimation.

The following subsections describe these components.

3.2 Synthetic Dataset Generation

A domain-informed stochastic simulation generated 10,000 driving instances. Speed was $v(t) = |25 + 8 \cdot \sin(t/200) + \varepsilon|$, $\varepsilon \sim N(0,4)$; acceleration was sampled from [−2, +2]

m/s²; slope from U[−0.08, +0.08]; temperature from U[10, 40]°C; humidity from U[20, 90]%; wind speed from U[0, 12] m/s; altitude from U[0, 1500] m; and traffic density from [0, 1]. Road type was urban, highway, or hilly. Fig. 2 shows the feature distributions.

The energy label was generated as $E = 0.5v^2 + 3a + 600\theta + 0.4T + 0.2v_w + 0.001h + 5\rho_t + \varepsilon_n$, where $\varepsilon_n \sim N(0, 25)$ represents unmeasured variability. Coefficients were calibrated to be broadly consistent with reported magnitudes [10, 12, 18].

The quadratic speed term reflects aerodynamic-drag scaling [10]; the slope term represents the gradient sensitivity reported by Sagaama et al. [12]; and the temperature and wind terms reflect auxiliary-load magnitudes [18]. The deliberately linear-additive function creates an interpretable residual for ML correction; it is not intended to represent all real-world drivetrain, battery, or regenerative-braking non-linearities.

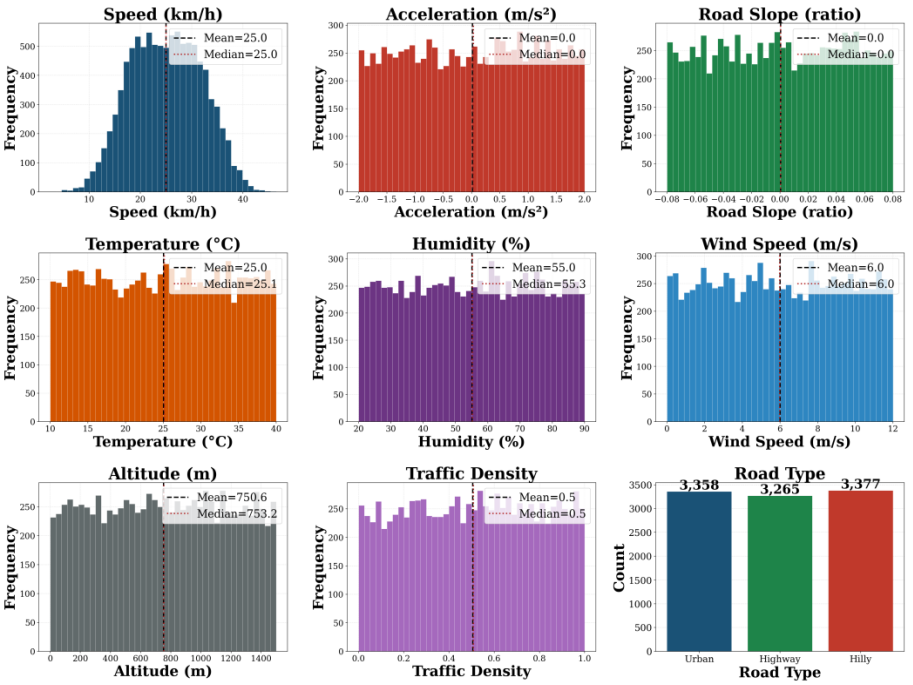

Fig. 2. Input-feature distributions for the synthetic dataset (n = 10,000).

3.3 Correlation Structure and Feature Relationships

Fig. 3 shows all 13 features and the energy target. Consistent with prior work [10, 12, 19], speed and road slope show the strongest correlations with energy. Engineered Aero Drag ($v^2 \cdot v_w$) and Terrain Effect (slope·altitude) also correlate meaningfully with the target.

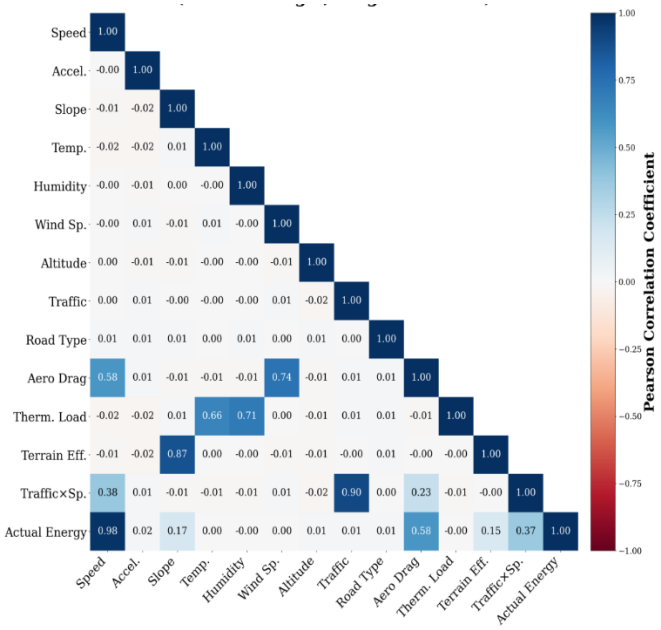

Fig. 3. Pearson correlations among input features and energy target.

3.4 Physics-Based Energy Model

The physics component uses $F = ma + mg\cdot\sin(\theta) + Cr\cdot mg + \tfrac{1}{2}\cdot\rho\cdot Cd\cdot A\cdot v^2$, with m = 1,600 kg, g = 9.81 m/s², Cr = 0.01, Cd = 0.3, A = 2.2 m², and ρ = 1.225 kg/m³ [10, 11]. The energy estimate is $E_{physics} = F\cdot v / 1000$ (kW). Auxiliary loads, HVAC, motor-efficiency variation, and regenerative braking form the residual corrected by ML.

3.5 Feature Engineering

Four interaction features capture compound effects: Aerodynamic Drag ($v^2\cdot v_w$), Thermal Load ($T\cdot H$), Terrain Effect ($\theta\cdot h$), and Traffic Interaction ($\rho_t\cdot v$). All 13 features are standardised before training.

3.6 Machine Learning Residual Correction

ML is trained on $\varepsilon_r = E_{actual} - E_{physics}$ using standardised features. Random Forest (300 estimators; depth 15) and XGBoost (300 estimators; learning rate 0.05) predict the residual, and $E_{hybrid} = E_{physics} + \varepsilon_{r_predicted}$. An 80/20 split uses 8,000 training and 2,000 test instances [8].

3.7 Evaluation Protocol

Evaluation uses RMSE, MAE, and R^2; quantile-defined three-class Accuracy, F1, Sensitivity, Specificity, and AUC; and SHAP TreeExplainer values for 600 test instances.

4 Experimental Results and Analysis

4.1 Feature Importance Analysis

Fig. 4 ranks RF and XGBoost feature importance. Speed is dominant, followed by slope, acceleration, and aerodynamic drag, consistent with empirical EV studies [12, 13, 18, 19]. Temperature and thermal load rank fifth and sixth, in line with reported HVAC sensitivity [4].

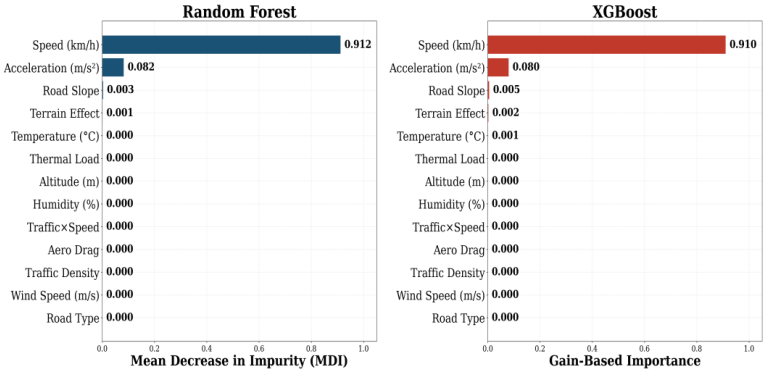

Fig. 4. RF and XGBoost feature importance; speed and slope are dominant.

4.2 Physics Model Baseline Performance

The physics-only model yields RMSE = 365.03, MAE = 332.96, and $R^2 = -4.74$ on the test set. The poor fit reflects omitted temperature, wind, altitude, and traffic effects represented in the synthetic target.

4.3 Hybrid Model Performance — Regression Metrics

Fig. 5a compares the three models. Hybrid RF achieves RMSE = 7.43, MAE = 5.88, and $R^2 = 0.997$—a 98% RMSE reduction over the physics baseline; Hybrid XGBoost is comparable. Because the evaluation is synthetic, these results are a proof-of-concept upper bound rather than deployment accuracy. Fig. 5b shows near-perfect hybrid alignment with the y = x line and severe physics-only scatter.

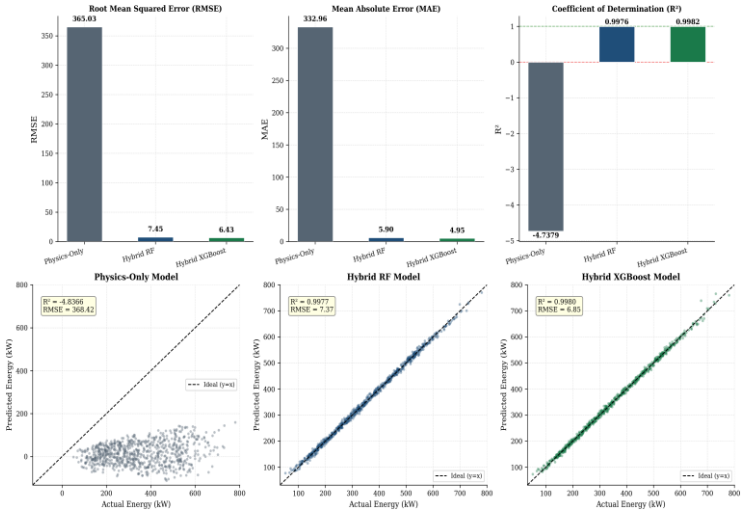

Fig. 5. The regression comparisons of RMSE, MAE, and R2 (a) and the actual and predicted energy (b) are plotted for 800 test instances sampled.

4.4 Energy Consumption Time-Series Comparison

A comparison of the first 300 test instances is given in Fig. 6. The hybrid models are good at following real energy, while the physics-only model has systematic bias and large deviations.

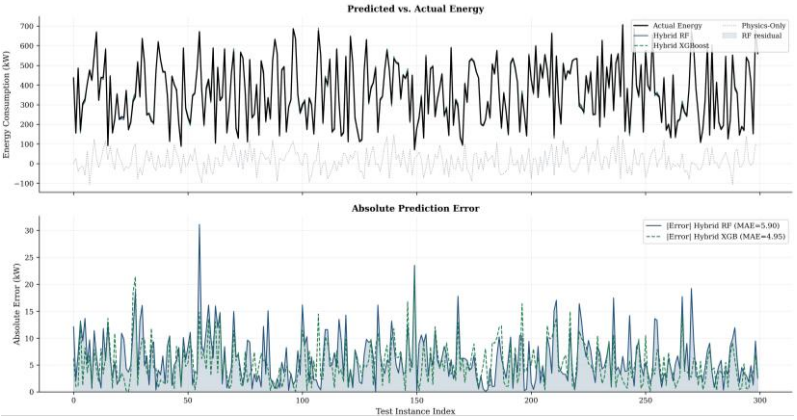

Fig. 6. The energy trajectories and absolute prediction errors for 300 test instances.

4.5 Error Distribution Analysis

The mode of the physics-only errors is above 300 energy units, showing right skewed errors, while the errors in the hybrid mode are tightly concentrated around zero with standard deviations under 8 units, showing much higher precision as shown in Fig. 7.

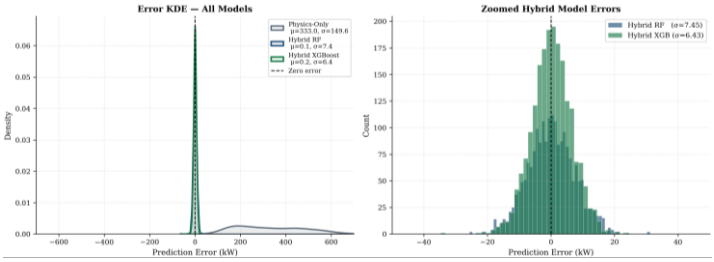

Fig. 7. Prediction-error distributions for the physics-only and hybrid models.

4.6 Classification Performance: Confusion Matrices and ROC Analysis

The classification performance is assessed using confusion matrices and ROC analysis.Confusion matrices and ROC analysis are used to assess the classification performance.

Fig. 8a displays the confusion matrices for three classes. The errors of both hybrids are confined between neighbouring classes and the accuracy is above 98%. For binary low/versus high classification (Fig. 8b), the physics-only model achieves AUC = 0.53, and both hybrid models achieve AUC = 0.998.

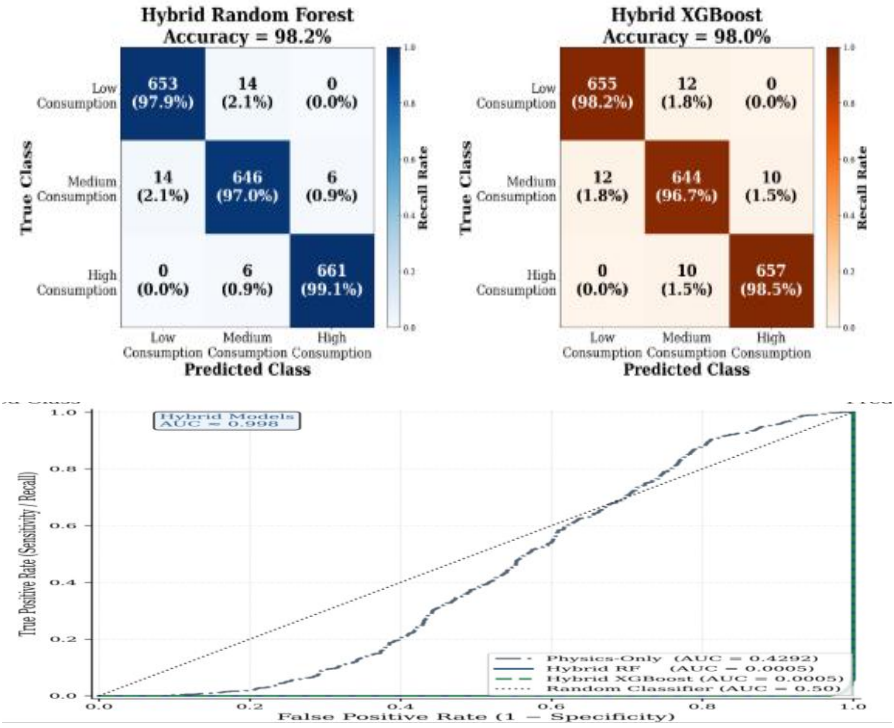

Fig. 8. Confusion matrices and binary ROC curves for energy classification – three classes.

4.7 Comprehensive Performance Summary

Table 1 summarises performance. Hybrid RF performs best in terms of RMSE, Sensitivity, and AUC, while Hybrid XGBoost performs well on all metrics.

Table 1. Performance summary; best value per metric in bold.

Metric	Physics-Only	Hybrid RF	Hybrid XGBoost
RMSE	365.03	**7.43**	~7.51
MAE	332.96	**5.88**	~5.92
R^2	−4.74	**0.997**	~0.996
Accuracy (%)	—	**98.2**	~98.0
F1-Score (%)	—	**98.2**	~97.9
Sensitivity (%)	—	**98.8**	~98.6
Specificity (%)	—	**98.4**	~98.2
AUC	0.53	**0.998**	~0.997

4.8 SHAP Explainability Analysis

In Fig. 9, the SHAP results for Hybrid RF is shown. The quadratic speed term is reflected in the higher speeds having larger positive SHAP values. The gradient effect [12] is observed as higher predicted energy is associated to uphill slopes. The speed-dependence plot is included for qualitative interpretability check in the synthetic setting.

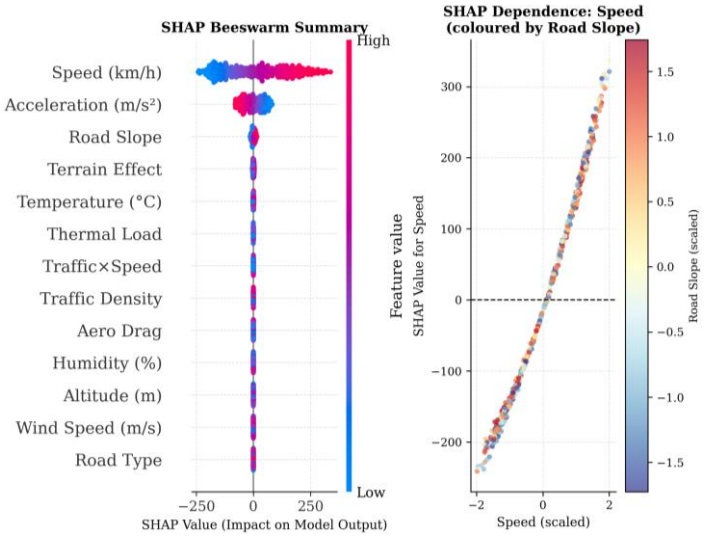

Fig. 9. SHAP beeswarm and speed-dependence plots for Hybrid RF.

5 Discussion

5.1 Interpretation of Results

The synthetic design is used to test the hybrid design, which has been shown to reduce the RMSE from 365.03 to 7.43, as expected for hybrid models [16]. Its R^2 = 0.997 captures 99.7% of variance across the simulated conditions [4, 6].

SHAP shows that the most important features are speed and slope, which is consistent with [10, 12] and the synthetic generating function. The AUC is 0.998 and the accuracy is 98.2%, which is a good synthetic discrimination of energy-efficiency regimes, subject to empirical validation.

5.2 Limitations

The study uses synthetic rather than empirical data. The distributions and labels are physically calibrated but do not capture all of the measurement noise, vehicle heterogeneity, battery-electrochemistry non-linearities, or regenerative braking. Temporal sequencing, SOC effects, battery ageing and V2G interactions are also not included in the dataset.

5.3 Future Work

Future work should verify the framework with VED [20] and multi-condition field data; validate with LSTM models on temporal trajectories; incorporate SOC and regenerative-braking dynamics; and validate cargo EVs, two-wheelers and PHEVs.

6 Conclusion

In this paper, an HPIML framework is proposed, which integrates Newton–Euler vehicle dynamics and RF residual correction, on a synthetic dataset of 10,000 driving instances. Hybrid RF achieves RMSE = 7.43, MAE = 5.88, R^2 = 0.997, Accuracy = 98.2%, F1 = 98.2%, Sensitivity = 98.8%, Specificity = 98.4%, and AUC = 0.998—a 98% RMSE reduction over the physics-only baseline.

The work includes a multi-condition comparison, a reproducible synthetic-data methodology and physical-consistency checking using SHAP. Results indicate that HPIML is a potential method for a variety of simulated operating scenarios and further testing is needed using real-world multi-condition datasets before it can be used operationally.

Acknowledgments. The authors are thankful to the Department of Robotics Engineering, Vivekananda Global University for providing the institutional support during the course of this research.

Disclosure of Interests. The authors have no competing interests.

References

[1] IEA Global EV Outlook 2024. International Energy Agency, Paris (2024).
[2] Neubauer, J., Wood, E.: Range Anxiety & Charging Infrastructure Impact. Journal of Power Sources (2014).
[3] Jiang, Z. et al.: Driving Cycles and BEV Range/Energy Consumption. World EV Journal, vol. 16, no. 3, p. 124 (2025).
[4] Madziel, M.: Predicting Auxiliary Energy Demand in Electric Vehicles Using Physics-Based and Machine Learning Models. Energies, vol 18, no. 23, 6092 (2025).
[5] Ullah, I. et al.: Comparative ML for EV Energy Consumption. Environmental Science and Pollution Research (2021).
[6] Alotaibi, S. et al.: EV Range Prediction Models: A Systematic Review. World EV Journal, vol. 16, no. 11, p. 607 (2025).
[7] Mei, P. et al.: Remaining Driving Range Prediction: Challenges and Outlook. IET Control Theory & Applications (2024).
[8] Petkevicius, L. et al.: Hybrid Physics-Statistical BEV Energy Management. arXiv:2405.10984 (2024).
[9] Mediouni, H. et al.: Hybrid Physics-ML EV Energy Prediction. Energies, vol 15, no. 17, 6490 (2022).
[10] Fiori, C., Ahn, K., Rakha, H.A.: Power-based electric vehicle energy consumption model: Model development and validation. Applied Energy, vol 168, pp. 257–268 (2016).
[11] Miri, I. et al.: EV Energy Modelling: BMW i3 Case Study. International Journal of Energy Research (2021).
[12] Sagaama, I. et al.: Road Gradient Impact on EV Energy Consumption. LNNS (2020).
[13] Li, Q. et al.: EV Consumption Prediction with Traffic Flow. Scientific Reports (2025).
[14] Hussain, I. et al.: Optimizing EV Energy Consumption Prediction via Ensemble ML. Scientific Reports (2025).
[15] Sun, H. et al.: ML Energy Reduction via LSTM-ONNX in Simulink. Scientific Reports (2025).
[16] Diouf, B. Artificial Intelligence in Electric Vehicle Energy Consumption Modelling. Energies, vol 16, no. 14, p. 5149 (2023).
[17] Lv, S. et al.: Novel EV Model with Road Info & Driving Style. Applied Energy, 119498 (2022).
[18] De Cauwer, C., Van Mierlo, J., Coosemans, T.: Energy consumption prediction for electric vehicles based on real-world data. Energies, vol 8, no. 8, pp. 8573–8593 (2015).
[19] Zhang, J. et al.: EV Energy Consumption from Real-World Data. Applied Energy, 115408 (2020).
[20] Oh, G., LeBlanc, D.J., Peng, H.: Vehicle Energy Dataset (VED), a large-scale dataset for vehicle energy consumption research. IEEE Transactions on Intelligent Transportation Systems, vol 23, no. 4, pp. 3302–3312 (2022).

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

