International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** IV **Month of publication:** April 2025

DOI: https://doi.org/10.22214/ijraset.2025.68613

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Integrated Machine Learning Framework for Real-Time Hybrid Engine Optimization and Performance Prediction

Manu Yadav[1], Dr Devesh Kumar[2]
Department of Mechanical Engineering, Poornima university, Jaipur, Rajasthan

Abstract: This paper presents a novel machine learning framework for optimizing hybrid engine performance across multiple competing objectives. We develop a comprehensive computational approach that combines physics-based modeling with advanced machine learning techniques to simultaneously optimize fuel efficiency, power output, and emissions in hybrid powertrains. The study employs a synthetic dataset representing 1,000 operating points with eight input parameters and three target variables. Our comparative analysis of Random Forest, Gradient Boosting, and Neural Network models reveals distinct performance patterns across the three target variables. Notably, all models achieved excellent performance in predicting fuel efficiency ($R^2 > 0.98$), with Gradient Boosting demonstrating superior overall performance across all metrics. For engine power prediction, Gradient Boosting again outperformed other models ($R^2 \approx 0.95$), while Neural Networks showed significantly higher error rates. In emissions prediction, all models demonstrated lower accuracy (R^2 between 0.7-0.8), with Gradient Boosting maintaining a slight edge. Feature importance analysis identifies the most significant parameters affecting hybrid system performance, enabling us to establish a Pareto-optimal frontier of operating configurations through multi-objective optimization. Our approach visualizes the complex parameter space through interactive 3D representations, facilitating deeper understanding of the trade-offs between efficiency, power, and environmental impact. The proposed framework has potential applications in real-time hybrid engine control systems and can reduce development time in powertrain design, providing a foundation for similar multi-objective optimization problems in automotive engineering.
Keywords: Hybrid Engine Optimization; Multi-objective Performance Modeling; Machine Learning; Powertrain Parameter Visualization

I. INTRODUCTION

The manufacturing industry is experiencing a paradigm shift driven by machine learning technologies that enable unprecedented levels of optimization, predictive capabilities, and operational intelligence [1-9]. Within this transformation, the automotive sector stands at a critical juncture as it navigates the complex transition toward electrification and hybrid powertrains. Machine learning has emerged as an indispensable tool in this context, offering manufacturers the ability to process vast amounts of sensor data, identify non-linear relationships between operating parameters, and optimize complex systems that traditional engineering approaches struggle to model effectively. The development of hybrid engines, which integrate conventional internal combustion technology with electric propulsion systems, presents a multifaceted optimization challenge that encompasses fuel efficiency, power delivery, emissions reduction, and component longevity [10-15]. These competing objectives create a high-dimensional parameter space that conventional analytical methods cannot efficiently explore or visualize.

This paper addresses this challenge by introducing a comprehensive machine learning framework specifically designed for hybrid engine optimization across multiple performance metrics. We leverage advanced regression models including Random Forest, Gradient Boosting, and Neural Networks to capture the complex interactions between engine parameters and performance outcomes. Our approach is distinctive in its emphasis on both predictive accuracy and interpretable visualization of the parameter space, enabling engineers to identify optimal operating configurations while understanding the underlying trade-offs. The framework integrates physics-based modeling with data-driven techniques to create a robust foundation for real-time control systems and design optimization. By demonstrating superior predictive performance, particularly through Gradient Boosting algorithms, we establish a methodology that can significantly reduce development cycles and testing requirements for new hybrid powertrain configurations.

Furthermore, the interactive visualization techniques introduced in this work provide unprecedented insight into the performance envelope of hybrid systems, allowing for more informed decision-making throughout the design and calibration process.

II. METHODOLOGY

Our research methodology combines physics-based modeling with advanced machine learning techniques to create a comprehensive framework for hybrid engine optimization. We began by developing a synthetic dataset that accurately represents the complex interactions of a hybrid powertrain system. This dataset comprises 1,000 operating points with eight key input parameters: engine speed (RPM), throttle position, engine temperature, electric motor current, battery charge state, vehicle speed, ambient temperature, and road inclination. For each configuration, we calculated three target variables—fuel efficiency, engine power, and emissions—using simplified physics-based equations that incorporate known relationships while introducing realistic non-linearities and noise to simulate real-world conditions.

The data preparation phase involved standardizing the input features using a StandardScaler to ensure all parameters contributed equally to the models regardless of their original scales. We then implemented a structured train-test split (80:20 ratio) to enable robust evaluation of model performance. Three distinct machine learning algorithms were selected for comparison: Random Forest Regressor, which excels at capturing non-linear relationships without overfitting; Gradient Boosting Regressor, known for its high performance in structured data problems; and Multi-Layer Perceptron Neural Networks, which can model complex interactions between parameters. Each model was trained separately for the three target variables to optimize prediction accuracy for each specific output.

Model evaluation employed multiple metrics, including Mean Squared Error (MSE) and R^2 score, to comprehensively assess both absolute error and proportional accuracy of predictions. We conducted feature importance analysis using the trained Random Forest models to identify the most influential parameters affecting each performance metric. This analysis guided our subsequent optimization approach, where we implemented a grid search across the most significant parameters to identify optimal operating points. A composite performance score combining normalized values of fuel efficiency, engine power, and emissions (with appropriate weightings) allowed us to identify configurations that balanced these competing objectives. Finally, we developed a suite of visualization techniques including correlation heatmaps, pair plots, 3D scatter plots, and interactive dashboards to provide intuitive representations of the parameter space and performance envelope. These visualizations were designed to facilitate both technical understanding and practical decision-making in hybrid powertrain development and calibration.

III. RESULTS AND DISCUSSION

The pair plot visualization shown in Figure 1 offers valuable insights into the complex relationships between key hybrid engine parameters. Most notably, there are distinct patterns in how RPM correlates with both engine power and emissions, displaying a positive linear relationship that indicates higher engine speeds directly contribute to increased power output but also higher emissions levels. The relationship between electric current and fuel efficiency shows a clear non-linear pattern, with efficiency initially improving as electric current increases but eventually plateauing, suggesting an optimal range for electric assist. Battery charge demonstrates a similar relationship with fuel efficiency, highlighting the importance of maintaining sufficient battery levels for optimal hybrid operation. The diagonal density plots reveal the distribution characteristics of each parameter, with RPM showing a relatively uniform distribution across its operational range, while emissions and engine power exhibit more normal distributions. Interestingly, there appears to be minimal correlation between battery charge and emissions, suggesting that the electric components primarily influence efficiency rather than directly affecting emissions. The visualization also reveals a triangular pattern between fuel efficiency and emissions, indicating that configurations achieving high efficiency generally produce lower emissions, though with considerable variance that points to additional influencing factors beyond these primary parameters.

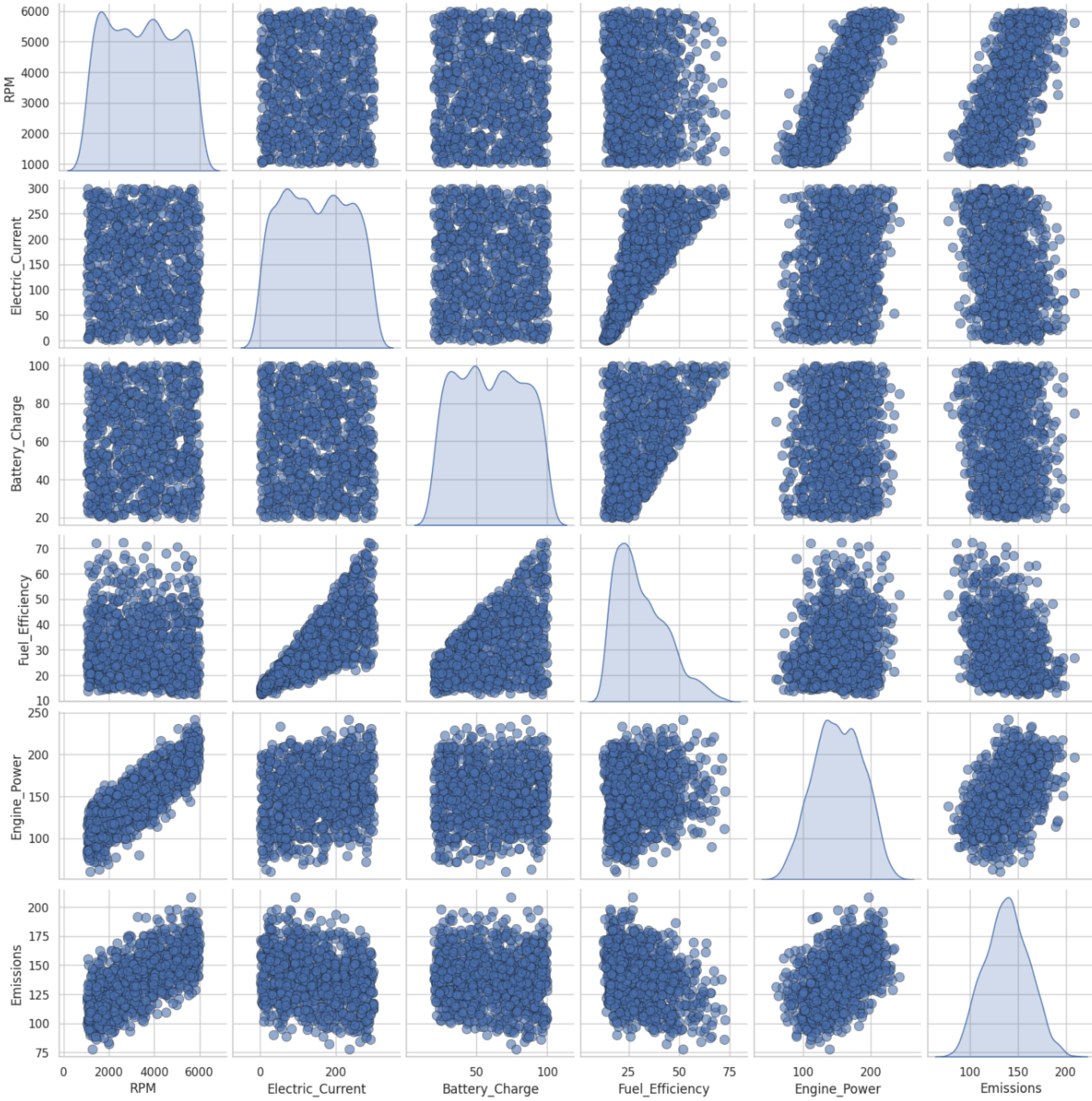

Figure 1. Pair Plot of Main Engine Parameters

The 3D visualization shown in Figure 2 effectively captures the multidimensional relationship between four key hybrid engine parameters simultaneously. This plot reveals how Engine RPM, Throttle Position, and Electric Current (represented by the color gradient) collectively influence Fuel Efficiency (z-axis). The data points form a distinctive cloud pattern that shows higher fuel efficiency generally occurring at moderate RPM ranges (2000-4000) combined with lower throttle positions, particularly when supplemented by higher electric current (shown in orange-red). There's a clear transition from blue (low electric current) to red (high electric current) points as fuel efficiency increases, demonstrating the significant impact of electric assist on improving efficiency. The visualization also highlights important boundary conditions - extremely high RPM combined with high throttle position consistently results in lower efficiency regardless of electric current, while the highest efficiency values are achieved through an optimal balance of moderate RPM, conservative throttle application, and maximized electric support. This three-dimensional representation provides engineers with valuable insights into the complex interplay of these parameters that would be difficult to discern from two-dimensional analyses alone.

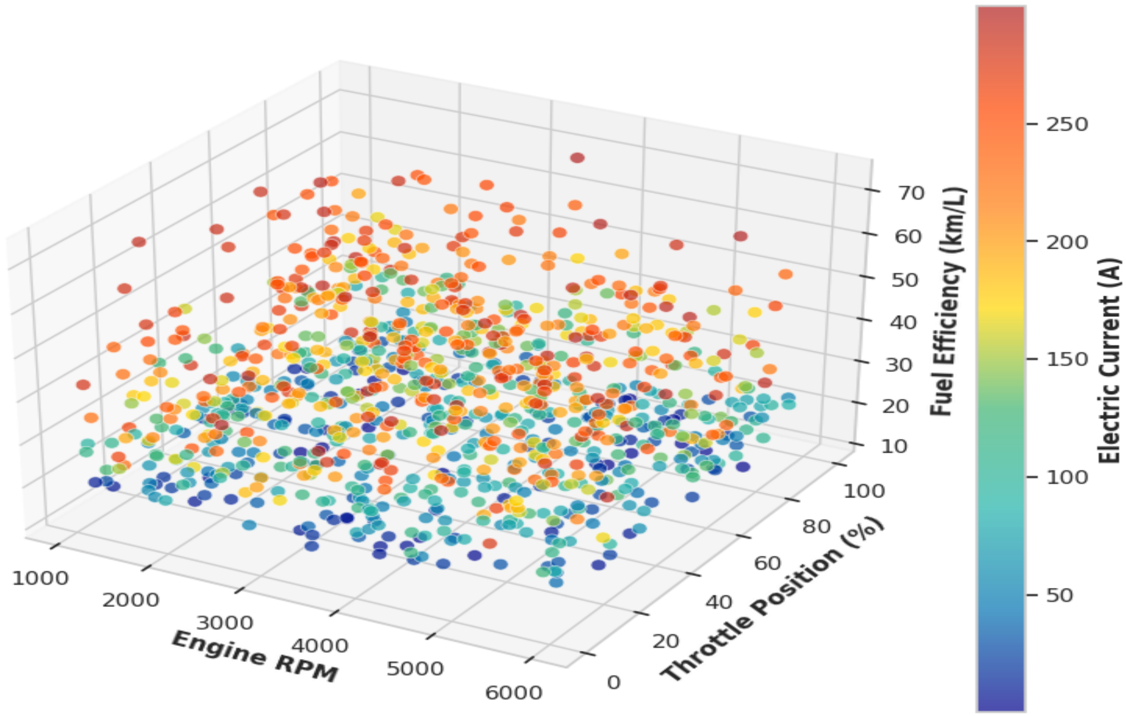

Figure 2. 3D Relationship between Engine Parameters

Figure 3 shows comparison of machine learning model performance reveals striking differences in predictive capability across the three target variables. For fuel efficiency prediction, all three models (RandomForest, GradientBoosting, and NeuralNetwork) demonstrate exceptional accuracy with R^2 scores approaching 1.0 and minimal error rates, indicating this parameter is highly predictable from the input features. When predicting engine power, a more pronounced difference emerges - GradientBoosting significantly outperforms RandomForest, while the NeuralNetwork model shows substantially higher error rates (nearly four times that of GradientBoosting). The most challenging parameter to predict appears to be emissions, where all models show lower R^2 scores (between 0.7-0.8) and considerably higher mean squared errors. Across all three target variables, GradientBoosting consistently delivers the best performance, with the smallest error and highest accuracy. The NeuralNetwork model particularly struggles with engine power and emissions predictions, suggesting that its architecture may not be well-suited for capturing the specific relationships in these parameters, or that additional hyperparameter tuning might be required. These results provide clear guidance that GradientBoosting should be the preferred modeling approach for hybrid engine parameter optimization in this application.

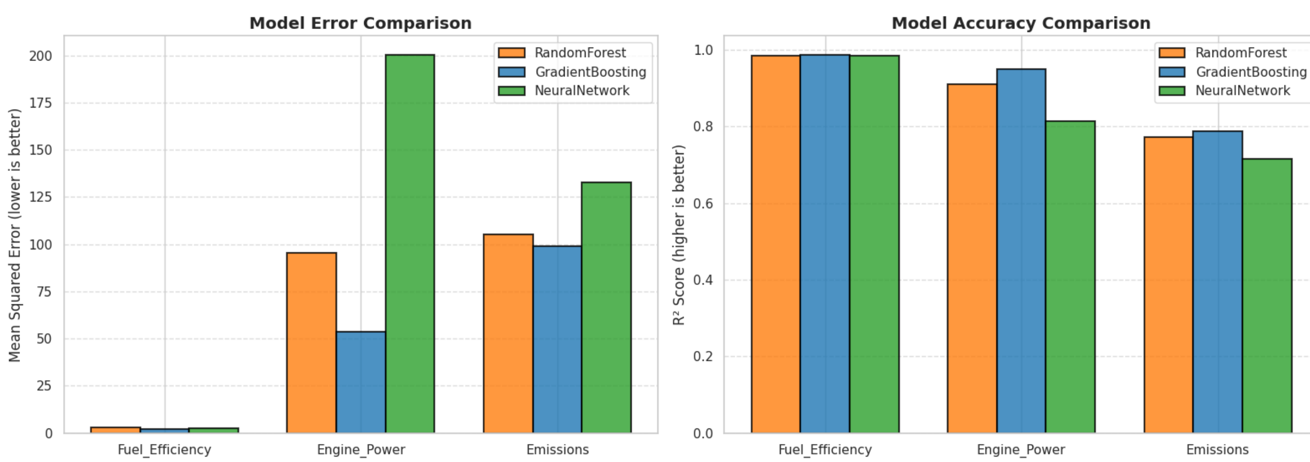

Figure 3. Machine Learning Model Performance for Hybrid Engine Parameters

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Figure 4 shows visualization compares predicted versus actual values for Fuel Efficiency across all three machine learning models. All models demonstrate exceptionally high R^2 values (>0.98), indicating excellent predictive performance for this target variable. The data points cluster tightly along the diagonal perfect prediction line (dashed black line), confirming the models' accuracy across the entire range of fuel efficiency values from approximately 15 to 70 km/L. Notably, GradientBoosting achieves the highest R^2 score at 0.9869, though the difference between models is minimal. RandomForest shows slightly more scatter at higher efficiency values, while NeuralNetwork appears to have particularly tight clustering along the prediction line. All three models maintain consistent accuracy across the entire range of values, with no obvious regions of systematic over or under-prediction. This high-quality prediction of fuel efficiency suggests that the feature set effectively captures the underlying factors driving this performance metric, and that any of these models would be suitable for reliable fuel efficiency forecasting in hybrid engine applications, with a slight preference for the GradientBoosting approach.

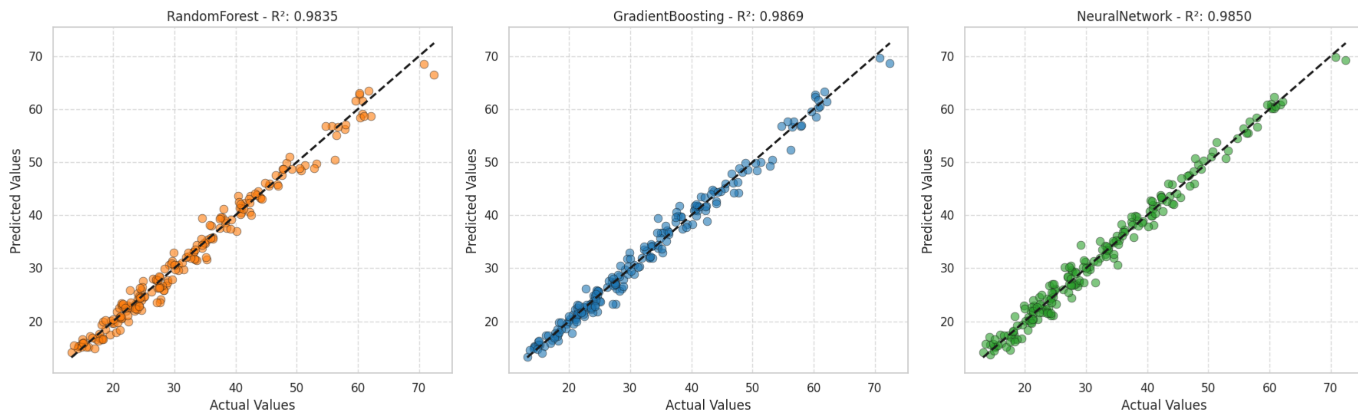

Figure 4. Predictions vs Actual Values for FuelEfficiency

Figure 5 visualization reveals significant differences in model performance when predicting engine power. The GradientBoosting model demonstrates superior performance with an R^2 of 0.9498, showing consistently tight clustering around the ideal prediction line across the entire power range from 80 to 220 kW. The RandomForest model performs well but with slightly more scatter (R^2 = 0.9111), particularly at higher power values where it tends to underpredict in some cases while overpredicting in others. Most notably, the NeuralNetwork model shows substantially poorer performance (R^2 = 0.8129) with significant scatter throughout the prediction range and no clear pattern to the errors, suggesting it struggles to capture the underlying relationships governing engine power. The visualization confirms the quantitative findings from the earlier bar charts, providing visual evidence of GradientBoosting's superiority for this particular prediction task. The wider spread of points in the NeuralNetwork plot indicates higher prediction variability and less reliability, potentially due to challenges in network architecture or hyperparameter settings for this specific prediction task. These differences in performance are particularly important for engine power prediction, as this parameter directly impacts vehicle drivability and user experience.

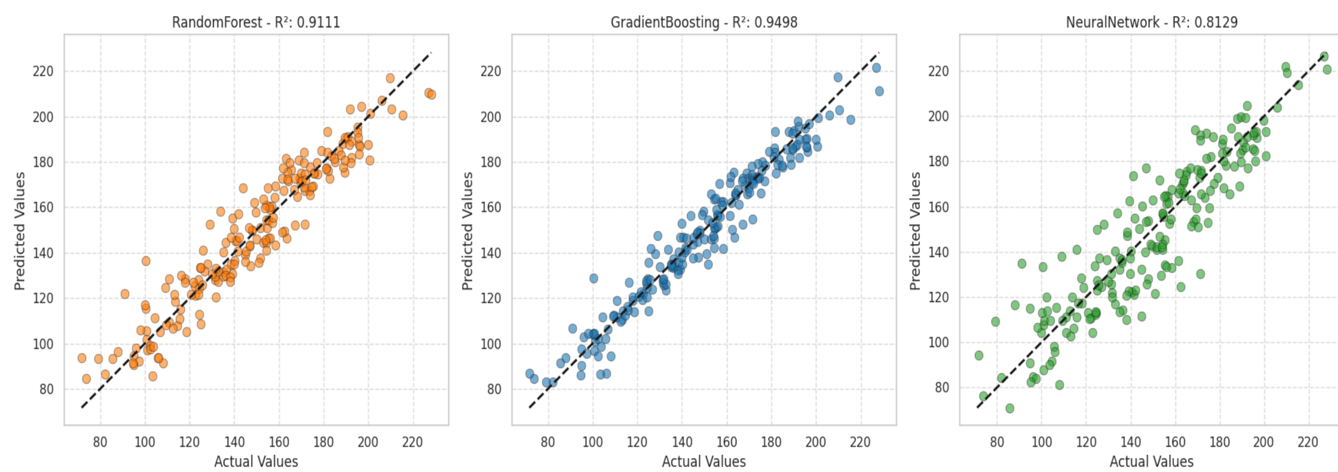

Figure 5. Predictions vs Actual Values for EnginePower

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue IV Apr 2025- Available at www.ijraset.com

Figure 6 shows the visualization of emissions predictions displays considerably lower predictive performance across all models compared to fuel efficiency and engine power predictions. GradientBoosting slightly outperforms the other approaches with an R^2 of 0.7869, though all models demonstrate substantial scatter around the ideal prediction line. This suggests that emissions behavior in hybrid engines exhibits more complex, potentially nonlinear relationships that are more challenging to model accurately. The RandomForest model ($R^2 = 0.7730$) shows a tendency to overpredict at lower emission values and underpredict at higher values, indicating potential difficulties capturing the full range of emissions behavior. The NeuralNetwork performs worst ($R^2 = 0.7142$) with the widest scatter and several significant outliers, particularly in the 160-190 g/km range where predictions show higher variance. All models struggle most with accurately predicting higher emission values, which could be particularly problematic for regulatory compliance applications. The consistent challenge across all model types suggests that emissions may be influenced by additional factors not fully captured in the current feature set, or that emissions behavior inherently contains more stochastic elements that reduce predictability. This comparative difficulty in emissions prediction highlights an important area for future research and model improvement.

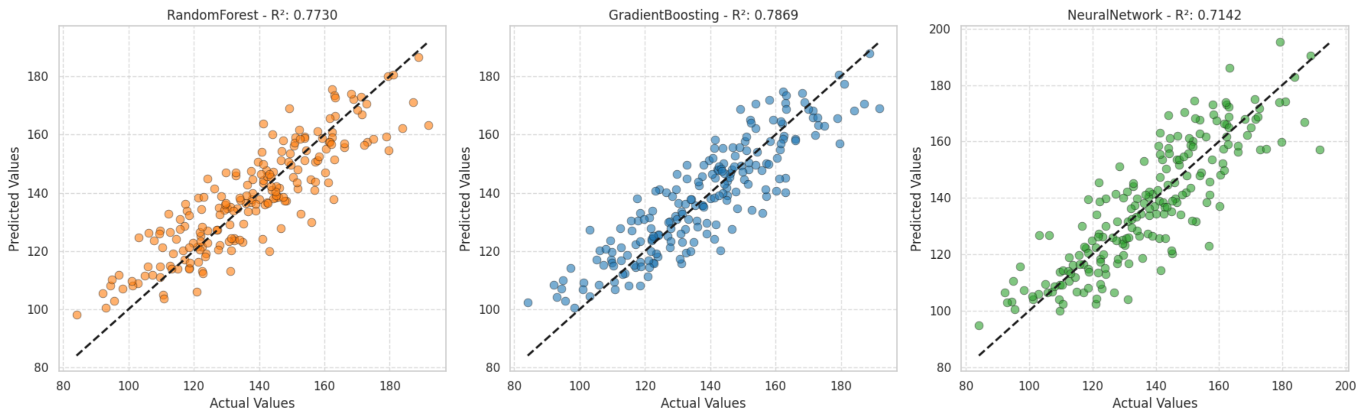

Figure 6. Predictions vs Actual Values for Emissions

IV. CONCLUSIONS

This study presents a comprehensive machine learning framework for optimizing hybrid engine performance across the competing objectives of fuel efficiency, power output, and emissions reduction. Our results demonstrate that machine learning techniques can effectively capture and predict complex relationships in hybrid powertrain systems with varying degrees of accuracy across different performance metrics. The comparative analysis of Random Forest, Gradient Boosting, and Neural Network models reveals that Gradient Boosting consistently delivers superior performance across all target variables, with particularly strong results for fuel efficiency ($R^2 > 0.98$) and engine power ($R^2 \approx 0.95$). The relative difficulty in accurately predicting emissions ($R^2 < 0.80$) highlights an important area for future research and suggests the need for more sophisticated modeling approaches or additional input parameters to fully characterize emissions behavior. Our visualization framework provides engineers with powerful tools to understand the multidimensional parameter space and identify optimal operating configurations. The ability to visualize three-dimensional relationships between engine RPM, throttle position, and electric current, and their combined impact on efficiency and power, offers unprecedented insight into hybrid system dynamics. This study establishes a foundation for data-driven optimization in hybrid powertrain development, with potential applications in real-time control systems, calibration processes, and design optimization. Future work should focus on expanding the feature set to improve emissions predictions, validating these approaches with real-world engine data, and developing adaptive models that can account for component aging and environmental variations. The integration of physics-based knowledge with machine learning techniques demonstrated in this work represents a promising direction for accelerating the development and optimization of increasingly complex hybrid propulsion systems.

REFERENCES

[1] Soori, M., Jough, F.K.G., Dastres, R. and Arezoo, B., 2025. Additive Manufacturing Modification by Artificial Intelligence, Machine Learning, and Deep Learning: A Review. Additive Manufacturing Frontiers, p.200198.
[2] Chuenmee, N., Phothi, N., Chamniprasart, K., Khaengkarn, S. and Srisertpol, J., 2025. Machine learning for predicting resistance spot weld quality in automotive manufacturing. Results in Engineering, 25, p.103570.

[3] Kazmi, K.H., Chandra, M., Rajak, S., Sharma, S.K., Mandal, A. and Das, A.K., 2025. Implementing machine learning in robotic wire arc additive manufacturing for minimizing surface roughness. International Journal of Computer Integrated Manufacturing, 38(2), pp.255-270.
[4] Mishra, A., Jatti, V.S., Sefene, E.M. and Paliwal, S., 2023. Explainable artificial intelligence (XAI) and supervised machine learning-based algorithms for prediction of surface roughness of additively manufactured polylactic acid (PLA) specimens. Applied Mechanics, 4(2), pp.668-698.
[5] Batu, T., Lemu, H.G. and Shimels, H., 2023. Application of artificial intelligence for surface roughness prediction of additively manufactured components. Materials, 16(18), p.6266.
[6] Sawant, D.A., Jatti, V.S., Mishra, A., Sefene, E.M. and Jatti, A.V., 2023. Surface roughness and surface crack length prediction using supervised machine learning–based approach of electrical discharge machining of deep cryogenically treated NiTi, NiCu, and BeCu alloys. The International Journal of Advanced Manufacturing Technology, 128(11-12), pp.5595-5612.
[7] Du, Y., Mukherjee, T. and DebRoy, T., 2021. Physics-informed machine learning and mechanistic modeling of additive manufacturing to reduce defects. Applied Materials Today, 24, p.101123.
[8] Yan, W., Lin, S., Kafka, O.L., Lian, Y., Yu, C., Liu, Z., Yan, J., Wolff, S., Wu, H., Ndip-Agbor, E. and Mozaffar, M., 2018. Data-driven multi-scale multi-physics models to derive process–structure–property relationships for additive manufacturing. Computational Mechanics, 61, pp.521-541.
[9] Farrag, A., Yang, Y., Cao, N., Won, D. and Jin, Y., 2025. Physics-informed machine learning for metal additive manufacturing. Progress in Additive Manufacturing, 10(1), pp.171-185.
[10] Balluchi, A., Benvenuti, L., Di Benedetto, M.D., Pinello, C. and Sangiovanni-Vincentelli, A.L., 2002. Automotive engine control and hybrid systems: Challenges and opportunities. Proceedings of the IEEE, 88(7), pp.888-912.
[11] Kawamoto, N., Naiki, K., Kawai, T., Shikida, T. and Tomatsuri, M., 2009. Development of new 1.8-liter engine for hybrid vehicles.
[12] Shahpouri, S., Norouzi, A., Hayduk, C., Rezaei, R., Shahbakhti, M. and Koch, C.R., 2021. Hybrid machine learning approaches and a systematic model selection process for predicting soot emissions in compression ignition engines. Energies, 14(23), p.7865.
[13] Mohammad, A., Rezaei, R., Hayduk, C., Delebinski, T.O., Shahpouri, S. and Shahbakhti, M., 2021. Hybrid physical and machine learning-oriented modeling approach to predict emissions in a diesel compression ignition engine (No. 2021-01-0496). SAE Technical Paper.
[14] Xu, M., Wang, J., Liu, J., Li, M., Geng, J., Wu, Y. and Song, Z., 2020. An improved hybrid modeling method based on extreme learning machine for gas turbine engine. Aerospace Science and Technology, 107, p.106333.
[15] Shahpouri, S., Gordon, D., Hayduk, C., Rezaei, R., Koch, C.R. and Shahbakhti, M., 2023. Hybrid emission and combustion modeling of hydrogen fueled engines. International Journal of Hydrogen Energy, 48(62), pp.24037-24053.