TELKOMNIKA Telecommunication Computing Electronics and Control
Vol. 23, No. 6, December 2025, pp. 1543~1554
ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v23i6.27128

Application of artificial intelligence in emission prediction for hybrid electric vehicles: integrating ANN and GPR

Heru Priyanto[1], Rizqon Fajar[1], Yaaro Telaumbanua[1], Ariyanto[1], Mohammad Mukhlas Af[1], Sigit Tri Atmaja[1], Muhammad Samsul Maarif[1,2], Kurnia Fajar Adhi Sukra[1,2], Fauzi Dwi Setiawan[1]
[1]Research Center for Transportation Technology, National Research and Innovation Agency, South Tangerang, Indonesia
[2]Mechanical Precision Engineering Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia (UTM), Kuala Lumpur, Malaysia

Article Info

Article history:

Received Apr 14, 2025
Revised Sep 11, 2025
Accepted Oct 19, 2025

Keywords:

Artificial neural networks
Emission prediction
Gaussian process regression
Hybrid electric vehicles
Prediction uncertainty
Vehicle emissions

ABSTRACT

In recent years, hybrid electric vehicles (HEVs) have emerged as a promising solution to mitigate vehicular emissions and improve fuel efficiency. This study focuses on the Toyota Prius HEV, employing advanced artificial neural networks (ANN) and Gaussian process regression (GPR) to develop a predictive model for vehicle emissions. The model considers multiple pollutants, including carbon monoxide (CO), carbon dioxide (CO_2), hydrocarbons (HC), and nitrogen oxides (NOx), measured under diverse driving conditions. The ANN model predicts emission trends, while GPR estimates prediction uncertainty, enhancing the model's robustness. The GPR models achieved uncertainty levels of ±0.829 ppm for CO, ±9.978 ppm for HC, ±0.144 ppm for NOx, and ±411.256 ppm for CO_2, respectively, underscoring the robustness of the integrated approach for emission prediction. This research aims to support the development of more sustainable vehicle technologies and inform policy making for environmental sustainability (e.g., Euro 6/Euro 7 standards). Overall, the study addresses how artificial intelligence (AI) can be utilized to achieve accurate multi-pollutant emission predictions in HEVs. The findings reveal that an integrated ANN-GPR approach yields superior predictive performance (R^2 values approaching 1.0) with quantifiable uncertainty, outperforming a stand-alone ANN model and providing a robust solution to the emission prediction challenge.

Corresponding Author:

Heru Priyanto
Research Center for Transportation Technology, National Research and Innovation Agency
St. Puspiptek No. 60, Setu, South Tangerang, Banten, 15314, Indonesia
Email: heru013@brin.go.id

1. INTRODUCTION

The increasing number of motor vehicles in recent decades has led to significant environmental concerns, particularly regarding air pollution from exhaust emissions. Vehicles powered by internal combustion engines (ICE) release harmful gases pollutants including carbon dioxide (CO_2), nitrogen oxides (NOx), and particulate matter (PM) lead to poor air quality and climate change. These pollutants negatively impact public health, ecosystems, and global climate stability. In response, governments and the automotive industry have been actively exploring environmentally friendly solutions to mitigate vehicle emissions. One promising innovation is the development of hybrid electric vehicles (HEVs), can enhance energy efficiency and lower pollution by combining an electric motor with an internal combustion engine. compared to conventional fossil-fuel-powered vehicles [1], [2].

Despite their potential benefits, HEVs still present challenges in accurately measuring and predicting emissions. Variations in driving patterns, road conditions, and driver behavior can cause fluctuations in emission levels, making it difficult to develop consistent regulatory and environmental strategies. To address these complexities [3], [4]. By leveraging real-time data from vehicle sensors, artificial intelligence (AI) algorithms can identify emission trends and provide accurate predictions under different operating conditions, offering a more effective approach to managing vehicle emissions [5].

This study explores how HEVs contribute to reducing emissions and examines the role of AI in enhancing emission prediction accuracy. The research aims to uncover the synergy between hybrid vehicle technology and AI-driven predictive models in achieving more sustainable and low-emission transportation systems [1], [6]. Motor vehicle emissions, including those from hybrid vehicles like the Toyota Prius, remain a major contributor to air pollution and climate change. While HEVs help mitigate emissions, their effectiveness varies due to operational conditions. A reliable predictive model is needed to address these inconsistencies and provide accurate estimations of emissions.

Artificial neural networks (ANN) are commonly utilized to predict emission trends, yet their inability to quantify prediction uncertainty can limit their effectiveness. In contrast, Gaussian process regression (GPR) is capable of estimating uncertainty, enhancing the reliability of predictive models [7], [8]. By integrating ANN and GPR, a more comprehensive and robust hybrid prediction model can be developed. This study seeks to answer key research questions: what are the emission characteristics of hybrid vehicles under diverse driving conditions? how can predictive models be designed to provide high-accuracy emission forecasts with measurable uncertainty?

Vehicular emission modeling is commonly employed to predict the amount of exhaust emissions produced by different vehicle types (carbon monoxide (CO), CO_2, hydrocarbons (HC), and NOx). Such models play a crucial role in environmental impact assessment, transportation policy development, and the advancement of sustainable automotive technologies. They also help analyze the influence of key factors fuel type, operating conditions, vehicle specifications on emissions. The urgency of addressing vehicular emissions stems from their direct implications on environmental sustainability and public health. While HEVs offer a promising solution, their emissions remain variable due to factors like driving behavior and road conditions. As a result, accurate and reliable emission prediction models are essential to facilitate better emission control strategies, inform policymakers, and support the development of sustainable transportation technologies.

This research aims to develop a multi-pollutant emission prediction model for hybrid vehicles using a combination of deep learning techniques and GPR. The model focuses on predicting emissions of CO, CO_2, NOx, and HC while ensuring both accuracy and reliability [9], [10]. By validating model performance against real-world data under United Nations ECE R83 standard driving cycles, this research offers insights on the strengths and weaknesses of different prediction techniques. Furthermore, analyzing the factors contributing to emission prediction uncertainty will enhance the model's practical applicability in emission management.

The expected outcomes of this research are significant for the development of HEV technology and the formulation of more effective emission-reduction strategies. The study is anticipated to yield an accurate and efficient multi-pollutant emission prediction model by integrating ANN and GPR. This model will facilitate deeper analysis of factors influencing emissions and offer better estimations for HEVs. To ensure practical applicability, model validation is conducted using test data compliant with United Nations Economic Commission for Europe (UN ECE) R83 standards, ensuring robustness in handling emission variability [11], [12].

In summary, this study addresses the critical question of how AI-based models can improve the accuracy of HEV emission predictions. To answer this question, we developed an integrated ANN-GPR model that combines the pattern-recognition strength of ANN with GPR's ability to quantify uncertainty. The approach involves training the hybrid model on real-world emission data from a Toyota Prius under standard urban and extra-urban driving cycles, and evaluating performance with rigorous validation metrics (root mean squared error (RMSE), mean absolute error (MAE), R^2). The results show that the ANN-GPR hybrid model achieves superior accuracy with R^2 values approaching 1.0 across all targeted pollutants and effectively captures prediction uncertainty, outperforming a standalone ANN in tracking emission trends. Consequently, this integrated modeling approach bridges a key gap in hybrid vehicle emission modeling, providing a robust predictive tool that can inform emission control policies and guide the development of cleaner vehicle technologies for sustainable transportation. Various stakeholders stand to benefit from these outcomes: the research community gains new insights into AI-driven emission modeling, the automotive industry can implement improved predictive controls for lower emissions, policymakers obtain a scientific basis for stricter emission standards, and society at large will benefit from improved air quality and environmental sustainability.

2. METHOD

This section describes the development of the emission prediction model using and GPR. It covers the data collection process, model development, and training procedures.

2.1. ANN model

Multiple ANN architectures were tested (5-20 neurons, single/multi hidden layers) using grid search and cross-validation. The final architecture with 10 neurons in one hidden layer was selected as it balanced accuracy with computational efficiency. ANN is a parallel processing approaches that can specifically describe non-linear and complex interactions using input-output data set training pattern [13], [14]. The ANN model operates in three stages: feed forward, back-propagation, and weight adjustment, calculated based on established equations [15], [16]. In the first stage, each input node receives an input value X_i (i = 1, 2, 3, ..., n) and forwards the signal to all nodes in the hidden layer. Each hidden layer node sums all weighted input signals (Z_{in_j}), where xi is multiplied by the weight (V_{ij}) and added to the received bias b_j (j= 1, 2, 3, ..., p) as shown in (1). The output signal from the hidden layer node is calculated using the activation function as in (2).

$$Z_{in_j} = b_j + \sum_{i=1}^{n} X_i V_{ij} \tag{1}$$

$$Z_j = f(Z_{in_j}) \tag{2}$$

In (1) and (2) represent the weighted sum and activation functions in the hidden layer.

Each output layer node y_k (k = 1, 2, 3, ..., m), sums all signals from the hidden layer nodes (multiplied by weights wjk, and added to the bias b_k) as in (3). The output signal from the output node is then calculated using the activation function (4).

$$y\_in_k = b_k + \sum_{j=1}^{p} z_i w_{jk} \tag{3}$$

$$y_k = f(y\_in_k) \tag{4}$$

In (3) and (4) are for weighted sum and activation in output layer.

In the second stage (back-propagation), the error information (δ_k) between each output node (y_k) and the target value (t_k) related to the training data is calculated as in (5). To adjust the weights and biases, corrections for weights (Δw_{jk}) and biases (Δb_k) are calculated using a predetermined learning rate (α) as given in (6) and (7).

$$\delta_k = (t_k - y_k) f'(y\_in_k) \tag{5}$$

$$\Delta w_{jk} = \alpha\ \delta_k Z_j \tag{6}$$

$$\Delta b_k = \alpha \delta_k \tag{7}$$

The error calculation and weight/bias update for the output layer are expressed in (5)-(7).

In the third stage (weight adjustment), the error information (δ_j) between each hidden layer node (Z_j, j = 1, 2, 3, ..., p) and the input layer nodes is calculated as in (8). Weight corrections (ΔV_{ij}) and bias corrections (Δb_j) are computed to adjust the weight (V_{ij}) and bias (b_j) values as shown in (9) and (10) using the learning rate (α).

$$\delta_j = \left(\sum_{k=1}^{m} \delta_k w_{jk}\right) f'\left(Z_{in_j}\right) \tag{8}$$

$$\Delta V_{ij} = \alpha\ \delta_j Z_j \tag{9}$$

$$\Delta b_j = \alpha \delta_j \tag{10}$$

In (8)-(10) are error propagation and update for hidden layer.

After these adjustments, each output unit is updated with new weight and bias values as in (11) and (12), and similarly each hidden unit's parameters are updated as in (13) and (14).

$$w_{jk}(new) = w_{jk}(old) + \Delta w_{jk} \tag{11}$$

$$b_k(new) = b_k(old) + \Delta b_k \tag{12}$$

$$V_{ij}(new) = V_{ij}(old) + \Delta V_{ij} \tag{13}$$

$$b_j(new) = b_j(old) + \Delta b_j \tag{14}$$

In (11)-(14) are final weight and bias update equations. The architecture of the ANN is shown in Figure 1.

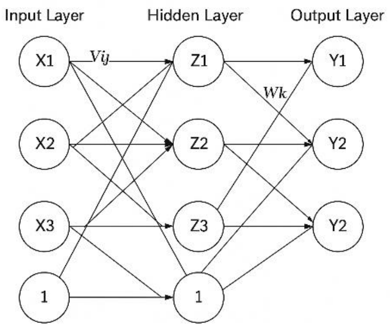

Figure 1. ANN architecture

2.2. Experimental setup and data collection

The vehicle tests were conducted at the Thermodynamics Motor and Propulsion Systems Laboratory of the National Research and Innovation Agency. The tests followed the current Indonesian standards, specifically the UN ECE R83 regulations [17], [18]. The vehicle was placed on a chassis dynamometer and operated using two drive cycles: urban drive cycle (UDC) and extra UDC (EUDC). The car was driven for approximately 1,180 seconds, and measurements were taken for several emission parameters, including CO, CO_2, HC, and NOx. Additionally, environmental parameters such as temperature, pressure, gas flow rate, humidity, and vehicle speed were recorded. Figure 2 illustrates the UN ECE R83 driving cycle.

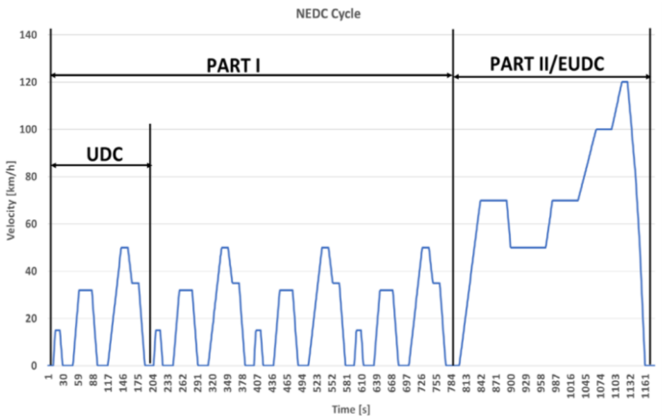

Figure 2. UN ECE R83 driving cycle

2.3. GPR model

GPR is a non-parametric Bayesian regression method widely used in machine learning due to its effectiveness with small datasets and its ability to provide predictive uncertainty measures [19], [20]. GPR models the distribution of an unknown target function based on training data, under the assumption that the data can be described by a GPR. It defines a prior over functions and uses a covariance function (kernel) to describe the relationship between any two data points. The choice of kernel function determines the shape of the inferred relationship. Several kernels (radial basis function (RBF), squared exponential (SE), and combined) were compared. The polynomial kernel was chosen due to its superior performance on emission data, yielding lower RMSE and higher coefficient of determination (R^2) compared to others while maintaining reasonable computational cost. In this study, a polynomial kernel was employed for the GPR model to capture non-linear trends in the emission data. By training the GPR model on the available data, we obtain not only a prediction for each output (emission level) but also an estimate of the uncertainty (confidence interval) associated with that prediction. This feature of GPR is particularly useful for understanding the reliability of the model's estimates under various conditions.

2.4. Model development and training

The integrated modeling approach was developed and evaluated using MATLAB (neural network fitting tool (NFTtool) for ANN training, combined with custom routines for GPR). Twelve input parameters were selected as model features, encompassing vehicle dynamics and environmental conditions: vehicle speed, air pressure, relative humidity, absolute humidity, dry-bulb air temperature, wet-bulb temperature, dew point temperature, exhaust gas flow volume, exhaust gas pressure, exhaust gas temperature, exhaust gas flow rate, and a correction factor for NOx. The target outputs for prediction were the emission concentrations of CO, CO_2, HC, and NOx. The twelve chosen input parameters were justified: vehicle speed affects load, humidity/temperature influence combustion chemistry, and exhaust parameters capture engine-out and after-treatment performance.

Using the ANN NFTool, we configured a feed-forward neural network with 12 input nodes, a single hidden layer containing 10 neurons, and 1 output node (for each emission variable, a separate network was trained). The Levenberg-Marquardt (LM) algorithm was employed for network training (Figure 3 shows the chosen network architecture). The dataset consisted of 1,180 data points recorded from the drive cycle experiments; we partitioned this data into 70% for training, 15% for validation, and 15% for testing (as illustrated in Figure 4). The LM training algorithm was chosen for its fast convergence and efficient memory usage (Figure 5 depicts the training performance curves) [21], [22]. After training the ANN, we recorded the mean squared error (MSE) and coefficient of determination (R^2) on the training and validation sets as preliminary performance indicators. Additionally, the trained models were evaluated on the independent test dataset using RMSE and MAE to quantify predictive performance [23], [24]. These validation metrics, together with R^2, provide a comprehensive assessment: RMSE and MAE measure the average deviation of predictions from actual values (with RMSE giving greater weight to larger errors), whereas R^2 reflects the proportion of emission variance explained by the model. After training, the resulting ANN model provides point predictions for emissions but does not inherently quantify prediction uncertainty.

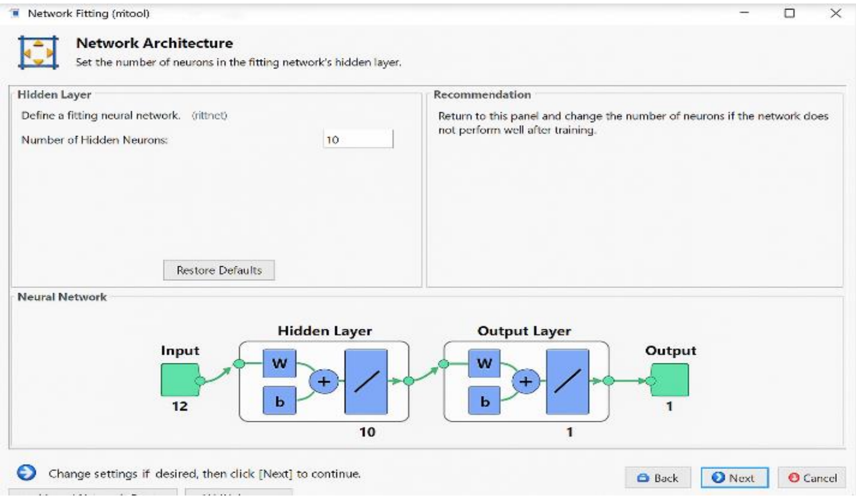

Figure 3. Size of the network formed

For the GPR model, the same training and test datasets were used to learn the relationship between the input features and each emission output. The GPR was configured with a polynomial kernel and optimized hyperparameters to balance bias and variance for the given data. Training the GPR model yields a predictive mean function (which gives the emission estimate) and an associated standard deviation (which can be translated into a prediction interval for uncertainty quantification) [25], [26].

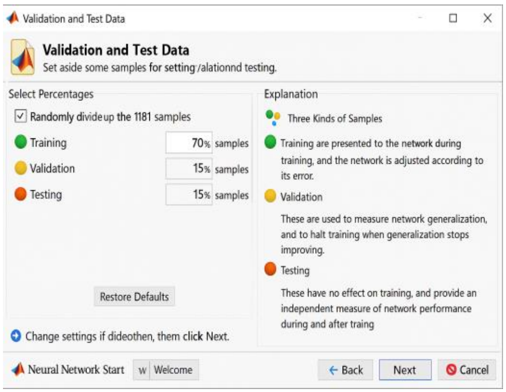

Figure 4. Menu and test data

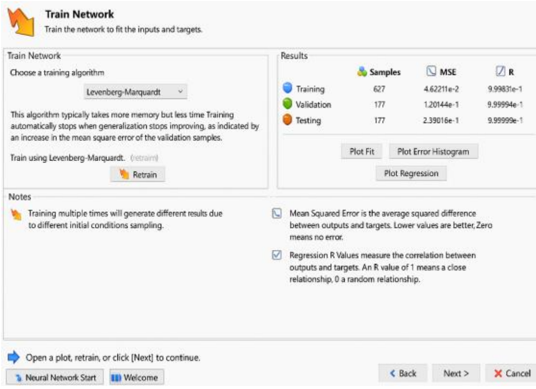

Figure 5. Network training results

Finally, we combined the strengths of both models: the ANN was used to capture the underlying nonlinear mapping from inputs to emission outputs, and the GPR was applied in parallel to estimate the uncertainty of those predictions. The performance of the hybrid ANN-GPR approach was then validated on the test set. We evaluated the model's generalization by comparing predicted emission values against the actual measured values for the test data, using RMSE, MAE, and R^2 as key performance metrics. High R^2 values alongside low RMSE and MAE on the test set would serve as empirical evidence that the model accurately captures emission patterns, while the GPR's output provides insight into the confidence of each prediction.

3. RESULTS AND DISCUSSION

3.1. Model performance and evaluation

The performance of the ANN and GPR models was evaluated using several metrics, including RMSE, MAE, and the coefficient of determination (R^2) [27], [28]. For these metrics, lower RMSE and MAE values correspond to more accurate predictions, and the R^2 value closer to 1 show that the model explains a larger portion of the variance in the emission data. The models were validated using the held-out test dataset to assess predictive performance for each pollutant, and their results were compared against each other as well as against expectations from traditional modeling approaches.

Benchmarking was also conducted with support vector machine (SVM), random forest, and gradient boosting models. These alternatives performed adequately but were less accurate and more computationally expensive than the ANN-GPR hybrid, reinforcing the superiority of the proposed method. Instances of underperformance (e.g., NOx spikes during acceleration) indicate unmodeled catalyst efficiency; these insights can guide sensor or feature expansion. R^2 values close to 0.999 mean that the model almost perfectly replicates emission dynamics. In practice, this accuracy ensures emission control systems in HEVs can anticipate pollutant spikes and comply with Euro 6/7 standards.

Table 1 presents a quantitative comparison between the ANN based model and the GPR model for each pollutant. The R^2 values in this Table 1 indicate each model's ability to explain variance in the emission data, while the MSE provides a measure of prediction error. The GPR model consistently outperformed the ANN model across all pollutants, as evidenced by higher R^2 and lower error values. The substantially lower MSE values for GPR (especially notable in CO_2 and NOx predictions) highlight its superior accuracy. Overall, GPR provided more reliable estimations with lower prediction errors, making it a robust choice for emission modeling in this context.

Table 1. Comparison between GPR and ANN result

Emission	Method	MSE	R^2
CO	GPR	0.1479	0.999
	ANN	17.6873	0.982
CO_2	GPR	411.2558	0.997
	ANN	203205.31679	0.967
NOx	GPR	0.1435	0.999
	ANN	245.58293	0.987
HC	GPR	37.6483	0.996
	ANN	0.141539	0.995

(Note: CO, CO_2, NOx and HC emissions values in ppm)

3.2. Prediction accuracy and trends

Beyond the aggregate error metrics, the models' behavior over time and across driving conditions was examined. Figures 6 to 9 illustrate the measured vs. predicted emission trajectories for CO, CO_2, HC, and NOx, respectively, over the course of the urban and EUDC. The CO emission predictions show that the hybrid ANN-GPR model successfully captures the fluctuation patterns observed in the actual measurements for both drive cycles. The predictive traces closely follow the real emission curves, with only minor deviations. These small discrepancies can be attributed to external environmental factors (such as transient changes in temperature or pressure) that were not fully captured by the model inputs. Overall, the high agreement between predicted and observed CO values confirms the model's effectiveness in learning the emission dynamics of HEV operation.

Similarly, for CO_2 emissions, throughout the driving cycles, the model's predictions closely match the empirical data. In particular, the GPR component of the model provides confidence intervals (shaded regions in Figure 7) that remain narrow in most segments, indicating high certainty in the predictions. Even during segments where the vehicle transitioned between urban and extra-urban phases (which can cause significant variability in CO_2 output due to changing engine loads and regenerative braking in the HEV), the model maintained accurate estimates. This performance reflects the model's ability to account for variations in combustion efficiency and energy management between different driving regimes.

For HC and NOx emissions, the hybrid model again demonstrates strong predictive capabilities. The predicted HC levels Figure 8 track the measured data well, capturing rapid changes that occur during acceleration and deceleration events. The GPR uncertainty bounds for HC are slightly wider during certain high variability segments, suggesting that the model rightly expresses less confidence in predictions under those transient conditions potentially due to unmodeled factors like catalyst efficiency variation. In the case of NOx Figure 9, the model's predictions not only follow the actual emission trend but also highlight specific points in the drive cycle where uncertainty is elevated. These typically correspond to instances of aggressive acceleration or high engine load, which are known to produce spikes in NOx emissions. Identifying these regions of higher uncertainty is valuable: it indicates where the model (and by extension, emissions control systems) might need to pay extra attention or where additional sensor inputs might be required to improve accuracy.

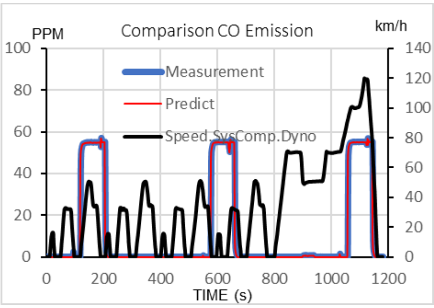

Figure 6. Measurement and prediction graphs of CO emissions

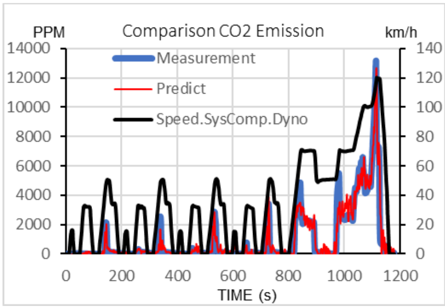

Figure 7. Measurement and prediction graphs of CO_2 emissions

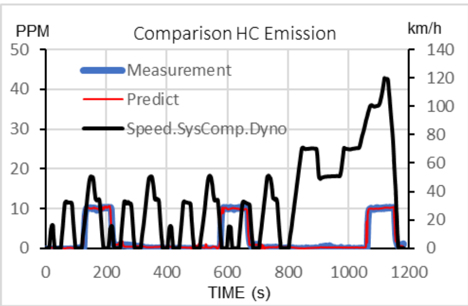

Figure 8. Measurement and prediction graphs of HC emissions

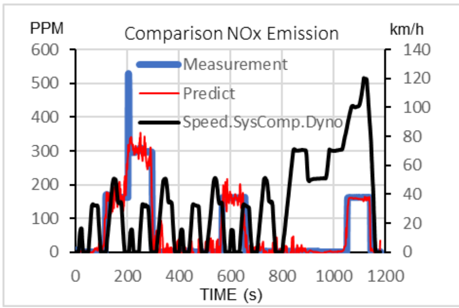

Figure 9. Measurement and prediction graphs of NOx emissions

3.3. Uncertainty analysis

A distinct advantage of the GPR component is its ability to quantify predictive uncertainty. To evaluate this aspect, the prediction uncertainties of the ANN and GPR models were compared for each pollutant, as summarized in Table 2 the GPR model demonstrated consistently lower uncertainty (narrower prediction intervals) across all pollutants compared to the ANN model, reinforcing the reliability of the integrated approach in emission modelling [29], [30]. These results show that not only does GPR achieve higher accuracy, but it also provides more confidence in each prediction, an important consideration for practical applications.

Table 2. Comparison of uncertainty result

Emission	Method	Uncertainty (ppm)
CO	GPR	±0.829
	ANN	±1.237
HC	GPR	±9.978
	ANN	±13.756
NOx	GPR	±0.144
	ANN	±0.199
CO_2	GPR	±411.256
	ANN	±426.763

Being able to quantify uncertainty is particularly valuable for regulatory and operational purposes. Policymakers can utilize these insights to refine vehicle emission standards and develop targeted mitigation strategies for high-emission scenarios. For instance, if the model identifies a certain driving condition that leads to high uncertainty in NOx predictions, regulators might investigate that scenario further or set specific guidelines to manage it. Automotive manufacturers and engineers can also leverage an uncertainty-aware model to improve real-time emission control systems in HEVs, ensuring that the vehicle's control algorithms account for conditions where emissions are less predictable. In summary, incorporating GPR-based uncertainty analysis adds a layer of transparency and robustness to emission prediction that purely deterministic models (like a standard ANN) would lack. Table 2 predictive uncertainty comparison between GPR and ANN models for each pollutant. (Uncertainty values represent an approximate ± range around the predicted mean).

Compared with previous works, this study advances the literature by combining ANN's nonlinear mapping capability with GPR's strength in uncertainty quantification. GPR's superiority for CO_2 prediction in hybrids and potential in laboratory engines, the present study extends these insights to real-world hybrid conditions [27], [31], [32]. Overall, the integrated ANN-GPR model shows strong potential as a decision-support tool for emission regulation, adaptive control strategies, and sustainable vehicle technology development.

3.4. Limitations and future work

The proposed ANN-GPR emission prediction model shows promise, but several limitations remain. It was trained and tested only on one HEV model (Toyota Prius) under specific urban and extra-urban cycles, limiting its generalizability to other vehicles or conditions such as rural driving, high-speed cruising, or extreme weather. Furthermore, validation relied on UN ECE R83 standards, which may not represent global driving cycles. Another challenge lies in the computational intensity of GPR, which could hinder real-time applications as data size increases.

Future research should address these constraints by expanding datasets to include multiple HEV types, additional driving cycles, and extreme environmental conditions. Broader data coverage, aligned with international standards, would improve robustness across regions. On the computational side, applying sparse or approximate GPR methods (e.g., inducing point approximations) could significantly reduce complexity while maintaining accuracy. Moreover, exploring hybrid frameworks such as ANN-bayesian neural network (BNN) or ANN-long short-term memory (LSTM) would enrich uncertainty estimation and capture temporal dependencies, offering stronger predictive capabilities for emission dynamics.

Another important direction is the integration of adaptive learning mechanisms. An onboard system that updates itself with real-time sensor data would allow the model to maintain accuracy as vehicles age or encounter new conditions. Such adaptability could reduce uncertainty during atypical scenarios and enhance predictive reliability. Ultimately, broadening data scope, improving efficiency, and enabling adaptive learning will strengthen the model's applicability for real-world emission control and sustainable transportation.

4. CONCLUSION

This study developed and validated an integrated ANN-GPR framework for predicting multi-pollutant emissions in HEV under standard driving cycles. The results showed that ANN was able to capture general emission trends, but GPR consistently outperformed it with higher coefficients of determination and lower prediction errors across all pollutants. In addition, the GPR component provided quantifiable uncertainty estimates, making the predictions not only accurate but also more reliable. These findings confirm that integrating ANN and GPR enhances model robustness, offering a more effective approach than either method alone. Beyond its technical contribution, this integrated framework provides a scientific foundation for emission management and regulatory development toward Euro 6/7 standards. The ability to quantify predictive uncertainty makes it suitable for real-time monitoring and adaptive emission control in HEVs.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to Mr. Dr. Aam Muharam, M.Eng., Head of the Research Center for Transportation Technology National Research and Innovation Agency, for his invaluable guidance and continuous support throughout this research. His expertise and insightful suggestions have significantly contributed to the development of this study. We also extend our deepest appreciation to Mr. Mokhtar, S.T., M.T., Head of the Thermodynamics, Engines, and Propulsion Laboratory National Research and Innovation Agency, for providing the necessary research facilities and technical support that made this work possible. His encouragement and assistance have been instrumental in ensuring the smooth progress of our experiments and analyses. we would also like to express our sincere gratitude to Universiti Teknologi Malaysia (UTM) for providing access to the software facilities that greatly supported the completion of this research work. Furthermore, we would like to acknowledge all colleagues and staff members who have contributed, directly or indirectly, to the completion of this research. Their valuable input and cooperation are greatly appreciated. Finally, we are grateful for the support from our institution and ELSA funds, which have facilitated this research endeavor.

FUNDING INFORMATION

This research was made possible through the collective financial contributions of the co-authors, utilizing ELSA funds to support the experimental activities conducted at the Thermodynamics, Engines, and Propulsion Laboratory. We sincerely appreciate the dedication and commitment of all contributors who have played a crucial role in ensuring the successful completion of this study.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Heru Priyanto	✓	✓	✓	✓	✓			✓	✓	✓				✓
Rizqon Fajar	✓	✓				✓		✓		✓	✓	✓		✓
Yaaro Telaumbanua	✓		✓	✓	✓		✓		✓		✓	✓	✓	✓
Ariyanto	✓		✓		✓	✓				✓		✓	✓	✓
Muhammad Mukhlas Af		✓		✓		✓	✓		✓			✓	✓	✓
Sigit Tri Atmaja					✓		✓	✓	✓	✓				✓
Muhammad Samsul Maarif		✓	✓	✓		✓		✓	✓					✓
Kurnia Fajar Adhi Sukra	✓		✓		✓		✓			✓	✓	✓	✓	✓
Fauzi Dwi Setiawan		✓		✓	✓			✓		✓				✓

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest regarding the publication of this paper.

INFORMED CONSENT

This study does not involve human participants, personal data, or confidential information requiring informed consent.

ETHICAL APPROVAL

This research does not involve human participants, animals, or any sensitive data requiring ethical approval.

DATA AVAILABILITY

The data supporting the findings of this study are available upon reasonable request. Interested researchers may contact the first author for access to the data.

REFERENCES

[1] G. Keegan, P. Nelendran, and O. Oluwafemi, "Modeling and simulation of hybrid electric vehicles for sustainable transportation: insights into fuel savings and emissions reduction," *Energies*, vol. 17, no. 20, 2024, doi: 10.3390/en17205225.

[2] S. Khalid, "A comprehensive review of advanced control strategies for hybrid electric vehicles," *Hybrid Electric Vehicles and Distributed Renewable Energy Conversion: Control and Vibration Analysis*, pp. 55-98, 2024, doi: 10.4018/979-8-3693-5797-2.ch002.

[3] C. Maino, D. Misul, A. Di Mauro, and E. Spessa, "A deep neural network based model for the prediction of hybrid electric vehicles carbon dioxide emissions," *Energy and AI*, vol. 5, p. 100073, 2021, doi: 10.1016/j.egyai.2021.100073.

[4] Y. Mehta, R. Xu, B. Lim, J. Wu, and J. Gao, "A review for green energy machine learning and AI services," *Energies*, vol. 16, no. 15, p. 5718, 2023, doi: 10.3390/en16155718.

[5] J. Udoh, J. Lu, and Q. Xu, "Application of machine learning to predict CO2 emissions in light-duty vehicles," *Sensors*, vol. 24, no. 24, p. 8219, 2024, doi: 10.3390/s24248219.

[6] S. Katreddi and A. Thiruvengadam, "Trip based modeling of fuel consumption in modern heavy-duty vehicles using artificial intelligence," *Energies*, vol. 14, no. 24, pp. 1–12, 2021, doi: 10.3390/en14248592.

[7] D. Zhong, B. Liu, L. Liu, W. Fan, and J. Tang, "Artificial intelligence algorithms for hybrid electric powertrain system control: A review," *Energies*, vol. 18, no. 8, 2025, doi: 10.3390/en18082018.

[8] S. R. Yoo, J. W. Shin, and S. H. Choi, "Machine learning vehicle fuel efficiency prediction," *Scientific Reports*, vol. 15, no. 1, pp. 1–20, 2025, doi: 10.1038/s41598-025-96999-0.

[9] A. A. Alhussan, M. Metwally, and S. K. Towfek, "Predicting CO2 emissions with advanced deep learning models and a hybrid greylag goose optimization algorithm," *Mathematics*, vol. 13, no. 9, 2025, doi: 10.3390/math13091481.

[10] J. C. Ferreira and M. Esperança, "Enhancing sustainable last-mile delivery: The impact of electric vehicles and AI optimization on urban logistics," *World Electric Vehicle Journal*, vol. 16, no. 5, pp. 1–15, 2025, doi: 10.3390/wevj16050242.

[11] R. Yang *et al.*, "An artificial neural network model to predict efficiency and emissions of a gasoline engine," *Processes*, vol. 10, no. 2, pp. 1–20, 2022, doi: 10.3390/pr10020204.

[12] D. Tena-Gago, G. Golcarenarenji, I. Martinez-Alpiste, Q. Wang, and J. M. Alcaraz-Calero, "Machine-learning-based carbon dioxide concentration prediction for hybrid vehicles," *Sensors*, vol. 23, no. 3, 2023, doi: 10.3390/s23031350.

[13] V. Yadav, A. K. Yadav, V. Singh, and T. Singh, "Artificial neural network an innovative approach in air pollutant prediction for environmental applications: A review," *Results in Engineering*, vol. 22, no. May, p. 102305, 2024, doi: 10.1016/j.rineng.2024.102305.

[14] V. Chandran, C. K. Patil, A. Karthick, D. Ganeshaperumal, R. Rahim, and A. Ghosh, "State of charge estimation of lithium-ion battery for electric vehicles using machine learning algorithms," *World Electric Vehicle Journal*, vol. 12, no. 1, 2021, doi: 10.3390/wevj12010038.

[15] Y. E. Altun and O. A. Kutlar, "Energy management systems' modeling and optimization in hybrid electric vehicles," *Energies*, vol. 17, no. 7, 2024, doi: 10.3390/en17071696.

[16] W. M. Ashraf *et al.*, "Machine learning assisted improved desalination pilot system design and experimentation for the circular economy," *Journal of Water Process Engineering*, vol. 63, p. 105535, 2024, doi: 10.1016/j.jwpe.2024.105535.

[17] United Nations Economic Commission for Europe (UNECE), "Regulation no. 83: Uniform provisions concerning the approval of vehicles with regard to the emission of pollutants according to engine fuel requirements," *Official Journal of the European Union*, vol. 84, no. 11, pp. 20-226, 2012.

[18] K. F. A. Sukra, H. Priyanto, D. Indriatmono, M. A. Wijayanto, I. Y. Ikhsanudin, and Y. A. Ermansyah, "Impact of road load parameters on vehicle CO2 emissions and fuel economy: A case study in Indonesia," *Journal of Mechatronics, Electrical Power, and Vehicular Technology*, vol. 14, no. 1, pp. 87–93, 2023, doi: 10.14203/j.mev.2023.v14.87-93.

[19] R. C. Mamat, A. Ramli, and A. B. Bawamohiddin, "A comparative analysis of Gaussian process regression and support vector machines in predicting carbon emissions from building construction activities," *Journal of Advanced Research in Applied Mechanics*, vol. 131, no. 1, pp. 185–196, 2025, doi: 10.37934/aram.131.1.185196.

[20] M. O. Oyedeji, M. Aldhaifallah, H. Rezk, and A. A. A. Mohamed, "Computational models for forecasting electric vehicle energy demand," *International Journal of Energy Research*, vol. 2023, no. 01,p. 1934188, 2023, doi: 10.1155/2023/1934188.

[21] Z. Yan, S. Zhong, L. Lin, and Z. Cui, "Adaptive Levenberg–Marquardt algorithm: A new optimization strategy for Levenberg–Marquardt neural networks," *Mathematics*, vol. 9, no. 17, 2021, doi: 10.3390/math9172176.

[22] R. L. G. Kakaï, M. Egah, C. G. Hounmenou, "Effect of dataset partition and normalization methods in the hyperparameters

optimization phase on the efficiency of the Levenberg Marquardt algorithm," *Computer Science and Mathematics*, pp. 0–16, 2024, doi: 10.20944/preprints202411.2210.v1.
[23] A. Rashid and S. Kumari, "Performance evaluation of ANN and ANFIS models for estimating velocity and pressure in water distribution networks," *Water Supply*, vol. 23, no. 9, pp. 3925–3949, 2023, doi: 10.2166/ws.2023.224.
[24] M. Li, M. Yan, H. He, and J. Peng, "Data-driven predictive energy management and emission optimization for hybrid electric buses considering speed and passengers prediction," *Journal of Cleaner Production*, vol. 304, p. 127139, 2021, doi: 10.1016/j.jclepro.2021.127139.
[25] X. Shi, D. Jiang, W. Qian, and Y. Liang, "Application of the Gaussian process regression method based on a combined kernel function in engine performance prediction," *ACS Omega*, vol. 7, no. 45, pp. 41732–41743, 2022, doi: 10.1021/acsomega.2c05952.
[26] S. Srivastava, Z. Xu, Y. Li, W. N. Street, and S. Gilbertson-White, "Gaussian process regression and classification using International classification of disease codes as covariates," *Stat*, vol. 12, no. 1, pp. 1–18, 2023, doi: 10.1002/sta4.618.
[27] F. Ricci, M. Avana, and F. Mariani, "A deep learning method for the prediction of pollutant emissions from internal combustion engines," *Applied Sciences*, vol. 14, no. 21, 2024, doi: 10.3390/app14219707.
[28] A. A. Ajala, O. L. Adeoye, O. M. Salami, and A. Y. Jimoh, "An examination of daily CO2 emissions prediction through a comparative analysis of machine learning, deep learning, and statistical models," *Environmental Science and Pollution Research*, vol. 32, no. 5, pp. 2510–2535, 2025, doi: 10.1007/s11356-024-35764-8.
[29] S. Lim and J. Oh, "Hybrid neural network-based maritime carbon dioxide emission prediction: Incorporating dynamics for enhanced accuracy," *Applied Sciences*, vol. 15, no. 9, 2025, doi: 10.3390/app15094654.
[30] M. Holzenkamp, D. Lyu, U. Kleinekathöfer, and P. Zaspel, "Evaluation of uncertainty estimations for Gaussian process regression based machine learning interatomic potentials," *Machine Learning*, 2024, doi: 10.48550/arXiv.2410.20398.
[31] M. Mądziel, A. Jaworski, H. Kuszewski, P. Woś, T. Campisi, and K. Lew, "The development of CO2 instantaneous emission model of full hybrid vehicle with the use of machine learning techniques," *Energies*, vol. 15, no. 1, p. 142, Dec. 2021, doi: 10.3390/en15010142.
[32] F. Millo, L. Rolando, L. Tresca, and L. Pulvirenti, "Development of a neural network-based energy management system for a plug-in hybrid electric vehicle," *Transportation Engineering*, vol. 11, p. 100156, 2023, doi: 10.1016/j.treng.2022.100156.

BIOGRAPHIES OF AUTHORS

Heru Priyanto holds a Master of Engineering in Electronic Control from University of Indonesia. His affiliation is the National Research and Innovation Agency (BRIN), Indonesia at Transportation Technology Research Center. His research focused on control systems, emission vehicle, data aquitition, machine learning and simulation. He can be contacted at email: heru013@brin.go.id.

Rizqon Fajar affiliation is the National Research and Innovation Agency (BRIN), Indonesia at Transportation Technology Research Center. He has published on fuel combustion, lubricant performance, and data science in engineering. He can be contacted at email: rizq001@brin.go.id.

Yaaro Telaumbanua is a researcher at BRIN, Serpong. He earned a Bachelor's in Electronics and Instrumentation from Politeknik ITB and a Master's in Electro-Telecommunication from Gunadarma University. His interests include transportation technologies, machine learning, control systems, artificial intelligence, and ICT applications. He can be contacted at email: yaar001@brin.go.id.

Ariyanto received his M.Sc. in Instrumentation Science from the University of Indonesia, Indonesia. He is an Engineer in the Research Center for Transportation Technology, National Research and Innovation Agency, South Tangerang, Indonesia. His research interests include instrumentation and measurement, machine learning, and data science. He can be contacted at email: ariy001@brin.go.id.

Mohammad Mukhlas Af is a researcher at BRIN, Serpong. He holds a Bachelor's in Electronics and Instrumentation from Politeknik Brawijaya and a Master's in Physics Instrumentation from the University of Indonesia. His work covers transportation technologies, control systems, instrumentation, machine learning, data science, and AI. He can be contacted at email: mohm001@brin.go.id.

Sigit Tri Atmaja holds a Master of Engineering (M.T.) from the University of Indonesia. He is an expert in intelligent control systems, the internet of things, and machine learning in the Research Center for Transportation Technology, National Research and Innovation Agency, South Tangerang, Indonesia. He can be contacted at email: sigit.tri.atmaja@brin.go.id.

Muhammad Samsul Maarif is an assistant researcher at BRIN, Serpong, and an M.Phil. student at Universiti Teknologi Malaysia (UTM). He holds a Bachelor's in Electrical Engineering from Universitas Gadjah Mada. His research focuses on machine learning, control systems, AI, instrumentation, and IoT for transportation and energy. He can be contacted at email: muha127@brin.go.id.

Kurnia Fajar Adhi Sukra is a researcher at BRIN and a Ph.D. student at Universiti Teknologi Malaysia. He earned a Master's in Mechanical Engineering from Institut Teknologi Bandung (ITB). His research expertise includes combustion engines, thermodynamics, and machine learning applications. He can be contacted at email: kurn021@brin.go.id.

Fauzi Dwi Setiawan is an assistant researcher at BRIN, Serpong. He holds a Bachelor's in Electronics and Instrumentation from Universitas Gadjah Mada and a Master's in ICT Policy Convergence from Handong Global University, South Korea. His research interests include machine learning, data science, AI, and ICT applications for transportation. He can be contacted at email: fauz015@brin.go.id.