Utilizing Data-Driven Techniques to Improve Predictive Modeling of Connected Electric Vehicle Energy Consumption

Ahmed-Ramzi Houalef, Florian Delavernhe, Sidi-Mohammed Senouci, El-hassane Aglzim
Université de Bourgogne, DRIVE Lab, Nevers, France, email: {firstname.lastname}@u-bourgogne.fr

***Abstract*—Electric vehicles (EVs) have emerged as a promising solution for environmental preservation. However, a major hurdle remains: range anxiety caused by limited driving range and inaccurate energy consumption estimates. This is because energy use in EVs varies significantly based on environmental factors (like temperature) and driving conditions (like traffic congestion). To address this challenge, this paper leverages the recent trend of connected electric vehicles (CEVs). These vehicles are equipped with sensors and onboard computers that gather real-time data on various aspects that affect energy consumption. The paper then introduces a Tri-component data-driven machine-learning model that utilizes this data from connected EVs. The model focuses on optimizing route planning for energy efficiency by predicting energy consumption along different road segments. It considers various external factors like temperature, traffic congestion, and road incline to predict three key elements: velocity (how fast the vehicle will go), traction power (energy needed to move the vehicle), and auxiliary power (energy used by features like A/C and onboard computer). Tested in real-world scenarios, the model demonstrates a significant reduction in energy consumption estimation errors, with a remarkable 2.38% error rate for battery state-of-charge (SoC) and 0.52 kWh for energy consumption.**

***Index Terms*—electric vehicles, route planning, energy efficiency, machine learning, energy consumption.**

I. Introduction

The rise of electric vehicles (EVs) is a boon for the environment, but range anxiety – the fear of running out of charge – remains a major hurdle. Consumers worry about travel distance, especially under varying conditions. This apprehension significantly impacts EV adoption. To address range anxiety, various strategies are employed. These include increasing battery capacity, improving energy management systems, expanding charging infrastructure, and optimizing charging times [1]. However, another powerful tool exists: smart route planning [2] specifically designed for EVs. Traditional route planning focuses on distance or time, but it doesn't guarantee energy efficiency for EVs. Factors like speed, acceleration, road incline, and regenerative braking all play a role. This is where energy-efficient route planning comes in, considering energy consumption for each road segment and using that data for route optimization. Existing models for energy-efficient route planning fall into three categories based on their energy calculation methods: physics-based (using scientific principles), statistics-based (finding relationships in historical data), or AI-based (learning complex relationships from data), each with varying levels of variable aggregation. High aggregation uses averages over long periods, while low aggregation uses frequent individual data points for more precise predictions. Physics-based and statistics-based often lack the adaptability needed for real-world scenarios. Here's where CEVs come in. Connected EVs are equipped with a wealth of sensors and onboard computers that gather real-time data on driving conditions, energy usage, and environmental factors. This data stream is a goldmine for AI-based route planning models. Unlike traditional models, data-driven machine learning, can analyze vast amounts of historical energy consumption data collected from connected EVs. This allows them to identify complex relationships between factors like traffic, weather, and driving behavior, empowering them to adapt predictions to ever-changing real-world conditions. By leveraging the power of data and connectivity, connected EVs can integrate seamlessly with AI-based route planning models. This combination offers a powerful solution to address range anxiety. It allows for highly precise energy consumption estimates, leading to more efficient routes and ultimately, a more practical and appealing choice for consumers. This paper introduces a data-driven machine learning framework for optimizing EV energy consumption and enabling energy-efficient route planning. The framework leverages three key components:

- **Long-Term Velocity Model**: This model predicts the expected speed for each road segment, considering factors like traffic patterns and road grade.
- **Hybrid Traction Power Model**: Utilizing the predicted velocity, this model estimates the energy required to propel the vehicle along each segment, by combining physics-based model with an Artificial Neural Network (ANN).
- **Auxiliary Power Model**: This model accounts for energy consumption by auxiliary systems like air conditioning and heating, based on factors like ambient temperature and user preferences.

By combining these predictions, the framework generates an estimate of the vehicle's energy consumption for different routes. This allows drivers and route planning systems to identify the most energy-efficient paths, ultimately maximizing driving range and minimizing environmental impact. The rest of the paper is organized as follows. Related work is detailed in Section II, the adopted modeling approach is detailed in Section III. The three proposed models, namely

Traction Power, Auxiliary Power, and Velocity are discussed in Sections III.C, III.D, and III.E, respectively. The experimental validation results are summarized in Section IV. Conclusions are in Section V.

II. Related Work

Driven by concerns over "range anxiety", EV adoption necessitates sophisticated route planning (itinerary planning) to maximize operational range. Route planning can be categorized into problems focusing on different objectives: minimizing travel time, energy consumption, distance, or a combination of these. In this work, we prioritize energy optimization, aiming to identify the route that minimizes energy expenditure for the journey. Energy-based route planning systems rely on a three-step process to achieve this goal. A Geographic Information System (GIS) gathers detailed geographical data on road segments. An energy model uses this data to estimate energy consumption for each segment, considering various factors and assigning weights accordingly. A shortest path algorithm then utilizes these weights to determine the route with the lowest overall energy consumption. Energy-based route planning systems can be categorized according to the aggregation level and calculation method of their energy model, in one hand High Aggregation Models rely on simplified assumptions and average values to estimate energy consumption, in the other hand Low Aggregation Models incorporate more granular data and realistic values to achieve higher accuracy. Calculation methods can be categorized into three main types: physics-based models, statistical data-driven models, and machine learning-based data-driven models. In [3], the authors employed a physics-based model to estimate battery SoC, incorporating a statistical model to account for traffic conditions. However, auxiliary power was not taken into consideration, which may lead to less accurate results. in [4] a physics-based energy model has been used with the assumption of constant auxiliary power, this assumption does not reflect real-life scenarios. As for velocity, they used a linear regression model. In [5], the authors proposed a data-driven approach based on historical data clustering. However, due to the absence of real energy data, they relied on a physics-based model for data generation. Additionally, for estimating auxiliary power, they utilized a temperature-dependent function sourced from a Nissan Leaf, this approach poses deployment challenges due to its data-intensive nature. In [6] the authors proposed a physics-based model for energy while using the maximum allowable road segment velocity (speed limit) as a reference and topography data was obtained from the "USGS Earth". Using the speed limit as a reference for energy calculation fails to capture real-life scenarios, as it overlooks driver behavior's impact on acceleration and deceleration. The authors in [7] proposed Eco-Routing, Eco-Driving, and Energy Consumption Prediction with physics-based model for energy and ANN for optimal speed recommendation. [8] developed a real-time range estimation based on calculating the energy-optimal route, considering traffic influences. In [9], the authors proposed a strategy to optimize EV route planning considering traffic impedance information. In [10] the authors proposed a Q-learning approach with a physics-based energy model, this approach has some limitations when it comes to the data size. while the previously mentioned work tackled route planning directly, others have tackled the energy and speed modeling separately. The authors of [11] tackled energy estimation using Floating Car Data and statistical methods. In [12] the authors explored different machine learning (ML) techniques to estimate energy consumption and identify the key factors influencing it. [13] has independently tackled the speed prediction problem. Current state-of-the-art strategies for electric vehicles often rely on simplified models, particularly physics-based ones. These models may not account for all environmental variables and often use highly aggregated data, leading to reduced accuracy. Additionally, some complex models can pose deployment challenges due to their computational demands. This manuscript proposes a data-driven, tri-component hybrid approach for accurate electric vehicle energy estimation. By leveraging realistic data and incorporating environmental variables, this approach offers improved prediction capabilities. Furthermore, it can be applied to various driver profiles and car models, requiring only new data for integration.

III. Proposed Approach for Energy Consumption Estimation

In this section, we will discuss the global system architecture, different key components and the used data for energy consumption estimation.

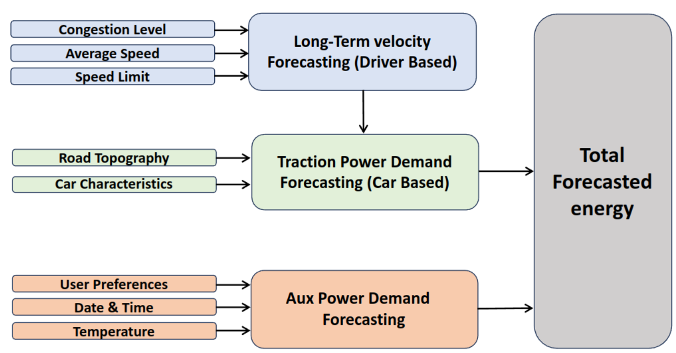


Fig. 1. Overview of the system architecture

A. System Architecture

As illustrated in Figure 1, an energy consumption estimation system leverages three key components. The first, velocity prediction, utilizes historical data on average segment speed, maximum allowable speed, traffic congestion levels and driver's to estimate the expected velocity for each road segment. This predicted velocity then serves as input for the traction power demand model, which calculates the energy required to propel the vehicle along the segment based on factors like vehicle weight and road topography. Finally, the auxiliary power model estimates the energy consumption of auxiliary systems like heating, ventilation, air conditioning (HVAC), and lighting, considering factors like ambient temperature and user preferences. The outputs from both the traction power demand model and the auxiliary power model are then combined to provide a comprehensive estimate of the vehicle's total energy consumption.

B. Data Processing

For our experiments, we used a Nissan Leaf dataset [14] that consists of 80 driving hours, with a frequency of 1 Hz, collected in Riverside, California. This rich dataset includes features like GPS points, speed, traction and auxiliary power consumption, altitude, ambient temperature, date, and time. Unfortunately, our approach necessitates additional data for implementation, notably the road slope (Grade). Due to the inadequacy of the available road slope data in the dataset, we decided to calculate it at each time step using. This involved utilizing data from the altitude sensor, time, and velocity, as outlined by Equation 1.

$$\theta = \arctan\left(\frac{Altitude_{\text{diff}}}{v \cdot t}\right) \tag{1}$$

Where $Altitude_{\text{diff}}$ denotes the difference between two consecutive altitudes. Other crucial information, such as average and max allowable speed, require matching each GPS point to a given road segment. Hence, we implemented an incremental matching algorithm with back-track using Open Street Map (OSM). By doing so, we can assign each collected data point to a road segment. The max allowable speed has been fetched from OSM directly while the global average speed for each road segment has been calculated using Equation 2. The details of the matching algorithm are not detailed in the present paper, as it is beyond the scope of this research paper.

$$v_l = \frac{1}{i \cdot j} \sum_{n=0}^{i} \sum_{t=0}^{j} v_n(t) \tag{2}$$

Where v_l is the average speed of the entire segment, v_n is the instantaneous speed at a given moment, i represents the number of times the segment is traversed across various driving cycles, and j denotes the duration of each passage through the segment. To determine whether the vehicle was operated during the night or day, thereby indicating the usage of car lighting, we conducted an analysis that integrated time and date information with sunset and sunrise data. This analytical process yielded a binary representation serving as an indicator of lighting usage. To measure the congestion level of each segment, we used the speed performance index (SPI) [15] as shown in Equation. 3. Congestion quantification levels are shown in Table. I

TABLE I
CONGESTION QUANTIFICATION LEVELS

Speed Performance Index	Traffic State Level
(0,25)	Heavy congestion
(25,50)	Mild congestion
(50,75)	Smooth
(75,100)	Very smooth

$$SPI = (v_{avg}/v_{max}) \cdot 100. \tag{3}$$

Where v_{avg} is the segment's average velocity in the corresponding passage, and v_{max} is the maximum allowable velocity in that segment. The acceleration can be directly derived from the speed.

C. Traction Power Consumption Forecasting

This subsection discusses the modeling approaches of the traction power model and compare the results of different approaches.

1) Modeling: To achieve a highly accurate power consumption model, we employed a multi-faceted approach that leverages physics-based, AI-based, and hybrid modeling techniques. The physics-based model, detailed in equations (4) to (9), calculates the power requirement based on fundamental physical principles by calculating different forces (Rolling, Climbing, Acceleration, and Aerodynamic).

$$F_{\text{roll}} = c_{\text{rr}} \cdot m \cdot g \tag{4}$$

$$F_{\text{climb}} = m \cdot g \cdot \sin(\theta) \tag{5}$$

$$F_{\text{aero}} = 0.5 \cdot c_d \cdot A \cdot \rho \cdot v^2 \tag{6}$$

$$F_{\text{acc}} = m \cdot a \tag{7}$$

$$F_{\text{Traction}} = F_{\text{aero}} + F_{\text{acc}} + F_{\text{grad}} + F_{\text{roll}} \tag{8}$$

$$P_{\text{Net}} = \begin{cases} F_{\text{Traction}} \cdot v \cdot \beta & \text{if } F_{\text{Traction}} < 0 \\ F_{\text{Traction}} \cdot v/\alpha & \text{if } F_{\text{Traction}} \geq 0 \end{cases} \tag{9}$$

$$P_{\text{Total}} = P_{\text{Aux}} + P_{\text{Net}} \tag{10}$$

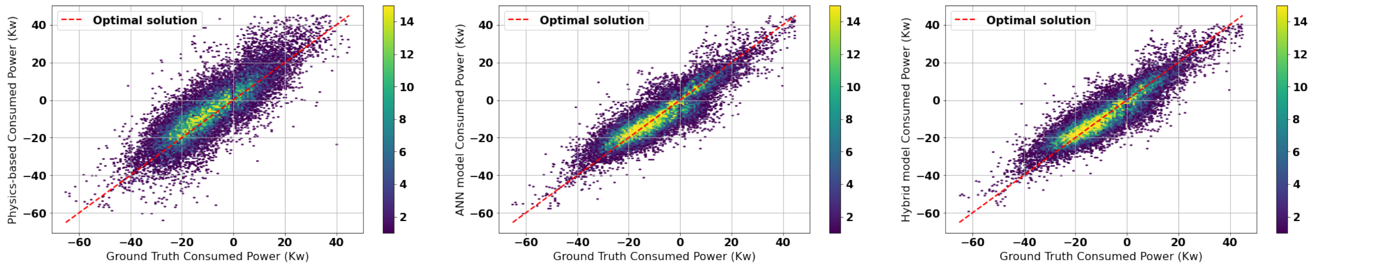
$$E_{\text{Total}} = \frac{\delta_t}{3600} \int_a^b P_{\text{Total}}(t)\, dt \tag{11}$$

For the AI-based model, we adopted a simple four-layer Artificial Neural Network (ANN) architecture. While the specific choice was informed by experimentation, the overall approach has strong potential for further refinement. The model takes three inputs – slope, velocity, and acceleration – to predict the net power (P_{Net}). We selected these inputs based on the Pearson Correlation Coefficient (PCC) analysis, as illustrated in Table III.

TABLE II
VARIABLE DESCRIPTIONS

Variable	**Description**
m	Mass
g	Gravity
c_{rr}	Friction Coefficient
v	Velocity
c_d	Drag Coefficient
A	Frontal Area
ρ	Air Density
a	Acceleration
α	Transmission Efficiency
β	Regeneration Efficiency
δ_t	Sampling Time

Building upon the physics-based model, the hybrid model leverages the same four-layer ANN architecture. However, instead of raw data, it utilizes the calculated forces (rolling



(a) Physics-based (b) ANN (c) Hybrid

Fig. 2. Hexagonal binned plots of Power Forecasting Comparison

TABLE III
PCC OF ANN MODEL FEATURES

	Velocity	Slope	Acceleration
Net Power	0.28	0.19	0.62

resistance, grade, aerodynamic, and acceleration) from equations (4) to (7) multiplied by the velocity as inputs. This refined input selection empowers the model to capture the relationships between these variables more effectively, leading to faster convergence. This is likely reflected in the improved PCC values observed in Table IV.

TABLE IV
PCC OF HYBRID MODEL FEATURES

	Roll	Aero	Climb	Acc
Net Power	0.28	0.26	0.35	0.74

2) Evaluation: To ensure a fair comparison between the models, we employed a stratified split on our data-set, dividing it into three partitions: training (60.3%), validation (29.7%), and test (10.0%). While the training and validation sets were primarily used to train and tune the AI-based and hybrid models, the test set was used for a global comparison of all three approaches (physics-based, AI-based, and hybrid). The effectiveness of the physics-based model can be significantly impacted by the chosen values for transmission and regeneration efficiency. To ensure a more realistic evaluation, we employed an exhaustive search to identify the optimal efficiency values. Since our primary focus is the comparison of power models, we leveraged the ground truth velocity data (actual measured velocity) for this evaluation. As shown in Fig. 2, the physics-based model exhibits a higher degree of variance in its predictions compared to the ANN and hybrid models. This is reflected in a wider distribution of results around the optimal solution, suggesting a greater presence of outliers. Conversely, the ANN and hybrid models demonstrate a higher concentration of predictions near the optimal solution, indicating a lower incidence of outliers.

The results presented in Table V reveal statistically significant improvements in power prediction achieved by both the ANN and hybrid models compared to the physics-based model, as evidenced by superior R-squared (R^2) and a lower Mean Squared Error (MSE) values. While the incremental

TABLE V
TRACTION POWER MODELS EVALUATION

	Physics-based	ANN	Hybrid
MSE	61.035	30.88	29.63
R-squared	0.679	0.779	0.781

improvement offered by the hybrid model relative to the ANN model might appear limited, it's pertinent to consider the cumulative effect of errors in energy estimation. Given that energy is obtained by integrating power over time, even a minor discrepancy in power prediction can significantly amplify the error in the final energy estimates. We conclude that the hybrid model has better performance. Therefore, it will be used for the remainder of this work.

D. Auxiliary Power Demand

We now evaluate the effectiveness of various modeling approaches for the auxiliary power model through a comparative analysis of their results.

1) Modeling: To develop a robust model for forecasting auxiliary power consumption, we embraced a holistic methodology. We combined the power demand from various sources, including air conditioners and other equipment, to obtain a comprehensive data-set of continuous auxiliary power consumption levels. Through this process, we unveiled 45 distinct values, encapsulating the spectrum of observed auxiliary power consumption. Our model utilizes a Random Forest Regressor. This choice is particularly suitable because it can effectively handle both continuous and categorical features. Our data includes several categorical features, such as month and binary indicators for air conditioner and light usage. The Random Forest Regressor will handle these features internally during the model building process. The model leverages five key inputs—ambient temperature (continuous), time of day (continuous), month (categorical), and binary indicators for air conditioner and light usage (categorical)—to predict the specific auxiliary power consumption value based on these influencing factors.

2) Evaluation: The model's performance is summarized in Table VI. It achieves a low Mean Absolute Percentage Error (MAPE), indicating high accuracy in predicting auxiliary power consumption values. Quantitatively, a lower MAPE signifies a smaller average difference between predicted and actual consumption, expressed as a percentage of the actual values. Additionally, the model obtained a good R-squared (R^2) score, suggesting a strong positive linear relationship between the predicted and actual values. This observation suggests that the model's predictions closely align with actual consumption patterns, surpassing a baseline approach that relies solely on the constant average auxiliary consumption.

TABLE VI
AUXILIARY POWER MODEL EVALUATION

	MSE	MAPE	R-squared
Random Forest Regressor	0.049	15.2%	0.808
Constant average Aux consumption	0.270	135.17%	-0.040

E. Long-Term Velocity Forecasting

This subsection investigates the efficacy of various modeling approaches for the velocity model by comparing their results.

1) Modeling: Velocity and acceleration are important variables for the determination of power consumption. To model the velocity we implemented three models and compared their performance. Our *first approach* leverages a multivariate time series model. This acknowledges that the velocity of a segment depends on the previous segment's velocity, along with congestion level and the global average velocity specific to that segment. We opted for a Long Short-Term Memory (LSTM) network with three LSTM cells followed by dense layers. The model is trained to predict the velocity for each segment sequentially, using the actual historical velocity of the previous segment, the average segment velocity and traffic level as inputs. The specific model architecture and hyper-parameters were chosen through experimentation. The *second approach* builds upon the first by employing a recursive strategy. We maintain the same LSTM network architecture. However, during training, instead of using the actual previous velocity of the segment as a training input for predicting the next one, we use the predicted velocity from the previous prediction. This approach aims to help the model learn the inherent error rate associated with its predictions and mitigate its influence on subsequent predictions. The *third approach* utilizes a seven-layer ANN to directly predict future velocity, bypassing the need for previous velocity data. This approach offers advantages in simplicity and computational efficiency.

In all three approaches, a safety check is applied using the maximum allowable velocity of each segment. This ensures that even if the predicted speed exceeds the limit, the final output is adjusted to stay within the safe range.

2) Evaluation: We evaluated the models on the same test set described earlier. As shown in Fig. 3, the ANN model achieves superior performance compared to both the other approaches and the baseline references. Notably, the ANN boasts significantly faster inference times compared to LSTM. This is because, in our case, LSTM requires each inference to consider the previous one, introducing a sequential dependency. While the results were promising, the forecasting method does not account for the sequential nature of the segments. Consequently, we observed abrupt changes in speed between segments. This can lead to significant acceleration and deceleration events, which translates to power spikes that negatively impact the overall power estimation accuracy. This phenomenon is clearly visible in Fig. 4 as we can see that the spectral components magnitude of the predicted acceleration is higher than the actual. To enhance the realism of the forecasted velocity, we employed the Exponential Moving Average (EMA) technique on both of the velocity and acceleration. This approach functions as a low-pass filter, attenuating the magnitude of high-frequency components in the data, leading to a smoother speed transition between road segments.

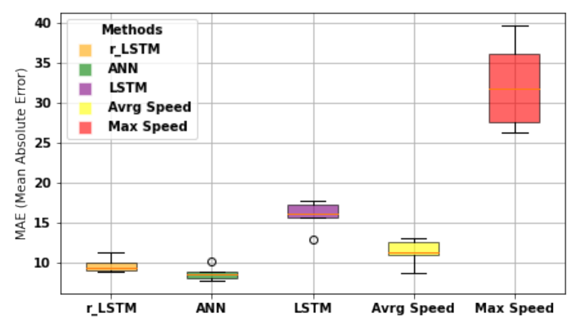


Fig. 3. Box plot of different velocity forecasting methods

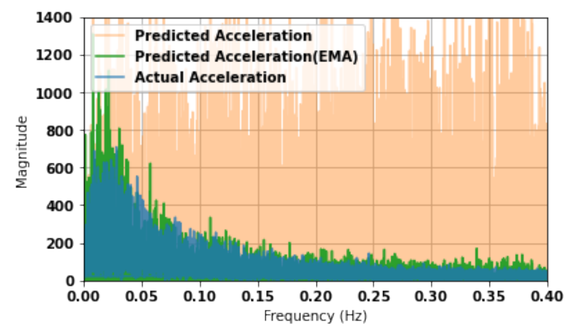


Fig. 4. Acceleration spectrum

IV. EVALUATION OF THE INTEGRATED MODEL

We conducted a comprehensive evaluation of the integrated model on the same test set described earlier. This test set comprised eight driving cycles, each lasting one hour. The evaluation assessed the performance of the integrated model, which incorporated an ANN velocity forecasting module alongside the hybrid traction power and Auxiliary power demand models. Energy consumption was calculated using Equations 11 and 10. By subtracting this estimated energy consumption from the initial battery capacity, we can obtain an approximate value for the remaining charge. Figure 5 depicts the results for a single driving cycle. As observed, the calculated SoC falls below the actual value. This discrepancy can be attributed to the inherent error rate of the power model.

Additionally, it is noteworthy that the Estimated Time of Arrival (ETA) is two minutes less than the Actual Time of Arrival (ATA) due to the model's velocity predictions. This difference in arrival times led to higher power consumption. Table VII details the performance of the integrated model across various driving cycles. Notably, all cycles share an ATA of 60 minutes and an average driven distance of 50 kilometers.

TABLE VII
PERFORMANCE ON DIFFERENT DRIVING CYCLES

Driving Cycle	Energy MAE	SoC MAE	ETA (Minutes)
1	0.579	2.734	58.07
2	1.026	4.490	62.22
3	0.628	2.702	76.02
4	0.555	2.834	59.76
5	0.601	2.568	61.28
6	0.204	0.970	63.63
7	0.307	1.558	72.24
8	0.311	1.253	67.79

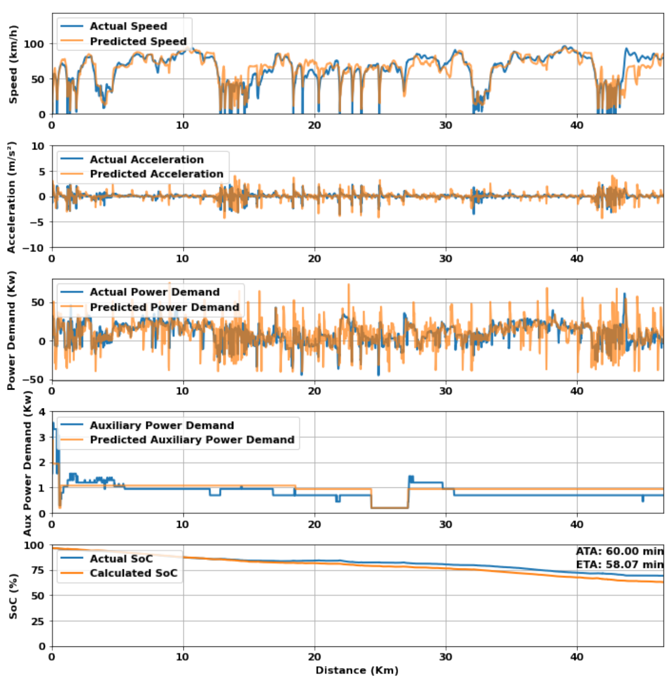


Fig. 5. Global evaluation on driving cycle 1

V. Conclusion

In conclusion, this study emphasizes the potential of a Data-Driven integrated model for both SoC estimation and energy consumption prediction in connected EVs. The model incorporates both velocity forecasting and power demand estimation. Furthermore, this model can be seamlessly integrated with route planning tools leveraging road data from various APIs. Its adaptability to different vehicles and drivers based on data availability enhances its versatility. Notably, this approach excels in in-vehicle navigation systems, as the onboard computer has continuous access to all necessary data. This data can be used to continuously train the model, ensuring it remains updated and captures changes in driving and consumption patterns. Overall, connected EVs, with their data-driven approach, represent a leap forward in sustainable transportation.

Acknowledgment

This paper is supported by the OPEVA project that has received funding within the Chips Joint Undertaking (Chips JU) from the European Union's Horizon Europe Program and the National Authorities (France, Czechia, Italy, Portugal, Turkey, Switzerland), under grant agreement 101097267. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Chips JU. Neither the European Union nor the granting authority can be held responsible for them.

References

[1] Y. An, Y. Gao, N. Wu, J. Zhu, H. Li, and J. Yang, "Optimal scheduling of electric vehicle charging operations considering real-time traffic condition and travel distance," *Expert Systems with Applications*, vol. 213, p. 118941, 2023.

[2] M. Baum, J. Dibbelt, A. Gemsa, and D. Wagner, "Towards route planning algorithms for electric vehicles with realistic constraints," *Computer Science-Research and Development*, vol. 31, pp. 105–109, 2016.

[3] A. L. Rhun, F. Bonnans, G. de Nunzio, T. Leroy, and P. Martinon, "An eco-routing algorithm for hevs under traffic conditions," vol. 53. Elsevier B.V., 2020, pp. 14 242–14 247.

[4] F. Morlock, B. Rolle, M. Bauer, and O. Sawodny, "Time optimal routing of electric vehicles under consideration of available charging infrastructure and a detailed consumption model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, pp. 5123–5135, 12 2020.

[5] A. M. Bozorgi, M. Farasat, and A. Mahmoud, "A time and energy efficient routing algorithm for electric vehicles based on historical driving data," *IEEE Transactions on Intelligent Vehicles*, vol. 2, pp. 308–320, 12 2017.

[6] T. Perger and H. Auer, "Energy efficient route planning for electric vehicles with special consideration of the topography and battery lifetime," *Energy Efficiency*, vol. 13, pp. 1705–1726, 12 2020.

[7] L. Thibault, G. D. Nunzio, and A. Sciarretta, "A unified approach for electric vehicles range maximization via eco-routing, eco-driving, and energy consumption prediction," *IEEE Transactions on Intelligent Vehicles*, vol. 3, pp. 463–475, 12 2018.

[8] G. De Nunzio and L. Thibault, "Energy-optimal driving range prediction for electric vehicles," in *2017 ieee intelligent vehicles symposium (iv)*. IEEE, 2017, pp. 1608–1613.

[9] J. Li, S. Tian, N. Zhang, G. Liu, Z. Wu, and W. Li, "Optimization strategy for electric vehicle routing under traffic impedance guidance," *Applied Sciences*, vol. 13, no. 20, p. 11474, 2023.

[10] M. Dorokhova, C. Ballif, and N. Wyrsch, "Routing of electric vehicles with intermediary charging stations: A reinforcement learning approach," *Frontiers in Big Data*, vol. 4, 5 2021.

[11] A. I. Croce, G. Musolino, C. Rindone, and A. Vitetta, "Energy consumption of electric vehicles: Models' estimation using big data (fcd)," vol. 47. Elsevier B.V., 2020, pp. 211–218.

[12] W. Achariyaviriya, W. Wongsapai, K. Janpoom, T. Katongtung, Y. Mona, N. Tippayawong, and P. Suttakul, "Estimating energy consumption of battery electric vehicles using vehicle sensor data and machine learning approaches," *Energies*, vol. 16, no. 17, 2023.

[13] Z. Zhou, Z. Yang, Y. Zhang, Y. Huang, H. Chen, and Z. Yu, "iscience a comprehensive study of speed prediction in transportation system: From vehicle to traffic."

[14] G. Wu, F. Ye, P. Hao, D. Esaid, K. Boriboonsomsin, and M. J. Barth, "Deep learning–based eco-driving system for battery electric vehicles[dataset]," 2019.

[15] F. He, X. Yan, Y. Liu, and L. Ma, "A traffic congestion assessment method for urban road networks based on speed performance index," *Procedia Engineering*, vol. 137, pp. 425–433, 2016, green Intelligent Transportation System and Safety.