vehicles

MDPI

Article

ML-Based Control Strategy for PHEV Under Predictive Vehicle Usage Behaviour

Aleksandr Doikin [1], Aleksandr Korsunovs [1], Felician Campean [1], Oscar García-Afonso [2,*] and Enrico Agostinelli [3]

[1] Automotive Research Centre, University of Bradford, Bradford BD7 1DP, UK; a.doikin2@bradford.ac.uk (A.D.); a.korsunovs2@bradford.ac.uk (A.K.); f.campean@bradford.ac.uk (F.C.)
[2] Departamento de Ingeniería Industrial, Escuela Superior de Ingeniería y Tecnología (ESIT), Universidad de La Laguna (ULL), Camino San Francisco de Paula, n° 19, 38200 San Cristóbal de La Laguna, Spain
[3] Jaguar Land Rover, Coventry CV34LF, UK; eagostin@jaguarlandrover.com
* Correspondence: ogarciaa@ull.es

Abstract: This paper introduces a novel strategy for an intelligent plug-in hybrid electric vehicle (PHEV) energy optimization strategy based on machine learning (ML) prediction of the upcoming journey, without recourse to navigation or other external data, which underpins many of the existing approaches. This study, based on extended real-world data (journeys history from 10 vehicles over 12 months), shows that trip patterns can be learnt quite effectively using classic ML classification algorithms. In particular, the RusBoosted ensemble classifier performed consistently well across the heterogeneous dataset (volume of data for training and variable imbalance in the datasets, reflecting the natural variability in the vehicle usage profiles), providing sufficiently accurate predictions for the proposed EMS strategy. Performance evaluation experiments were carried out using a model-in-the-loop (MIL) simulation set-up developed in this research. The results demonstrated that the proposed strategy has the potential to deliver significant reductions in engine running time (up to 76% on routine short journeys), with associated benefits in CO_2 consumption and tailpipe emissions, as well as enhanced engine reliability. The broader importance of this study is that it demonstrates the great potential of using predictive insights from computation-efficient and robust ML to learn vehicle usage patterns to optimize the control strategies without reliance on uncertain external inputs.

Keywords: vehicle electrification; plug-in hybrid vehicles; intelligent control strategy; machine learning

check for updates

Academic Editors: Osama A. Mohammed and Ahmed Mohamed

Received: 17 December 2024
Revised: 24 January 2025
Accepted: 21 February 2025
Published: 25 February 2025

Citation: Doikin, A.; Korsunovs, A.; Campean, F.; García-Afonso, O.; Agostinelli, E. ML-Based Control Strategy for PHEV Under Predictive Vehicle Usage Behaviour. *Vehicles* **2025**, *7*, 23. https://doi.org/10.3390/vehicles7010023

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Passenger cars and light commercial vehicles make up a significant portion of the on-road traffic, with a propulsion mix including traditional internal combustion engines (ICE), battery electric vehicles (BEV), and hybrid electric vehicles (HEV). In pursuit of reduced greenhouse gasses and harmful air pollutant emissions, governments are directing a significant shift towards electrification of the automotive sector, with clearly defined plans to phase out ICE-powered vehicles [1] and a roadmap for the development of supporting infrastructure [2]. However, given some of the known BEV adoption challenges, including range anxiety, long charging times and insufficiently developed infrastructure for charging and energy provision [3,4], it is not feasible to discontinue the use of ICE in the short to medium term.

In this context, HEVs offer a compromise solution. In particular, PHEV are an appealing solution, as they allow users to charge their electric batteries from the grid in addition to recovering energy on board. Modern PHEV passenger vehicles have relatively large battery capacities, with some models allowing up to a 70 mile range in the all-electric mode. This range is sufficient to cover most routine urban driving journeys, while still affording longer range trips with ICE propulsion.

While completing full journeys in EV mode is desirable for a PHEV, in practice, the blending of electric and ICE propulsion is more complex. For example, it is common that PHEVs use ICE power, e.g., to deliver higher torque demands from the driver. However, "cold" ICE engine starts can be problematic [5] and contribute to tailpipe emissions, while frequent ICE engine starts, in addition to contributions to fuel usage and emissions, are also associated with contributions to oil dilution and shorter service intervals and potential durability concerns [6].

Efforts to optimize the PHEV energy management strategies (EMSs) have been reported in the literature. For example, ref. [7] have presented a comprehensive approach to predictive energy optimization (PEO) for a PHEV based on a dynamic programming optimization, underpinned by live route navigation and traffic information. The review of the EMSs in [8] provided a comprehensive overview of machine learning and AI methods for HEV and PHEV optimization, and also the connectivity assumption in satellite navigation.

However, one practical limitation of these approaches is that drivers do not always connect to navigation services for routine repeated journeys. Such journeys are known to represent the majority of journeys made by the general public that are linked with daily routines such as commuting to work or to school. In terms of mileage or duration routine, journeys could be either short, long, or a mixture of both, as illustrated by the sample of real-world journey data for one vehicle shown in Figure 1. If navigation data are not available for routine journeys, which accounts the vast majority of vehicle usage for some users, the PEO strategy cannot be employed, and the base EMS is used instead. This not only results in suboptimal energy usage, but also durability concerns from frequent engine starts, in particular in conjunction with repeated short journeys.

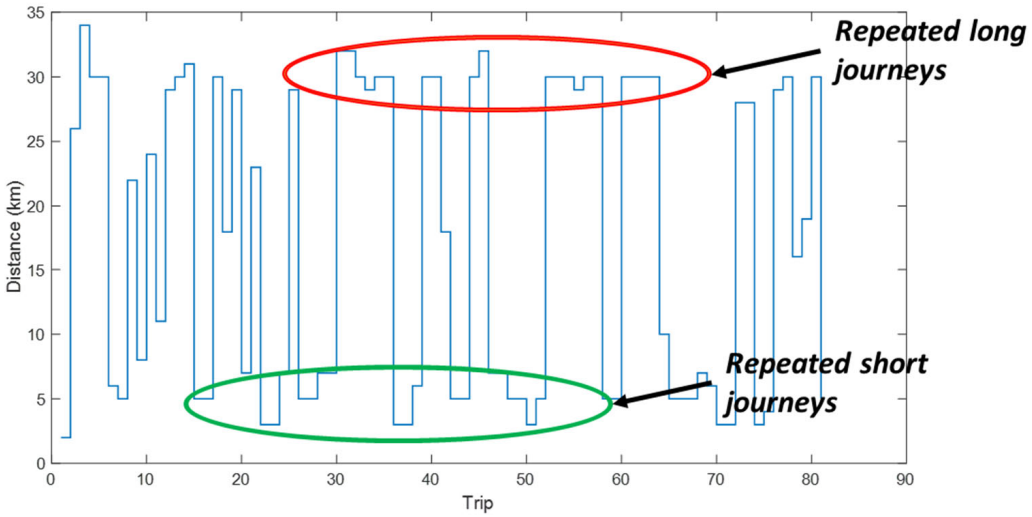


Figure 1. An example of a real-world vehicle journey pattern.

The research hypothesis underpinning the work presented in this paper is that machine learning can be deployed to data routinely collected from vehicles, through telematics or data-over-the-air methods, to "learn" the patterns of vehicle usage in terms of journeys and driver behaviour. If a journey can be robustly predicted based on the learnt pattern and current data, then an intelligent PHEV EMS can be designed to use this information to

maximize the use of the electric mode and minimize the ICE emissions. A case of interest is the prediction of shorter journeys that can be completed in electric mode only, subject to the battery's state of charge. For such cases, an "ECO" EMS option can be implemented, e.g., it is possible to force the EV mode, delivering a significant benefit in terms of both emissions and engine durability or service interval, provided that one or more cold starts are avoided within the journey.

This strategy is significantly novel compared to current PEO strategies used for PHEVs that are dependent on navigation information or other external information, and has potential to deliver significant benefits for common real-world vehicle usage patterns (as illustrated in Figure 1, where even the "long" journeys are shorter than the EV range on a fully charged battery). Therefore, the aim of the research is to evaluate the potential of this approach based on real-world journeys and drive cycle data, in conjunction with a model-in-the-loop simulation.

The challenges addressed in this study relate to three aspects: (i) the selection of features and ML algorithms that can be employed to predict the upcoming journey, with sufficient accuracy and robustness to be used for the proposed PHEV EMS; (ii) the development and validation of a PHEV simulation model, that can be employed for the model-in-the-loop (MIL) experiments; and (iii) setting up and conducting the MIL experiments to evaluate the performance of the proposed PHEV EMS comparatively against the standard PHEV EMS, with drive-cycle data reflecting the real-world usage of the vehicle.

This paper describes in detail the study of ML modelling based on real-world datasets of vehicle usage information (journeys and driver behaviour) and proposes an intelligent PHEV "ECO" EMS based on the predicted energy required for the upcoming trip and the current state of the vehicle. This paper also presents the development and validation of a detailed simulation model of a PHEV used for the evaluation of the PHEV control strategy in a virtual MIL environment. The detailed evaluation of the performance of the proposed strategy assessed based on MIL experiments considers criteria including the impact on engine running time, fuel consumption/CO_2 reduction benefits, and the vehicle's ability to meet the performance demands of the driver.

The paper is organized as follows: Section 1.1 of the article provides an overview of existing approaches towards PHEV EMSs, detailing their benefits, drawbacks, and opportunities for improvement. Section 2 outlines the methodology applied in this research; it includes the description of real-world test data used, the development of the machine learning-assisted vehicle usage model and its use for the PHEV EMS developed in this research, as well as the development and validation of the PHEV model within the MATLAB Simulink (2021a, 9.10) environment. Section 3 details the simulation results of the PHEV model with the proposed ML-assisted control strategy, which is also compared against the conventional PHEV EMS performance in terms of overall engine running time, CO_2 emissions, and performance capability. Section 4 provides a summary of outcome results, the conclusion, and the contributions involved.

1.1. Review of Related Work

1.1.1. Overview of PHEV EMSs

In specialist literature, the electric-only mode is usually referred to as the all-electric charge-depleting (CD) mode. In this mode, the propulsion power is solely provided by the electric motor until the battery state-of-charge (SOC) reaches a minimum allowable level or the required power exceeds the maximum electric hardware output (high speeds or aggressive accelerations). Once the battery level depletes to the critical threshold, the charge-sustaining (CS) mode is engaged, employing an ICE to maintain battery charge at a set minimum threshold. This leads to the loss of potential PHEV benefits, as in this

case the vehicle is essentially used as mainly ICE-powered full/mild hybrid, yet with increased weight due to relatively big electric battery and possibly suboptimal ICE [9,10]. All-electric CD-CS EMSs offer ease of implementation and are attractive due to their simplicity. However, their efficiency is limited [11]. Sun et al. [12] addressed the lack of optimality for CD-CS EMSs in simulation environments, where they demonstrated 90% fuel optimality in comparison to the 94–96% fuel optimality of the EMSs that were optimized for approaching traffic information, with 100% being the overall best result achieved through a dynamic programming approach. This was set as the benchmark.

An alternative EMS approach that can be used for PHEVs is Blended Mode (BM), which utilizes both ICE and electric motors throughout the trip [13]. Mechichi et al. [14] proposed a strategy that merges the use of CD, BM and CS mode selection based on the current SOC, which also allows a better fuel economy to be attained compared to the use of a single mode. In essence, BM EMSs are a combination of different operating modes, and they are dependent on the availability of upcoming trip-related information. Without a trip preview, the BM can be less efficient than well-tuned CD-CS EMSs [11]. Therefore, the main challenge associated with using BM strategies is trip information handling within the strategy, as well as the processing and control implementation, which can be computationally expensive.

Many researchers have attempted to deal with this challenge. For example, ref. [15] presented an efficient following-vehicle PHEV BM EMS, relying on trip data obtained from a lead vehicle. This, of course, requires interconnectivity between vehicles and therefore can only be used in a narrow field of possible applications, such as in EMSs for platooning freight trucks on motorways. However, this can also be useful for autonomous driving scenarios as discussed by [16].

Scholars have discussed scenarios, including a specific drive cycle/route, where it is possible to optimize the vehicle EMS to maximize efficiency. This could be highly effective for a PHEV city bus that follows the same route every time; thus, the use of an optimized BM strategy can lead to significant improvement in fuel economy and drivability [17,18]. However, most vehicles do not follow just one pre-defined route and therefore this approach has limited applicability.

Several researchers sought to use machine learning (ML) techniques to match the current trip conditions to sections of stored drive cycle information, for which optimal EMS parameters are known. For example, ref. [19] used artificial neural networks for drive cycle recognition (DCR) based on different operating parameters, such as maximum velocity, maximum acceleration, maximum deceleration, average driving velocity, and the driving distance. The output of this DCR was used as the reference input for the fuzzy energy management strategy (F-EMS) of the powertrain. Through simulation, this strategy showed a 21.39% reduction in fuel consumption and reductions of similar magnitudes in harmful emissions in comparison to the stock EMS during standard emissions drive cycles, whilst still using similar amounts of battery energy. Similar studies were presented in [20,21], showing similar benefits. In [22], a strategy was proposed that employed a backpropagation neural network to predict vehicle future velocity for the next 5 s based on the past 10 s of vehicle velocity data. This information was further used by the EMS to solve the power source torque distribution problem. Simulation over 6 consecutive China light-duty vehicle test cycles for passenger cars showed a fuel consumption reduction of 9.85% in comparison to the CD-CS strategy. However, within the model, at the start of the trip, the authors required navigation data to estimate the expected driving mileage for SOC consumption evaluation purposes, along with historical commuting conditions information as the reference velocity input of the driver to calculate torque demand. A similar approach was used by [23] to predict the near-future drive cycle, employing previ-

ously developed Economy Driving Pro System along with global driving cycle construction methodology [24]. Climent et al. [25] proposed a route-optimized BM EMS, which was also supported by ML and based on the Equivalent Consumption Minimization Strategy, that exploits driving history (represented by driving velocity and GPS position) to determine the driving conditions of sections of the cycle used in the current trip.

There have also been attempts to use neural networks as EMSs. For example, ref. [26] used a Deep-Q Network reinforcement learning algorithm to control ICE power output based on power demand at the wheel, the battery SOC, and the distance to the destination. This approach showed a 16.3% reduction in fuel consumption in comparison to the CD-CS mode. Several studies, e.g., [22,27], also considered battery ageing as an objective for EMS optimization, but no ML studies considered the impact on ICE engine durability.

In summary, while significant advancements have been made in developing intelligent PEO strategies for BM EMSs, their predictive horizon is limited to a route segment and/or also requires an external input, such as information from a lead vehicle, the GPS position, or traffic data information, thus requiring robust vehicle connectivity. Moreover, the implementation of optimization routines requires significant data processing and data storage capabilities that are frequently out of the scope of on-board offline applications. A recent comprehensive review of energy management strategies for HEVs, presented by Mohseni et al. [28], emphasized the challenges with the computational load and the issues associated with uncertainties and robustness, linked to both internal and external factors.

1.1.2. Review of Data-Driven Vehicle Usage Modelling

The growing amount of data available from the real-world usage of vehicles has provided opportunities to better understand vehicle usage profiles using data-driven modelling approaches. A wide range of machine learning techniques are available, ranging from classical statistical inference methods [29,30] to classification machine learning [31], including both supervised and unsupervised techniques [32,33], and, increasingly, deep learning methods. Taye [34] provided a detailed comparison between deep learning and classical machine learning methods, as summarized in Table 1.

Table 1. Classical machine learning vs. deep learning summary (based on Taye [34]).

Aspect	Deep Learning	Classical Machine Learning
Dataset size	Large datasets required	Small/medium datasets
Data type	Unstructured (images, text)	Structured (tabular data)
Feature engineering	Automatic	Manual
Computational effort	High (GPU/TPU)	Low to moderate
Interpretability	Low	High
Training time	Longer	Faster

A comprehensive review of data-driven modelling approaches for vehicle usage modelling purposes [35] showed that the choice of algorithm depends on multiple factors, with no immediate universal recommendation.

For example, ref. [36] applied a support vector machine to model driver identification based on a combination of information coming from the acceleration and brake pedals, demonstrating an accuracy of 95% based on data from a controlled experiment. Wang et al. [37] also looked at driver identification but used maneuver analysis in conjunction with neural networks. Zhang et al. [38] considered individual driving pattern analysis for the benefit of ECO-mode improvement. There are also several frameworks allowing

model drivers and replicate their behaviour, starting from just simulating behaviour at road intersections [39] to mimicking human behaviour in autonomous driving [40].

Predicting future driving behaviour patterns is a much more challenging task due to the higher degrees of uncertainty. Short-term prediction was attempted by [41], who used real-world data to predict the next driving action, e.g., steering or braking. In this work, a deep recurrent neural network was used to predict the next driver action to improve performance of advanced driver assistance systems (ADASs). The reported validation experiment claimed robust ability to make predictions 5 s ahead of driver's action. Kolachalama and Malik [42] have applied the deep learning NARX method for a predictive model using real-world field data, which forecasted engine operating point mapping. The prediction was then performed at an adaptive cruise control speed to enhance engine operating conditions.

A journey-based ML-based vehicle usage behaviour modelling methodology was introduced in [43], aiming to predict the characteristics of the upcoming journey based on learning from historic journeys, given a set of pre-defined classes of interest. This approach has the advantage of relatively low computational complexity as the prediction is based on ML models fitted offline. It also does not require online journey information (like GPS journey data) to optimize the control strategy. A diesel emissions aftertreatment system was employed to demonstrate practical application to particulate filter regeneration timing optimization, where the control strategy, adapted based on journey duration predictions, was proven to successfully minimize the number of interrupted regenerations. This in turn reduced the overall time spent in the regeneration mode, which led to increased fuel efficiency and reduced CO_2 emissions.

2. Materials and Methods

2.1. Proposed Intelligent PHEV EMS and Research Methodology

This research introduces a novel intelligent PHEV EMS, underpinned by predictive vehicle usage, for upcoming trip energy requirements. It is derived from the ML modelling of the pattern data of past trips. The principle of the proposed strategy is illustrated in Figure 2. The offline data-driven modelling module used historic trip data on journeys and driver behaviour to develop a predictive vehicle usage model. In the online module, considering the current journey and vehicle data, a prediction is made for the current trip. On this basis, an evaluation is made of the energy required to complete the journey. If the trip can be completed in the EV mode only, then, the proposed EMS will maintain EV mode only for the whole trip; otherwise, the base mixed-mode EMS is employed. If the strategy is successful, the expectation is that a higher EV mode utilization and a significant reduction in engine cold starts will be observed compared to the based strategy.

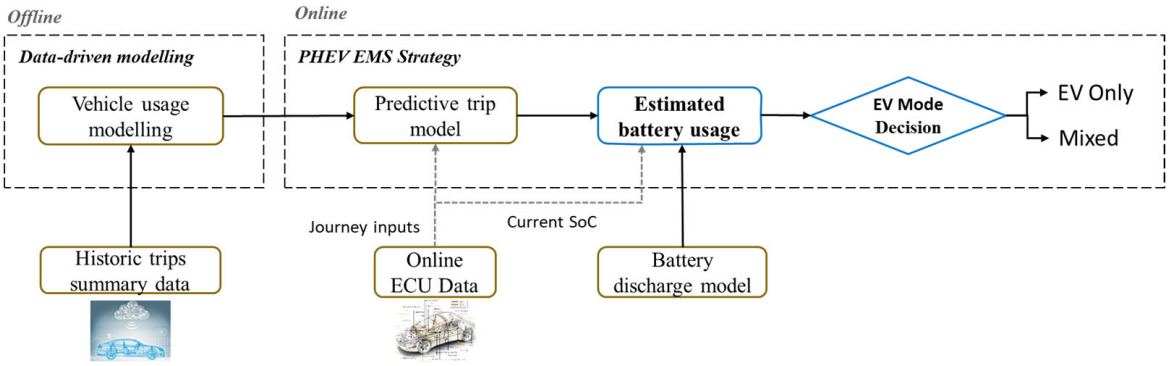


Figure 2. Proposed intelligent PHEV EMS based on ML-assisted vehicle usage model (ML-VUM).

This research aims to evaluate the feasibility and performance of this proposed strategy against these expectations. In order to achieve this aim, the following research questions need to be addressed:

1. Data-driven modelling of vehicle usage patterns: which journey features and ML models can be employed to achieve a robust predictive model for vehicle journeys?
2. Model-in-the-loop (MIL) simulation set-up: what PHEV simulation can be implemented to support the validation of the proposed EMS on an MIL set-up?
3. PHEV EMS validation: how can drive cycle data-driven experiments reflecting real-world usage be planned and run on the MIL to evaluate the performance of proposed PHEV EMSs against the standard mixed-mode EMSs?

2.2. Data-Driven Vehicle Usage Modelling

2.2.1. Data for Vehicle Usage Model

A set of historic real-world user trip summary data from 10 PHEV SUVs (sport utility vehicles) of a particular model, followed up over a period of 12 months, was made available for this study. The data, summarized in Table 2, consisted of summary statistics for each journey for a limited number of parameters: trip details, including the calendar timestamps for the start and end of the journey, the average vehicle speed, the trip duration, and the distance travelled. Table 2 also provides the average distance for the journeys undertaken by each vehicle in the sample.

Table 2. Summary of data available for vehicle usage model.

Vehicle Number	Total Trips	Average Trip Distance (km)
Vehicle 1	621	9
Vehicle 2	371	24
Vehicle 3	634	7.7
Vehicle 4	363	11.4
Vehicle 5	594	8.4
Vehicle 6	771	11.2
Vehicle 7	917	11.8
Vehicle 8	373	17.7
Vehicle 9	610	8.8
Vehicle 10	300	7.3

Detailed (high-frequency) data from the vehicle propulsion system were not available for this study. Such data are needed to calculate detailed energy demand or consumption profiles for the sections of the trip, which commonly inform PEO strategies. While high-frequency data were used in many of the PEO studies discussed in the literature, none of these studies considered such a large volume of real-world journeys data. Storing such high-resolution data for each trip and driver profile is problematic, for reasons including confidentiality as well as data storage costs. Therefore, this option is not generally available for open research. However, based on the data available, a simple overall energy consumption estimate per journey can be inferred based on the average velocity and duration of the trip using virtual vehicle models.

2.2.2. ML Model for Predicting Upcoming Trip Distance

The proposed intelligent PHEV EMS, shown in Figure 3, requires a predictive trip model for the current/upcoming journey (at the start of the journey) to estimate the energy required for completing the trip. The approach is predicated on the assumption that many users undertake routine vehicle usage patterns (commuting to work, school drop-off, also, completed without the use of navigation for route planning) with relatively short-distance journeys that could potentially be completed in the EV mode, depending on the battery SOC.

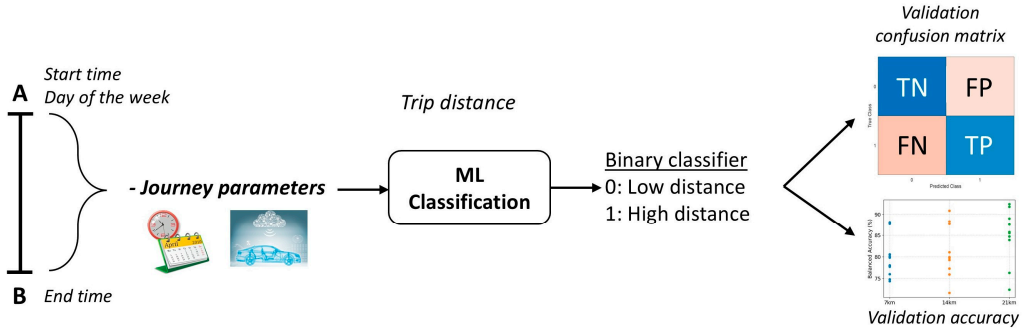


Figure 3. Trip distance modelling using classification algorithms.

Given the available parameters in the dataset, the trip distance could be set as the prediction variable as a function of the other variables. Looking at this as a classification ML problem, the first step is to define the basic classes of trips, based on the actual trip distance (Y). The natural approach is to classify trips into "short" (as trips that could be safely completed in EV mode) and "long", i.e., exceeding an arbitrarily defined limit value for the trip distance X_{lim}, in kilometres. Thus, a binary transformation can be applied to the distance travelled variable, as shown in Equation (1) below:

$$Y_i' = \begin{cases} 0, \; if \; Y_i < X_{lim} \; kilometres \\ 1, \; if \; Y_i \geq X_{lim} \; kilometers \end{cases} \quad (1)$$

Figure 3 illustrates the process of ML modelling implemented. The ML modelling was carried out within the MATLAB (2021a, 9.10) Classification Learner [44] app. The training of the model is based on the data available on historic journeys for each vehicle, using only the basic features of the journey data (start time, day of the week and vehicle park time, defined as the time interval from the end of the previous trip), with the labelled trip distance (obtained from the binary class transformation in Equation (1)) as the prediction variable of interest. The choice to restrict the predictor features to the basic journey information was based on the judgement that the PHEV EMS will require a prediction immediately after the key-on event for the current/upcoming journey, at which point only the calendar trip parameters will be available.

Using this architecture, vehicle usage models were fitted to each of the 10 vehicles with a known history of journeys, as summarized in Table 2. For each vehicle, three separate models (as each model is able to predict one particular distance) were trained to predict (on the basis of different pre-defined classes) distances of interest; in this case, we chosen distances of 7, 14, and 21 km. These distances were selected based on discussion with industry experts as the most representative based on the distribution analysis of vehicle trips in relation to the PHEV EV mode range. Given the size of the data and ECU requirements for fast computation time, preference has been given to the classical machine learning methods over the deep learning.

The vehicle usage data were partitioned as follows: the data corresponding to the first 9 months were used for model training, while the data for the remaining 3 months were reserved for model validation. Seven ML modelling algorithms (listed in Table 2), selected as the most used for classification problems [45], were considered for comparative performance analysis using the datasets available. Each classification ML algorithm was trained for each vehicle data using the training dataset (i.e., the first 75% of the records for each vehicle), and each of the three distance thresholds (7, 14 and 21 km). Thus, in total, 30 models were fitted for each algorithm. Each model was then validated using the validation dataset—i.e., the 25% of the data records set aside for each vehicle.

Table 3 summarizes the performance analysis of the algorithms as the average validation classification accuracy results across the 10 vehicles and the 3 distances. The performance metrics considered include the proportions of true positives (TPs) and true negatives (TNs) as predictions for the validation set (unseen data), and the balance accuracy calculated as the average of TPs and TNs. The preference for the balanced accuracy is justified by the observation that the data imbalance in vehicle usage can be on both sides (i.e., towards long or short journeys), depending on the individual driver vehicle usage patterns. This choice avoids pitfalls of the traditional accuracy measures for heavily imbalanced datasets, where the F_β metric is common, where β is a parameter that sets the balance of recall (i.e., the proportion of true positives correctly identified) relative to precision (i.e., the proportion of predicted positives that are true positives).

Table 3. Average validation accuracy for ML algorithms on unseen data.

Algorithm Tested	Balanced Accuracy	TP Accuracy	TN Accuracy
Decision tree	71.5%	69%	74%
Logistic regression	64%	59%	69%
Linear discriminant analysis	71%	67%	75%
Support vector machine	68.5%	57%	80%
k-NN classifier	64.5%	63%	66%
Neural network	76.5%	75%	78%
RusBoosted tree ensemble	81%	77%	85%

The analysis in Table 3 shows that the RusBoosted tree algorithm has the best performance, with an average accuracy of 81% on unseen data. This superior performance can be explained by the nature of imbalance present in the dataset, i.e., it is accounted for if the proportion of short trips after the binary transformation is much higher than that of the long ones, or the other way around. The efficiency of this algorithm for the type of data involved in vehicle usage modelling is justified by the boosting technique designed to address the issues with imbalanced datasets, i.e., when the number of records in target classes is significantly different. This is achieved through the conversion of multiple weak learners into a single composite robust classifier [46]. In the literature, it has also been discussed that boosting is a reliable method that generally avoids overfitting [47]. Therefore, RusBoosted tree was selected as the primary algorithm for model vehicle usage modelling.

Figure 4 shows the performance of the RusBoosted tree algorithm across the three classification distances for the set of 10 vehicles in terms of the balanced accuracy of the validation set predictions. While there is variability in the prediction accuracy across the vehicles and the distance threshold for classification, explained by the different resulting patterns of imbalance, in all cases, the models have reasonably good prediction accuracy, being usable from the point of view of the proposed PHEV EMS.

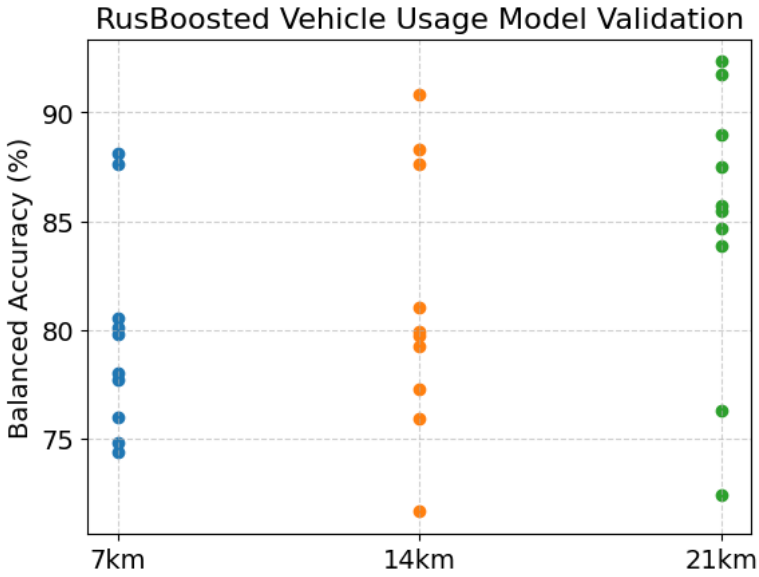


Figure 4. Vehicle usage model validation for 10 vehicles based on balanced accuracy measure—RusBoosted ensemble classifier.

2.3. PHEV Simulation Modelling

2.3.1. PHEV Simulation in Autonomie

The vehicle simulation tool used in this study was Autonomie (2021U1), which was developed by the Argonne National Laboratory. This MATLAB®-based software has been successfully applied in performance, fuel economy, and emission studies for both light- and heavy-duty vehicles [48–51]. A key feature of this framework is the detailed definition of the vehicle control algorithms that have been developed from extensive sets of chassis dynamometer data [52]. This fact, together with several examples of its successful application in hybrid assessments [53], ensures that the EMS is captured well, which becomes key in this research work.

As a starting point, the default Autonomie's library for parallel PHEV models is used. The driver model uses a look-ahead controller to model the accelerator and brake pedals, generating a torque demand that is interpreted by the supervisory controller, which sends commands to the vehicle plant models (engine and motor demand, engine on/off strategy, etc.).

The main hardware components are modified, according to Table 4, to match the details of the sport utility vehicle (SUV) PHEV used in this study (commercial details not provided to avoid mercantilism). Regarding the control side, the vehicle's propulsion control strategy is kept constant, as it is a verified and representative supervisory controller for parallel PHEVs.

Table 4. Main parameters of the simulated vehicle.

Parameter	Value
Frontal area [m^2]	3.1
Drag coefficient	0.34
Wheelbase [m]	2.9
Curb weight [kg]	2471
ICE power [kW]	221
e-motor power [kW]	85
Battery capacity [kWh]	13
Transmission	8-speed automatic
Fuel	Petrol

2.3.2. Data for Vehicle Simulation Model Validation

Running a detailed vehicle simulation model in Autonomie requires a much more detailed set of data, collected at a much higher frequency. As discussed earlier (in Section 2.2.1), such data are not normally/routinely available from data collected from vehicles in the field, unless the vehicle is instrumented specifically for this purpose.

Two sets of detailed drive cycle data were provided for Autonomie model validation. These were collected from instrumented development vehicles based on customer-representative journeys, as illustrated in the top row of Figure 5. The first test represents a low-demand mixed drive cycle (Figure 5a), with both urban and motorway driving conditions; the second test includes a dynamic driving style, with multiple accelerations and decelerations (Figure 5b).

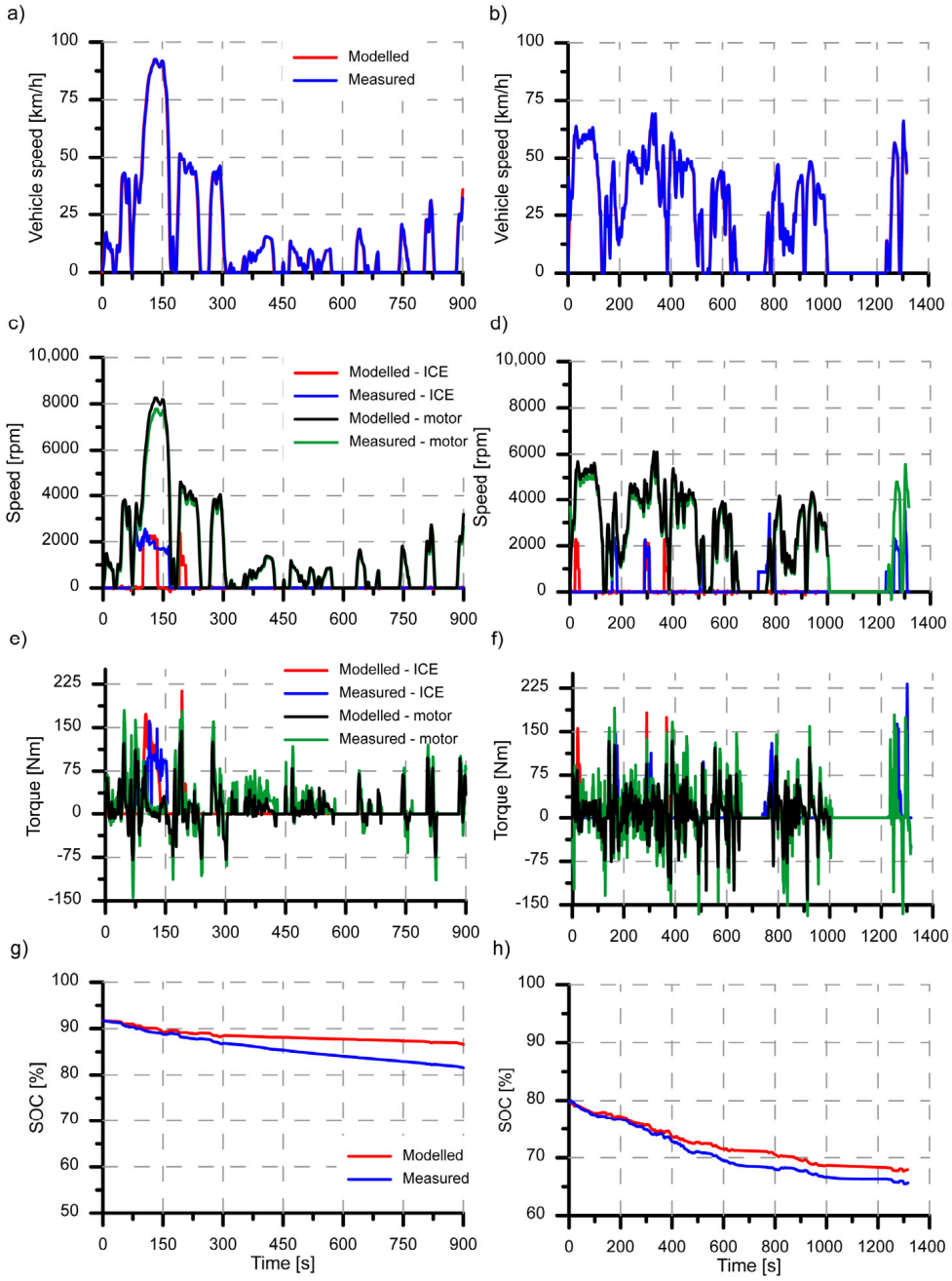


Figure 5. Autonomie PHEV model validation—comparison between modelled results and existing data in two real-world driving validation tests: vehicle speed (**a**,**b**); ICE and e-motor speed (**c**,**d**); ICE and e-motor torque (**e**,**f**); and SOC (**g**,**h**).

2.3.3. Autonomie Model Validation

Once configured, the PHEV model was validated against data from two real-world driving cycles.

The measured speed traces were used as targets in Autonomie without considering the gradient, as the altitude variations in those tests were negligible. For each driving test, the simulated ICE and e-motor operation (speed and torque) and the battery SOC are compared against the experimental data (see Figure 5). A good match between the measured and modelled speed profiles can be observed in Figure 5, with no trace missing, evidence of the accuracy of the driver model. This fact, together with the correct chassis configuration and transmission model, is critical in order to match the torque request. This can be seen in both the ICE and e-motor operation, defined by their corresponding speed (Figure 5c,d) and torque (Figure 5e,f). Although there appear to be minor disagreements in predicting the start of the ICE, the amount of time when the ICE is on duty is correctly predicted, with differences below 8% in both tests. These results confirm that the model provides a correct overview of the powertrain behaviour, demonstrating that the employed vehicle propulsion controller is representative of the real EMSs.

Finally, the electrical energy consumption, a key parameter to be captured when assessing alternative EMSs, is also predicted well. The maximum error showed by the model in predicting the battery SOC (see Figure 5g,f) at the end of the cycle is less than 7%, demonstrating adequate capture of battery energy depletion.

2.4. Methodology for PHEV EMS Validation

2.4.1. Methodology Overview

Figure 6 illustrates the overall methodology for the evaluation of the effectiveness of the proposed intelligent PHEV EMS based on ML vehicle usage modelling (ML-VUM). The MATLAB® Simulink environment was used for model-in-the-loop (MIL) integration, linking the Autonomie PHEV model with the ML-VUM decision block. The evaluation of the ML-VUM PHEV EMS was referenced to the performance of the standard PHEV EMS (i.e., when the ML-VUM decision block is not engaged) for the real-world journey data of 10 vehicles, described in Section 2.2.1. For each vehicle trip in the, the pre-fitted and stored ML-VUM models (Section 2.3) were used to predict the current journey distance in order to support the decision as to whether the ML-VUM EMS can used for the upcoming journey or not. In order to increase the robustness of trip distance prediction, all 3 classification thresholds (i.e., 7 km, 14 km and 21 km) were considered, and the associated RusBoosted tree ensemble models were fitted and validated (see Section 2.2.2). Table 5 summarizes the case-based approach used to select the current trip distance based on the ML-VUM models. The trip distance was used to generate a simple estimate of the energy required to complete the journey by assuming an average velocity for the upcoming journey as the mean of the distribution of observed average velocities for journeys in the respective distance class (e.g., D < 7 km, etc.).

Table 5. The use of ML-VUM to predict current journey distance.

Case	Predicted Trip Distance (D)	Distance Used to Estimate Trip Energy Requirements
1	$D < 7$ km	7 km
2	$7 \leq D < 14$ km	14 km
3	$14 \leq D < 21$ km	21 km
4	$D \geq 21$ km	ML-VUM strategy not applicable

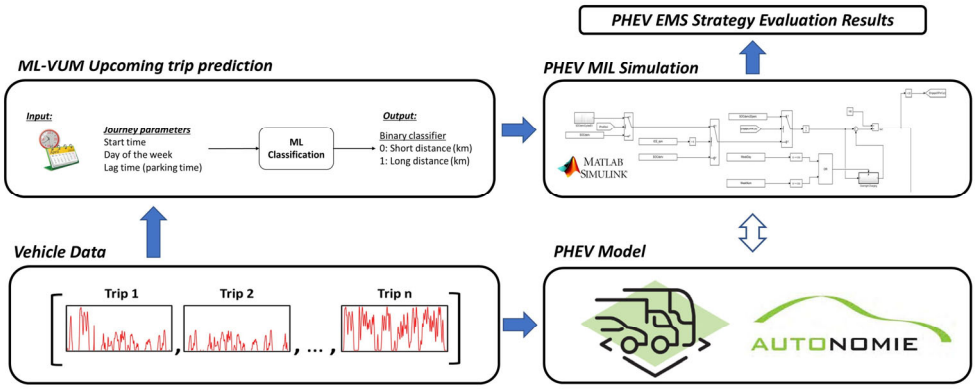


Figure 6. The overall methodology for the evaluation of the effectiveness of the proposed intelligent PHEV EMS.

The key decision guiding the implementation of the proposed ML-VUM EMS is based on comparing the predicted energy required for the upcoming trip with the current SOC, which is predicted by the vehicle model at the end of the previous journey. The following assumptions were considered for the MIL simulation approach:

- The battery is fully charged before the first trip of the day, and it is not charged until the end of the same day, i.e., no re-charging during the day was considered in this work;
- The minimum SOC allowed was set at 20%, at which point the switch from charge depletion to charge sustaining mode will be implemented; if the SOC limit is reached during a journey started under the ML-VUM strategy (e.g., if the trip distance was misclassified), the EMS will revert to the default strategy;
- The proposed ML strategy will be disabled when the vehicle speed exceeds 90 km/h (motorway conditions) to ensure optimal battery discharging;
- No route gradient has been considered in the simulation, as altitude data were not available.

For the MIL simulation set-up in Figure 6, drive cycle data were augmented synthetically to 1 Hz frequency (using the procedure described in the following section), and fed to the Autonomie PHEV simulation block to generate the current vehicle and battery status parameters, and simulate the vehicle performance over the trip, with either stock EMSs or the ML-VUM EMS.

2.4.2. Synthetic Drive Cycle Data for MIL Validation

A major challenge for the validation experiment was that the historic real-world trip data for the 10 vehicles (described in Section 2.2.1) only included summary statistics for each trip, rather than the detailed (1 Hz) data needed for running the MIL simulation. To overcome this challenge, an approach based on synthetic augmentation of the data was employed. This approach was inspired by studies discussed in the literature, e.g., [19,22], where algorithms were employed to identify segments of past drive cycles that were likely to occur as the next sequence based on the current vehicle status and data (e.g., velocity). The research team had access to a separate data pool from a development fleet of vehicles instrumented with on-board condition monitoring and sampling equipment, recording ECU signals from real-world drive cycles undertaken to replicate typical customer usage (e.g., urban, highway, countryside, and mixed drive cycles), sampled at a 1 Hz frequency. A set of nine reference drive cycles was selected, with durations of between 1800 and 3600 s. The procedure involved matching each trip for each of the 10 vehicles with one of the instrumented drive cycles, based on an assumed journey type inferred from the average velocity and journey distance. A drive cycle sequence was extracted to precisely match the journey duration and average velocity, also ensuring coherence and continuity. Where the

journey was longer than the corresponding drive cycle, the drive cycle was extended by repeating sequences while also assuring the matching of observed journey average velocity. This resulted in an appropriate sequence of vehicle trips data with 1 Hz frequency drive cycle data, feeding the Autonomie ML model, as illustrated in Figure 6.

3. Results

3.1. Impact of ML-VUM EMSs on the ICE Operation and Tailpipe CO_2 Emissions

In order to illustrate the impact of ML-VUM EMSs on the powertrain operation, Figure 7 includes the comparison of vehicle speed profile, engine operation, and SOC between the reference case (standard PHEV EMSs) and ML-VUM EMSs for one randomly selected journey of vehicle 3. The cycle is defined by the key-on events, shown by green lines in plot (c). In this specific case, at the key-on event, the proposed PHEV EMS controller allows the vehicle to run purely in EV mode, as the energy stored in the battery is sufficient to propel the vehicle with only the e-motor. As a result, the ICE is not switched on during the journey, as Figure 7b reveals. By contrast, under the standard EMS conditions, the ICE does power the vehicle during rapid accelerations. The experimental ICE speed is also included in the plot, confirming the good reproduction of the real vehicle energy management system.

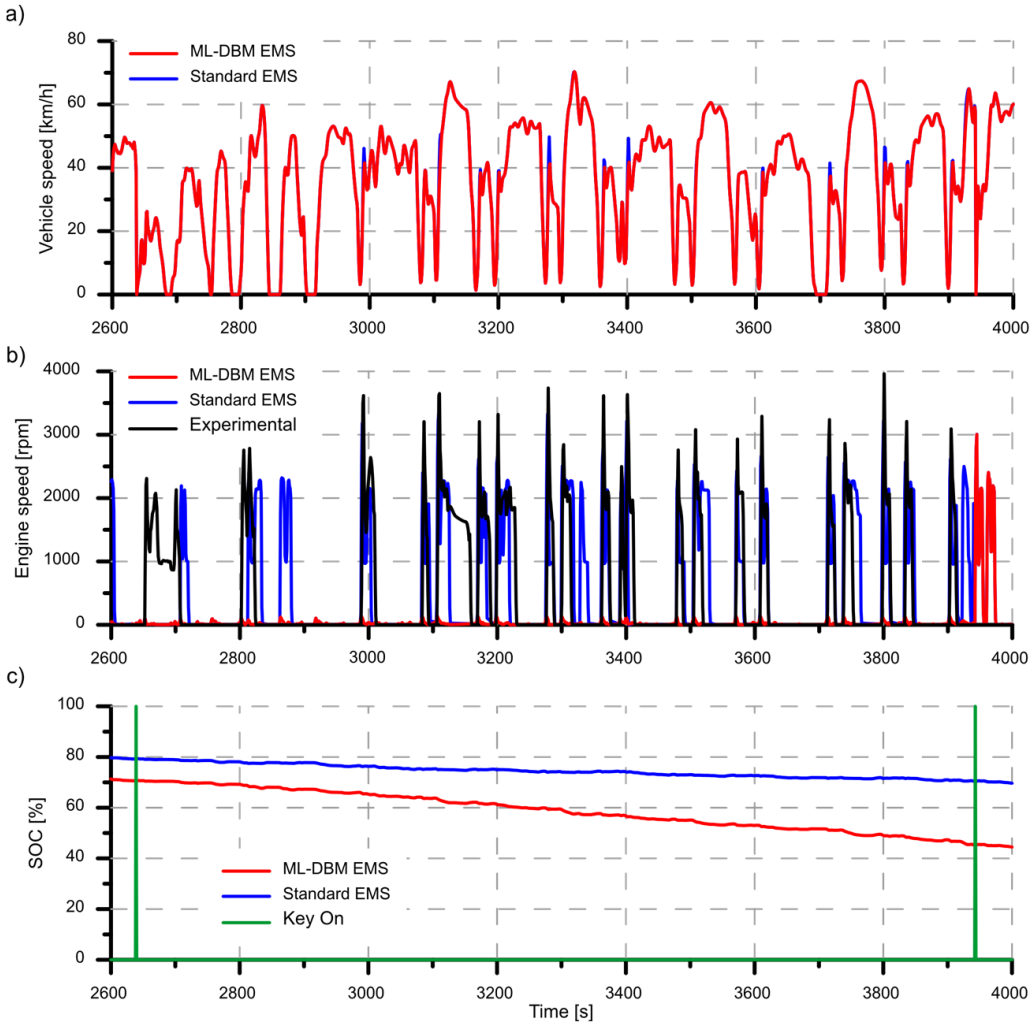


Figure 7. The impact of the ML-VUM EMSs on ICE operation over one driving cycle of driver 3: vehicle speed (**a**), ICE speed (**b**) and SOC (**c**).

As a result of the fully EV mode, the battery depletion is more severe, as depicted in Figure 7c. The lower SOC at the end of the trip could compromise the subsequent trips, especially if those are long or very demanding routes, leading to noticeable shares of unwanted operation under the CS mode. This behaviour points out that the ML-VUM EMS approach would be more efficient if implemented together with opportunity charging patterns in order to minimize the ICE operation.

The quantification of the cumulative time with the ICE on duty was performed over the whole dataset of each vehicle. This is shown in Figure 8, which includes the cumulative time with the ICE switched on for both the reference EMSs and ML-VUM EMSs over the whole driving time simulation representing 12 months of driving.

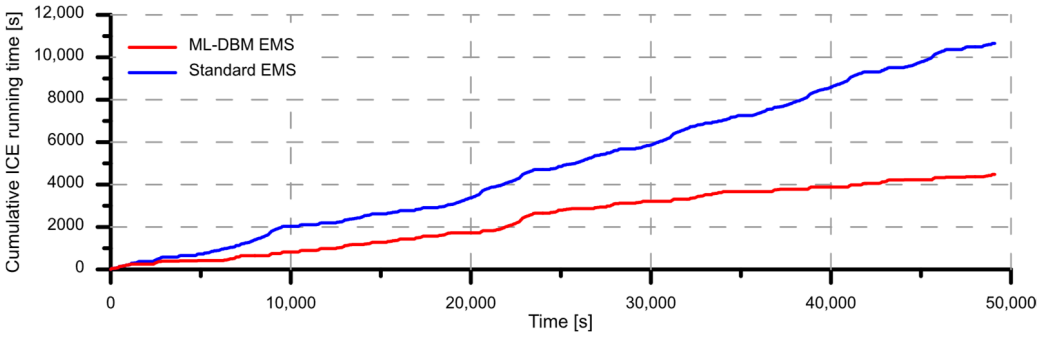


Figure 8. The cumulative time with the internal combustion engine switched on over the entire dataset: vehicle 10.

Based on the simulation results with the reference case, the ICE remains on duty for 11k seconds over the 50k seconds of the whole dataset, which means 22% of the total driving time. By contrast, employing the ML-VUM EMS, the operation of the ICE can be reduced to 4250 s, reaching a reduction figure of approximately 60% without any change in the hardware (larger battery pack) or the charging pattern (opportunity charging).

The evaluation of cumulative ICE operation across all 10 vehicles for the proposed ML-VUM versus the standard EMS is summarized in Figure 9. The plot also indicates the percentage reduction in the usage of ICE, ranging between 30% (vehicle 9) and 76% (vehicle 8). Given that assumptions were uniformly applied (e.g., the overnight charging events are the same), and the accuracy of the trip distance prediction is reasonably similar (Figure 4), variability is likely due to the journey and driving behaviour factors, including route type and aggressiveness.

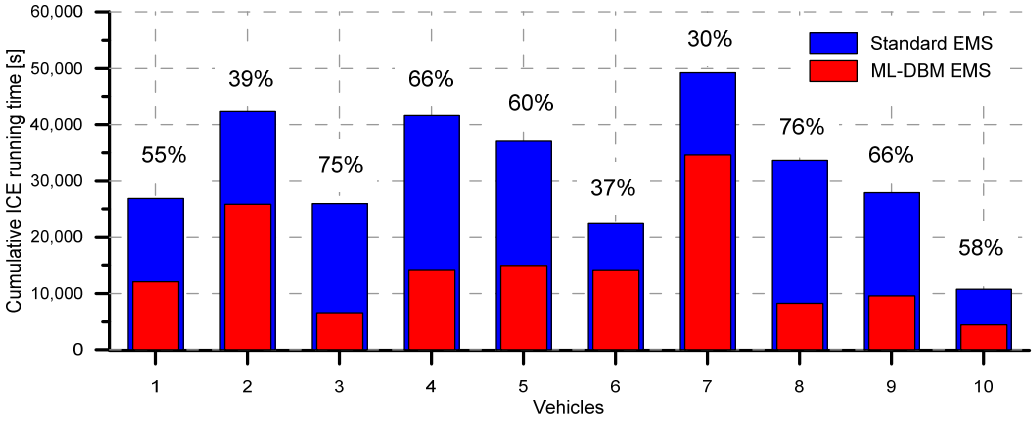


Figure 9. ICE operation reduction by applying ML strategy: cumulative time comparison and percentage reduction for every driver.

Key features for each vehicle (standard EMSs) are shown in Table 6, which allows a deeper analysis of the vehicle usage impact on the potential benefit of the ML-VUM EMS. The daily average distance travelled does play a key role. Patterns with daily distances longer than the EV range of the vehicle (approximately 30 km under the NEDC test procedure) would lead to frequent operations in CS mode. Drivers 2 and 7 present the highest daily average travelled distance (see Table 6), which is much larger than the rest of the vehicles. As a result, the percentages of days in which the vehicle operates under CS mode (69% and 44% respectively) reduce the potential number of trips in purely EV modes.

Table 6. Vehicle profile driving features.

Vehicle	Daily Average Trip Distance [km]	Average Speed [km]	Motorway Share [%]	RPA [m/s^2]	PKE [m/s^2]	CS Mode Is Reached [% of Days]
1	19.08	38.51	1.49	0.22	0.49	8
2	40.40	38.52	1.75	0.23	0.49	69
3	19.04	36.75	0.82	0.25	0.53	1
4	29.31	38.88	2.04	0.23	0.49	23
5	22.88	40.76	3.18	0.22	0.47	9
6	29.64	35.85	1.40	0.25	0.51	28
7	37.02	35.63	1.35	0.22	0.48	44
8	27.55	40.51	3.51	0.23	0.49	8
9	26.61	39.10	2.09	0.24	0.51	13
10	20.22	38.84	3.00	0.23	0.50	15

The opposite usage is observed in drivers 1 and 3, with daily travel distances below 20 km/day. In those cases, with very limited events running in CS mode, the benefit of the proposed EMS is relevant, with large reductions in the ICE operation.

Driving aggressiveness is also a factor that might influence the potential outcome of the ML-VUM EMS. On one hand, larger proportions of rapid acceleration are directly related to longer periods of ICE support with the standard EMS. Therefore, the ML-VUM EMS application should significantly reduce the ICE operation, at the expense of vehicle performance derating. On the other hand, energy consumption is directly related to the aggressiveness, accelerating the battery depletion, and compromising the following cycles until the next recharging event. Drivers' aggressiveness is quantified in Table 3 through the relative positive acceleration (RPA) and positive kinetic energy (PKE), widely accepted as relevant indicators of driving aggressiveness [54,55]. However, the calculated values do not show noticeable differences between and does not immediately point to clear conclusions. This is exemplified by comparing vehicles 1 and 3, and 4 and 6, respectively; despite showing similar travel distance patterns, the higher aggressiveness has a positive impact when comparing drivers 1 and 3, but a negative impact in the comparison of drivers 4 and 6.

From the MIL simulation exercise, it is possible to evaluate the fuel economy figures and the corresponding tailpipe CO_2 emissions for each vehicle with both the standard EMS and the proposed ML-VUM EMS. These results are shown in Figure 10. The average fuel consumption ranges from 3 to 4.2 L/100 km (see Figure 10a) for the standard PHEV EMS and progresses in accordance with the World Harmonized Light-duty Vehicle Test Procedure (WLTP) fuel economy figures expected for an SUV PHEV. If the ML-VUM EMS is applied, the fuel economy improves significantly to around 1.5 L/100 km, which is in the expected range of compact PHEVs [56].

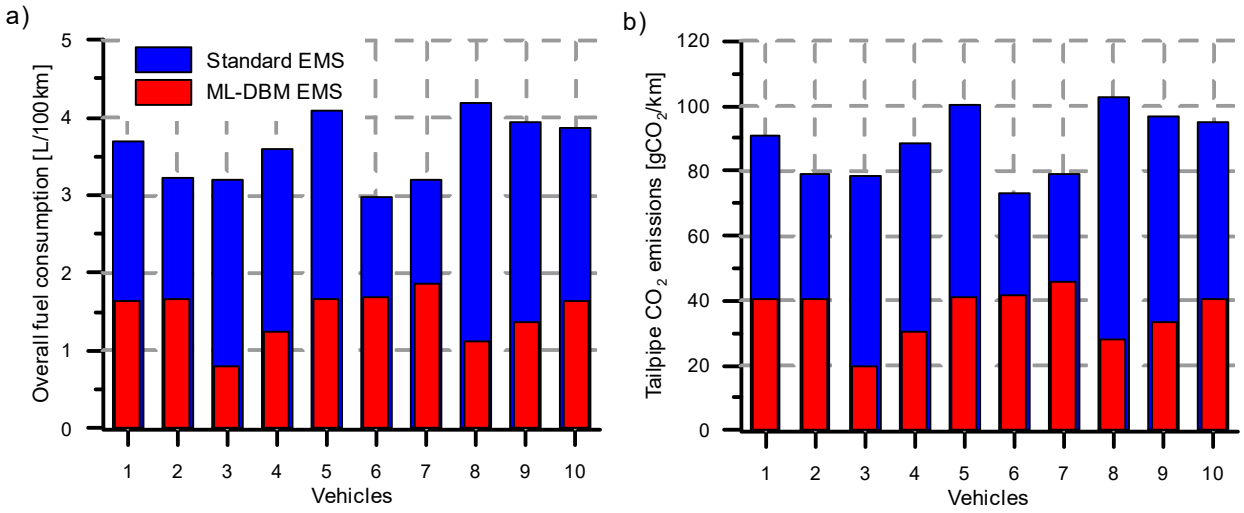


Figure 10. Overall fuel economy (**a**) and tailpipe CO_2 emissions (**b**) comparison between the reference and the ML strategy for every driver.

As a result of the lower fuel consumption, the tailpipe CO_2 emissions are reduced considerably, as Figure 10b depicts. From an average value of 88.3 gCO_2/km, the use of the ML-VUM EMS reduces the tailpipe CO_2 emissions to 36.1 gCO_2/km.

3.2. Impact of ML-VUM EMS on Vehicle Performance

The lack of ICE support in the event of demanding maneuvers has an impact on the vehicle performance, e.g., not being able to follow the reference speed trace. This behaviour can be seen in Figure 11, where vehicle speed and engine speed traces for two examples from vehicles 3 (plots (a) and (c)) and 9 (plots (b) and (d)) are illustrated.

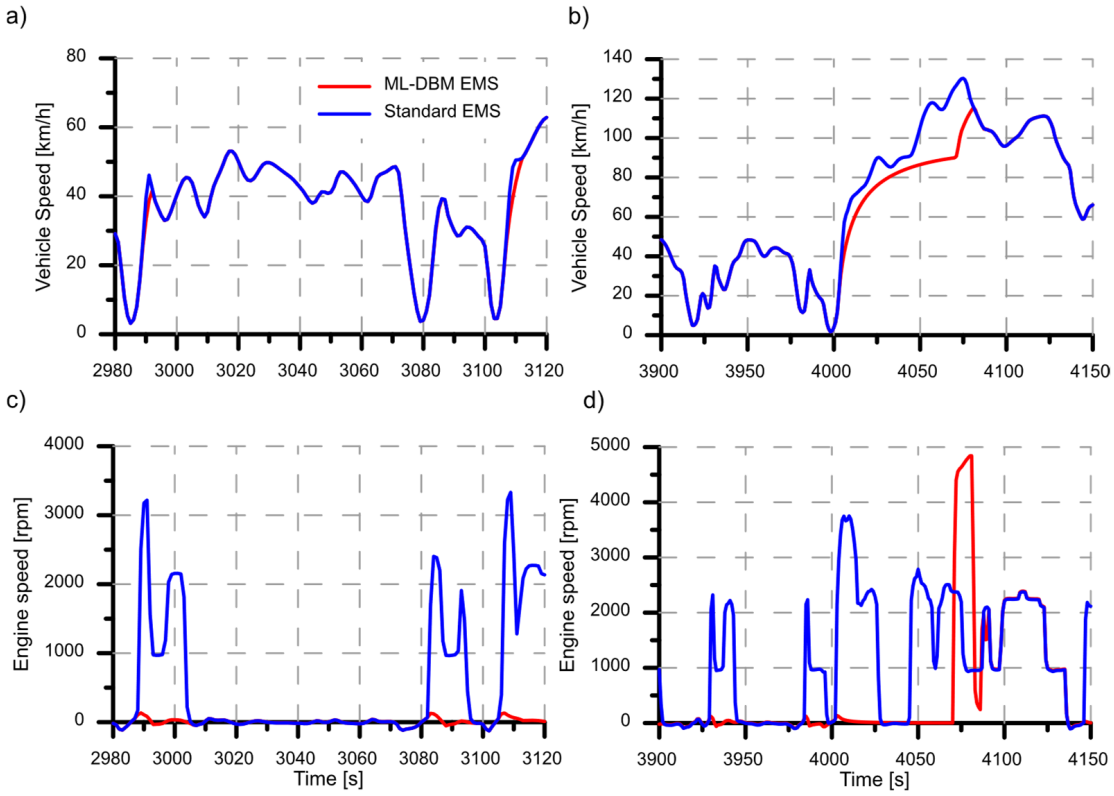


Figure 11. Example of performance loss when ML-VUM EMS is applied: vehicle and engine speed for vehicle 3 (**a**,**c**) and for vehicle 9 (**b**,**d**).

The sequence shown for vehicle 3 describes a common usage in urban modes. Three acceleration events are observed, where the ICE starts in the reference case to complement the torque provided by the e-motor (see Figure 11c). Without ICE assistance (ML-VUM EMS), the results show that, in two out of these three accelerations, the e-motor cannot provide enough power to follow the reference speed trace, indicating a vehicle performance loss.

The section of the trip analyzed for vehicle 9 represents an acceleration event from urban to motorway driving conditions. A clear lag during this maneuver can be appreciated when the ML-VUM EMS is used (see Figure 11b). The lack of ICE assistance from 0 to 90 km/h acceleration (see plot (d)), introduced as a boundary of the ML-VUM EMS operation window, impacts the vehicle's performance.

The examples of speed trace missed, leading to a loss of vehicle performance, might result in a lack of consumer acceptance. With the aim of quantifying this damage, the speed-trace-missed approach from homologation tests is applied. The WLTP type-approval test [57] establishes the speed trace tolerances as follows:

- Upper: 2.0 km/h higher than the highest point of the trace within ±1.0 s of the given point in time;
- Lower: 2.0 km/h lower than the lowest point of the trace within ±1.0 s of the given time.

The test procedure accepts speed tolerances larger than these as long as they are never exceeded for more than 1 s, with a maximum of 10 deviations per test. Based on the latter figure, the percentage of traces missed would be 0.55% (10 s in the entire 1800 of the cycle duration). This figure would be used as a reference to understand the level of performance derating for each driver.

Figure 12 shows the percentage of time with the speed trace out of the limits for all drivers' vehicles running with the ML-VUM EMS, defined by the previous tolerances. Around 50% of the analyzed vehicles presented speed-trace-missed values between 1% and 1.4%, which was relatively close to the figure extracted from the WLTP approval test (0.55%). This result points out that in general, the impact of using the proposed EMS was not severe in the case of those drivers whose data were analyzed. However, in 4 out of the 10 vehicles, this value exceeded 2% (approximately 4 times higher than the figure of WLTP). A common factor appears to be the larger proportion of motorway driving conditions, which cause the longest durations of missed traces due to the lack of power during acceleration (as illustrated with the example of vehicle 9 in Figure 11).

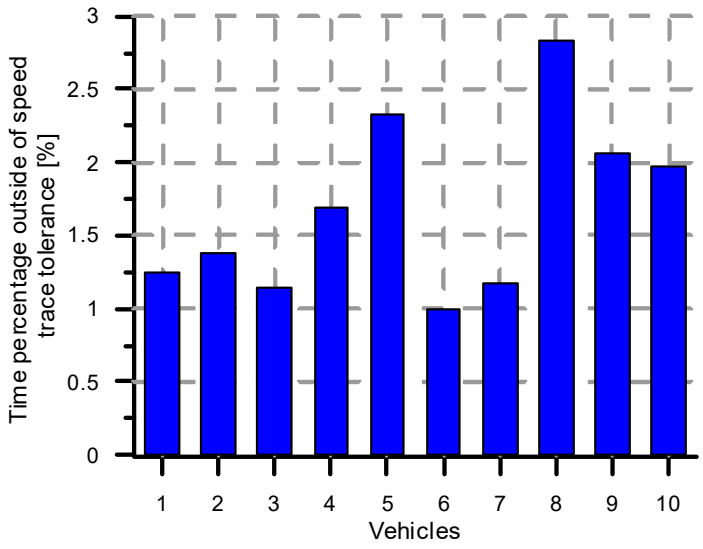


Figure 12. Time percentage outside of speed-trace tolerances for every vehicle.

4. Discussion

While the review of the literature showed significant progress with development of PEO approaches to enhance PHEV EMS and control, most of these methods significantly rely on the availability of external information (like GPS positioning and navigation data, or information from other vehicles or infrastructure). This not only raises issues with the robustness of the strategy in real-world driving, linked to the dependability of the communication services and even the availability of the information (specifically, navigation data are often unavailable for routine customer journeys), but also the computational costs associated with on-board processing.

In contrast, the approach proposed in this work relies on predicting the current journey characteristics based on the learnt usage patterns in terms of routine journeys and driving behaviour. Therefore, this method does not require the availability of online external information. While removing the reliance on external information has the potential to increase robustness, it is necessary to prove the predictive capability of the ML algorithms and the effectiveness of the strategy in real-world driving. The research in this paper sought to address these challenges by presenting a "proof-of-concept" study to evaluate the potential of the proposed strategy based on an MIL simulation experiment.

The first key research question to address was whether repeated routine customer journey behaviour can be learnt using classification ML in a way that gives sufficient prediction accuracy for the upcoming journey to support the implementation of the proposed EMS. The study was based on a large sample of real-world journeys (5554 trips recorded from 10 vehicles over one year) and considered 7 ML algorithms for classification with 3 different trip distance class thresholds. Classic machine learning was preferred due to the nature of the data and to reduce computational expense for online application. From an engineering point of view, an ML algorithm that provides consistent performance across a diverse range of data (in this case, in terms of the size of training sample from vehicle to vehicle, and the significantly different patterns of imbalance, reflecting heterogeneity in the vehicles; driving patterns and the multiple class thresholds) is preferred to a tailored approach, with different ML models used for each vehicle. The RusBoosted ensemble algorithm showed the best overall performance (as documented in Table 3) and a good prediction performance for all cases (as shown in Figure 6). It can support the deployment of the proposed EMS. The RusBoosted algorithm does not have high computational requirements compared to other algorithms [58], which makes it a good choice for limited ECU capabilities. This is in line with similar studies based on journey and driver behaviour characteristics (e.g., [35], based on a diesel SUV), which also found RusBosted ensemble classifier to perform well.

Considering further the vehicle implementation viewpoint, it can be envisaged that the ML models can be regularly updated using new data. From a practical implementation viewpoint, this can be carried out either using the ECU computational capacity (in this case, trained models are kept in memory and can be accessed offline for predictions at any time), or remotely on the server through data-over-the-air (DOTA) technologies. Other ML algorithms and strategies could be considered for evaluation under such scenarios, but this is beyond the proof-of-concept purpose of this work.

The MIL validation experiments were carried out on a comprehensive set of synthetically generated journeys data, simulating the real-world journeys of 10 vehicles over one year. The fact that synthetic data (generated using actual drive cycle data collected from development vehicles) was used instead of actual data (not commonly available for such a large sample) does not affect the validity of the comparative evaluation study between the two EMSs. A limitation of the drive cycle data used is that it does not include road quality and gradient information, which have an impact on the energy consumption, as the data available for the study did not include this information directly.

The vehicle MIL simulation exercise showed significant reductions in ICE operation, with values up to 76% when the ML-VUM EMS, with a corresponding significant reduction in tailpipe emissions from 88.3 gCO_2/km to 36.1 gCO_2/km. The reduction in ICE engine usage, in particular the significant reduction in engine starts achieved with the proposed strategy will also lead to higher reliability and longer oil service intervals. Thus, the MIL evaluation study has demonstrated strong potential benefits for the proposed ML-VUM EMS.

The detailed analysis of the simulation results showed a relatively low impact on vehicle performance, quantified by speed-trace deviations observed for some cases, in particular, in conjunction with more demanding road segments or traffic situations or more aggressive driving styles. The actual impact on the impact on the driveability perception was not possible from this MIL study, and was not part of the proof-of-concept scope for the work. This could be studied in more detailed within the implementation phase of the EMS with vehicle testing, and perhaps in conjunction with further refinements of the strategy—which could include blending the ML-VUM with the standard EMS. An alternative option is to study the user acceptance of the ML-VUM strategy as an intelligent-ECO strategy that provides fuel and emission benefits, as well as enhanced reliability, in exchange for a minor compromise in performance.

Overall, these results show that PHEV architectures with EMSs focused on energy savings are effective tools to reduce the well-to-wheel (WTW) CO_2 emissions, especially in geographical areas with high CO_2 grid intensity values [59,60]. In addition, these approaches can also reduce the gap between real-world fuel economy/emissions figures and the type approval ones for PHEVs, as is described well in the literature [61].

5. Conclusions

This paper has proposed a novel strategy for intelligent PHEV EMSs based on the ML prediction of the upcoming journey, to minimize the use of the ICE engine on repeat routine customer journeys of shorter distances, without recourse to navigation data or expensive storage of drive cycle data, commonly employed by other strategies.

The work has demonstrated that trip patterns can be learnt quite effectively using classic ML algorithms. The study based on an extended real-world data (journeys history from 10 vehicles over 12 months), found the RusBoosted ensemble classifier to perform consistently well across the heterogeneous dataset (different volumes of training data, and highly variable imbalance in the datasets reflecting the natural variability in the vehicle usage profiles by different users), providing in all vehicle cases sufficiently accurate predictions for the EMS.

The validation experiments carried out on an MIL set-up, based on a validated Autonomie model of the SUV PHEV developed for this research, showed that the proposed ML-VUM EMS has the potential to deliver a significant reduction in ICE usage on routine journeys, with associated benefits in terms of CO_2 consumption and emissions, as well as ICE engine reliability and extended service intervals. While the detailed analysis has shown a potential minor impact on drivability associated with more demanding road conditions or driver styles, these could be addressed through the detailed control strategy implementation. This study was set up as "proof of concept" for the proposed EMS, and the model demonstrated a very good potential.

The broader importance of this study is that it demonstrates the great potential of using predictive insight from ML deployed to learning vehicle usage patterns, to optimize the control strategies of the vehicle. Specifically, the study demonstrates the significant impact that can be achieved with using classic ML algorithms, with reduced computational expense, hence implementable in ECUs, and based on a very limited, carefully selected

predictor features. While more sophisticated ML approaches could be used, based on the ever-expanding volumes of data collected from the connected vehicles, these come with significant computational expense and dependability issues. In contrast, the approach described in this paper does not rely on external inputs at run time, and thus has inherent robustness by design.

Author Contributions: Conceptualization, A.D., A.K., O.G.-A. and F.C.; methodology, A.D., A.K., F.C. and O.G.-A.; software, A.D., A.K. and O.G.-A.; validation, A.D., A.K., F.C. and O.G.-A.; formal analysis, F.C. and E.A.; investigation, A.D., A.K. and F.C.; resources, O.G.-A.; data curation, A.D. and A.K.; writing—original draft preparation, A.D., A.K., F.C. and O.G.-A.; writing—review and editing, A.D., A.K., F.C., O.G.-A. and E.A.; supervision, F.C. and E.A. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge partial funding support for the research presented in this article from the UK Advanced Propulsion Centre/Institute of Digital Engineering proof-of-concept project JLR/aiR-FORCE/J14921 "AI for Reliability-based Feature Optimisation with Driver Contextual Intelligence", collaboration between Jaguar Land Rover and the University of Bradford (https://www.ide.uk/project-air-force).

Data Availability Statement: The datasets presented in this article are not readily available because of confidentiality. Requests to access the datasets should be directed to the author Enrico Agostinelli.

Acknowledgments: The authors also wish to acknowledge the technical input from Emanuele Angiolini and Chunxing Lin during the aiR-FORCE project.

Conflicts of Interest: Author Enrico Agostinelli was employed by the company Jaguar Land Rover. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

CAE	computer-aided engineering
CD	charge depletion
CS	charge sustained
DOTA	data over the air
VUM	vehicle usage modelling
EV	electrical vehicle
EMS	energy management system
ICE	internal combustion engine
NEDC	New European Driving Cycle
MIL	model in the loop
ML	machine learning
SOC	state of charge
PHEV	plug-in hybrid electric vehicle
WLTP	Worldwide harmonized Light vehicles Test Procedure
WTW	well-to-wheels emissions

References

1. Department for Transport (HM Government, UK). *Transitioning to Zero Emission Cars and Vans: 2035 Delivery Plan*; HM Government: London, UK, 2021.
2. Department for Transport (HM Government, UK). *Taking Charge: The Electric Vehicle Infrastructure Strategy*; HM Government: London, UK, 2021.
3. Zaino, R.; Ahmed, V.; Alhammadi, A.M.; Alghoush, M. Electric Vehicle Adoption: A Comprehensive Systematic Review of Technological, Environmental, Organizational and Policy Impacts. *World Electr. Veh. J.* **2024**, *15*, 375. [CrossRef]
4. Santana-Méndez, I.; García-Afonso, Ó.; González-Díaz, B. Full Road Transport Sector Transition Towards 100% Autonomous Renewable Energy Supply in Isolated Systems: Tenerife Island Test Case. *Appl. Sci.* **2024**, *14*, 9734. [CrossRef]

5. Smith, D.E.; Lohse-Busch, H.; Irick, D.K. A preliminary investigation into the mitigation of plug-in hybrid electric vehicle tailpipe emissions through supervisory control methods. *SAE Int. J. Engines* **2010**, *3*, 996–1011. [CrossRef]
6. Growney, D.; Joedicke, A.; Williams, M.; Robin, M.; Mainwaring, R.; Davies, M. Hybrid Electric Vehicle Engine Operation and Engine Oil Degradation: A Research Approach. *SAE Tech. Pap.* **2023**, *17*, 3–16. [CrossRef]
7. Plianos, A.; Jokela, T.; Hancock, M. Predictive energy optimization for connected and automated HEVs. *SAE Tech. Pap.* **2018**, 2018-01-1179.
8. Zhang, F.; Hu, X.; Langari, R.; Cao, D. Energy management strategies of connected HEVs and PHEVs: Recent progress and outlook. *Prog. Energy Combust. Sci.* **2019**, *73*, 235–256. [CrossRef]
9. Zhang, C.; Vahidi, A. Route Preview in Energy Management of Plug-in Hybrid Vehicles. *IEEE Trans. Control Syst. Technol.* **2012**, *20*, 546–553. [CrossRef]
10. Du, A.; Han, Y.; Zhu, Z. Review on multi-objective optimization of energy management strategy for hybrid electric vehicle integrated with traffic information. *Energy Sources Part A Recovery Util. Environ. Eff.* **2022**, *44*, 7914–7933. [CrossRef]
11. Martinez, C.M.; Hu, X.; Cao, D.; Velenis, E.; Gao, B.; Wellers, M. Energy Management in Plug-in Hybrid Electric Vehicles: Recent Progress and a Connected Vehicles Perspective. *IEEE Trans. Veh. Technol.* **2017**, *66*, 4534–4549. [CrossRef]
12. Sun, C.; Moura, S.J.; Hu, X.; Hedrick, J.K.; Sun, F. Dynamic Traffic Feedback Data Enabled Energy Management in Plug-in Hybrid Electric Vehicles. *IEEE Trans. Control Syst. Technol.* **2015**, *23*, 1075–1086.
13. Isenstadt, A.; Yang, Z.; Searle, S.; German, J. *Real World Usage of Plug-in Hybrid Vehicles in the United States*; The International Council on Clean Transportation: Washington, DC, USA, 2022.
14. Mechichi, O.; El Amraoui, L.; Trigui, R. Impact of Rule-based EMS parameters and charging strategies on Plug-in HEV efficiency. In Proceedings of the IEEE Vehicle Power and Propulsion Conference (VPPC), Gijon, Spain, 25–28 October 2021.
15. Yang, C.; Wang, M.; Wang, W.; Pu, Z.; Ma, M. An efficient vehicle-following predictive energy management strategy for PHEV based on improved sequential quadratic programming algorithm. *Energy* **2021**, *219*, 119595. [CrossRef]
16. Guanetti, J.; Kim, Y.; Borrelli, F. Control of connected and automated vehicles: State of the art and future challenges. *Annu. Rev. Control* **2018**, *45*, 18–40. [CrossRef]
17. Fan, L.; Zhang, Y.; Dou, H.; Zou, R. Design of an integrated energy management strategy for a plug-in hybrid electric bus. *J. Power Sourc.* **2020**, *448*, 227391. [CrossRef]
18. Liang, B.; Sun, C.; Liu, B. Connected PHEV Energy Management based on Global Driving Cycle Construction. In Proceedings of the 33rd Chinese Control and Decision Conference (CCDC), Kunming, China, 22–24 May 2021.
19. Zhang, Q.; Fu, X. A Neural Network Fuzzy Energy Management Strategy for Hybrid Electric Vehicles Based on Driving Cycle Recognition. *Appl. Sci.* **2020**, *10*, 696. [CrossRef]
20. Shi, D.; Li, S.; Liu, K.; Wang, Y.; Liu, R.; Guo, J. Adaptive Energy Management Strategy Based on Intelligent Prediction of Driving Cycle for Plug−In Hybrid Electric Vehicle. *Processes* **2022**, *10*, 1831. [CrossRef]
21. Lin, X.; Zhou, K.; Mo, L.; Li, H. Intelligent Energy Management Strategy Based on an Improved Reinforcement Learning Algorithm With Exploration Factor for a Plug-in PHEV. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 8725–8735. [CrossRef]
22. Li, S.; Gu, C.; Zhao, P.; Cheng, S. Adaptive energy management for hybrid power system considering fuel economy and battery longevity. *Energy Convers. Manag.* **2021**, *235*, 114004. [CrossRef]
23. Guo, J.; He, H.; Peng, J. Real-time Energy Management for Plug-in Hybrid Electric Vehicle based on Economy Driving Pro System. *Energy Procedia* **2019**, *158*, 2689–2694. [CrossRef]
24. Hongwen, H.; Jinquan, G.; Jiankun, P.; Huachun, T.; Chao, S. Real-time global driving cycle construction and the application to economy driving pro system in plug-in hybrid electric vehicles. *Energy* **2018**, *152*, 95–107. [CrossRef]
25. Climent, H.; Pla, B.; Bares, P.; Pandey, V. Exploiting driving history for optimising the Energy Management in plug-in Hybrid Electric Vehicles. *Energy Convers. Manag.* **2021**, *234*, 113919. [CrossRef]
26. Qi, X.; Luo, Y.; Wu, G.; Boriboonsomsin, K.; Barth, M. Deep reinforcement learning enabled self-learning control for energy efficient driving. *Transp. Res. Part C Emerg.* **2019**, *99*, 67–81. [CrossRef]
27. Zhang, Y.; Liu, H.; Zhang, Z.; Luo, Y.; Guo, Q.; Liao, S. Cloud computing-based real-time global optimization of battery aging and energy consumption for plug-in hybrid electric vehicles. *J. Power Sourc.* **2020**, *479*, 229069. [CrossRef]
28. Mohseni, N.A.; Bayati, N.; Ebel, T. Energy management strategies of hybrid electric vehicles: A comparative review. *IET Smart Grid* **2024**, *7*, 191–220. [CrossRef]
29. Levenberg, K. A method for the solution of certain non-linear problems in least squares. *Q. Appl. Math.* **1944**, *2*, 164–168. [CrossRef]
30. Mallet, A. A maximum likelihood estimation method for random coefficient regression models. *Biometrika* **1986**, *73*, 645–656. [CrossRef]
31. Stewart, M. The actual difference between statistics and machine learning. *Towards Data Sci.* **2019**, *24*, 19.
32. Sinaga, K.P.; Yang, M.S. Unsupervised K-means clustering algorithm. *IEEE Access* **2020**, *8*, 80716–80727. [CrossRef]

33. Malik, A.; Tuckfield, B. *Applied Unsupervised Learning with R: Uncover Hidden Relationships and Patterns with K-Means Clustering, Hierarchical Clustering, and PCA*, 1st ed.; Packt Publishing: Birmingham, UK, 2019.
34. Taye, M.M. Understanding of machine learning with deep learning: Architectures, workflow, applications and future directions. *Computers* **2023**, *12*, 91. [CrossRef]
35. Doikin, A.; Campean, F.; Priest, M.; Angiolini, E.; Lin, C.; Agostinelli, E. Driving Behavior Modelling Framework for Intelligent Powertrain Health Management. *SAE Int. J. Engines* **2022**, *16*, 443–462. [CrossRef]
36. Marchegiani, L.; Posner, I. Long-term driving behaviour modelling for driver identification. In Proceedings of the 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 4–7 November 2018.
37. Wang, Y.; Ho, I.W.-H. Joint deep neural network modelling and statistical analysis on characterizing driving behaviors. In Proceedings of the IEEE Intelligent Vehicles Symposium (IV), Changshu, China, 26–30 June 1018.
38. Zhang, L.; Zhu, Z.; Zhang, Z.; Song, G.; Zhai, Z.; Yu, L. An improved method for evaluating eco-driving behavior based-on speed-specific vehicle-specific power distributions. *Transp. Res. D Transp. Environ.* **2022**, *113*, 103476. [CrossRef]
39. Li, Z.; Lu, C.; Gong, C.; Gong, J.; Li, J.; Wei, L. Driver behavior modelling at the urban intersection via canonical correlation analysis. In Proceedings of the 3rd International Conference on Unmanned Systems (ICUS), Harbin, China, 27–28 November 2020.
40. Zhu, M.; Wang, X.; Wang, Y. Human-like autonomous car-following model with deep reinforcement learning. *Transp. Res. Part C Emerg.* **2018**, *97*, 348–368. [CrossRef]
41. Amata, H.; Miyajima, C.; Nishino, T.; Kitaoka, N.; Takeda, K. Prediction model of driving behavior based on traffic conditions and driver types. In Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems, St. Louis, MI, USA, 4–7 October 2009.
42. Kolachalama, S.; Malik, H. Predictive model of adaptive cruise control speed to enhance engine operating conditions. *Vehicles* **2021**, *3*, 749–763. [CrossRef]
43. Doikin, A.; Campean, F.; Priest, M.; Lin, C.; Angiolini, E. A reliability inspired strategy for intelligent performance management with predictive driver behaviour: A case study for a diesel particulate filter. *Proc. Des. Soc.* **2021**, *1*, 191–200. [CrossRef]
44. The MathWorks, Inc. Classification Learner. Available online: https://uk.mathworks.com/help/stats/classificationlearner-app.html (accessed on 21 January 2025).
45. Aliyu, R.; Mokhtar, A.A.; Hussin, H. A Study on Comparison of Classification Algorithms for Pump Failure Prediction. *J. Adv. Ind. Technol. Appl.* **2023**, *4*, 48–65.
46. Mounce, S.; Ellis, K.; Edwards, J.; Speight, V.; Jakomis, N.; Boxall, J. Ensemble decision tree models using RUSBoost for estimating risk of iron failure in drinking water distribution systems. *Water Resour. Manag.* **2017**, *31*, 1575–1589. [CrossRef]
47. Zhou, Z.H. *Ensemble Methods: Foundations and Algorithms*, 1st ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2012.
48. Kivekäs, K.; Lajunen, A.; Vepsäläinen, J.; Tammi, K. City Bus Powertrain Comparison: Driving Cycle Variation and Passenger Load Sensitivity Analysis. *Energies* **2018**, *11*, 1755. [CrossRef]
49. Xu, X.; Aziz, H.M.A.; Guensler, R. A modal-based approach for estimating electric vehicle energy consumption in transportation networks. *Transp. Res. D Transp. Environ.* **2019**, *75*, 249–264. [CrossRef]
50. Yao, J.; Moawad, A. Vehicle energy consumption estimation using large scale simulations and machine learning methods. *Transp. Res. Part C Emerg.* **2019**, *101*, 276–296. [CrossRef]
51. Garcia-Afonso, O. Impact of powertrain electrification on the overall CO_2 emissions of intercity public bus transport: Tenerife Island test case. *J. Clean. Prod.* **2023**, *412*, 137365. [CrossRef]
52. Simulate State-of-the-Art Vehicle Energy Consumption, Performance, and Cost with Autonomie. Available online: https://vms.taps.anl.gov/tools/autonomie (accessed on 13 January 2023).
53. Shen, D.; Karbowski, D.; Jeong, J.; Kim, N.; Rousseau, A. Energy-Efficient Cruise Control Using Optimal Control for a Hybrid Electric Vehicle. In Proceedings of the 30th International Electric Vehicle Symposium and Exhibition, Stuttgart, Germany, 10 October 2017.
54. Gallus, J.; Kirchner, U.; Vogt, R.; Benter, T. Impact of driving style and road grade on gaseous exhaust emissions of passenger vehicles measured by a Portable Emission Measurement System (PEMS). *Transp. Res. D Transp. Environ.* **2017**, *52*, 215–226.
55. Luján, J.M.; Guardiola, C.; Pla, B.; Pandey, V. Impact of driving dynamics in RDE test on NOx emissions dispersion. *Proc. Inst. Mech. Eng. D J. Automob. Eng.* **2020**, *234*, 1770–1778. [CrossRef]
56. Plötz, P.; Link, S.; Ringelschwendner, H.; Keller, M.; Moll, C.; Bieker, G.; Dornoff, J.; Mock, P. *Real-World Usage of Plug-In Hybrid Vehicles in Europe: A 2022 Update on Fuel Consumption, Electric Driving, and CO_2 Emissions*; The International Council of Clean Transportation: Berlin, Germany, 2022.
57. European Parlament and the Council of the European Union. Regulation (EC) No 715/2007 of the European Parliament and of the Council of 20 June 2007 on type approval of motor vehicles with respect to emissions from light passenger and commercial vehicles (Euro 5 and Euro 6) and on access to vehicle repair and maintenance information. *The Official Journal of the European Union*, 29 June 2007, pp. 1–16.

58. Seiffert, C.; Khoshgoftaar, T.M.; Van Hulse, J.; Napolitano, A. RUSBoost: Improving classification performance when training data is skewed. In Proceedings of the 19th International Conference on Pattern Recognition, Tampa, FL, USA, 8–11 December 2008.
59. McLaren, J.; Miller, J.; O'Shaughnessy, E.; Wood, E.; Shapiro, E. CO_2 emissions associated with electric vehicle charging: The impact of electricity generation mix, charging infrastructure availability and vehicle type. *Electr. J.* **2016**, *29*, 72–88. [CrossRef]
60. García-Afonso, Ó.; Santana-Méndez, I.; Delgado-Torres, A.M.; González-Díaz, B. On the road to a sustainable transport mobility in isolated power systems: The role of light-duty powertrain electrification. *J. Clean. Prod.* **2021**, *320*, 128646. [CrossRef]
61. Plötz, P.; Moll, C.; Bieker, G.; Mock, P. From lab-to-road: Real-world fuel consumption and CO_2 emissions of plug-in hybrid electric vehicles. *Environ. Res. Lett.* **2021**, *16*, 054078. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.