

## Article

# Hybrid Electric Vehicle Characteristics Change Analysis Using Mileage Interval Data

Jiyoung Woo <sup>1</sup>, Inbeom Yang <sup>2,\*</sup> and Chongun Pyon <sup>3</sup><sup>1</sup> Department of Big Data Engineering, Soonchunhyang University, Asan 31538, Korea; jywoo@sch.ac.kr<sup>2</sup> Department of Smart Automobile, Soonchunhyang University, Asan 31538, Korea<sup>3</sup> Department of Humanities and Social Science, Khalifa University, Abu Dhabi 127788, UAE; chongun.pyon@ku.ac.ae

\* Correspondence: ibyang@sch.ac.kr; Tel.: +82-41-530-1528

Received: 1 July 2020; Accepted: 7 August 2020; Published: 10 August 2020



**Abstract:** In this work, the relationship between the accumulated mileage of a hybrid electric vehicle (HEV) and the data provided from vehicle parts has been analyzed. Data were collected while traveling over 70,000 km in various paths. The collected data were aggregated for 10 min and characterized in terms of centrality and variability. It has been examined whether the statistical properties of vehicle parts are different for each cumulative mileage interval. When the cumulative mileage interval is categorized into 30,000–50,000, 50,000–60,000, and 60,000–70,000, the statistical properties contributed in classifying the mileage interval with accuracy of 92.68%, 82.58%, and 80.65%, respectively. This indicates that if the data of the vehicle parts are collected by operating the HEV for 10 min, the cumulative mileage interval of the vehicle can be estimated. This makes it possible to detect abnormality or characteristics change in the vehicle parts relative to the accumulated mileage. It also can be used to detect abnormal aging of vehicle parts and to inform maintenance necessity. Furthermore, a part or module that has a significant change in characteristics according to the mileage interval has been identified.

**Keywords:** hybrid electric vehicle (HEV); driving data; mileage interval; machine learning; characteristics change analysis

## 1. Introduction

There has been continuous development and distribution of high-performance electric vehicles, including smart cars featured with artificial intelligence and partial self-driving functions. These are based on innovative technologies in various fields, such as semiconductor and information communication technology. These technologies, in combination with social demands, have led to the progress of a hyperconnected society, resulting in changes in perception in the economy, society, and culture. A sharing economy represents these changes, and new business models must be developed in new vehicle-sharing systems in large cities. These changes will continue to be incorporated in innovations in the automotive industry, related upstream or downstream industry, and automobile culture in the future.

The scope of innovation in the automotive industry will be expanded to cars owned by corporates as well as those owned by individuals. It will include an effective and safe maintenance and management system so that periodic maintenance and post-repair management will be transformed into preventative maintenance and predictive management. Therefore, techniques such as condition-based maintenance (CBM) and prognostic health monitoring (PHM) are being gradually applied to the automotive industry.

Besides CBM and PHM, various other methodologies and approaches have been discussed to predict the condition of cars, based on real driving data. Because it is physically limited to save,

transmit, and analyze all types of data for a car in real-time, it is feasible to transmit the analyzed data through self-examination. Hence, the development of a basis for self-examination and derivation of a standardized analysis method using real driving data is crucial.

With the commercialization of electric vehicles, including HEVs and smart vehicles capable of partially autonomous driving, the complexity of vehicle functions and structures is increasing. Therefore, various studies have been conducted to monitor and predict the conditions of vehicle parts, modules, and systems as well as the driver's condition for driving safety and maintenance. For example, since the health level of a HEV battery is complexly linked to a driving environment and various parts, a new analysis method based on data is required.

It was difficult to collect data from vehicles in real-time, and hence, the data of vehicles based on the simulation were significantly different from the actual values. Nowadays, many studies are using the real driving data of vehicles. However, the studies on HEVs are in the beginning stage. There were few attempts to identify the car mileage by using the in-vehicle controller area network (CAN) data.

The US government has been conducting a project called Advanced Vehicle Testing Activity, which studies the mutual interconnection method between plug-in hybrid electric vehicles (PHEVs) and electric grids, evaluates vehicle monitoring, and facilitates the commercialization of PHEVs [1]. The US Department of Energy has been performing real-time monitoring and data analysis of hybrid cars and fuel cell vehicles through the Hydrogen Program and HyFLEET-CUTE Project [2,3].

Rezvanizani et al. [4] provided an overall review of PHM technology and practical solutions, including hybrid electric battery life prediction. You et al. [5] and Guo et al. [6] proposed an efficient method to estimate the state of a hybrid electric battery with high accuracy, using in-vehicle data. Mi et al. [7] estimated and verified the SOH of a hybrid electric battery using in-vehicle data and a genetic algorithm. Laayouj et al. [8] conducted a study to predict the remaining useful life of a hybrid electric battery using physical models and in-vehicle data.

Wakita et al. [9], Miyajima et al. [10], Nishiwaki [11], and Kwak et al. [12] studied the discrimination of drivers by collecting the telecommunication data of vehicle parts. Given that the data value of vehicle parts varies according to the driver's driving pattern, they constructed machine learning models to discriminate drivers.

Meng [13] modeled drivers' behavior changes by using a probabilistic model based on the data obtained from accelerators, brakes, and steering wheels, and predicted the users' driving patterns. Choi [14] proposed a model to detect drivers' carelessness using CAN data. Wahab [15] found that the pattern of stepping on the accelerator and the pressure of stepping on the brakes are prominent factors in determining drivers' profiling. Kedar-Dongarka [16] classified the drivers' characteristics into conservative, neutral, and aggressive by using the data retrieved from accelerators, brakes, and gears. Enev et al. [17] classified the road condition into the driving section and parking section based on the position of brake pedals, angle of steering wheel, long-term acceleration, speed of rotation, speed of driving, gear shifting, position of accelerator pedals, speed of engine, maximum engine torque, fuel consumption rates, and position of fuel control valve.

In this study, certain criteria for determining the vehicle condition according to the mileage have been developed by analyzing long-term driving data of a HEV. We also utilized a platform for collecting real-time driving data provided by an on-board diagnostics II (OBD-II) port for the long term. The driving data of major parts, such as motors, inverters, and high-voltage batteries, of the test HEV in terms of seasons and paths have been collected by using a real-time driving monitoring system. The changes in the conditions of the parts and modules according to the cumulative mileage-based collected data have been analyzed. A model has been constructed for identifying the mileage interval by using these variables and identified a part or module that has a significant change in characteristics according to the mileage interval.

This study can be used for estimation of aging status, detection of malfunctions, and lifetime prediction of parts and modules by classifying and identifying the change in characteristics according to the cumulative driving mileages of hybrid cars. Furthermore, this study can be used for the

development of optimized efficient HEV control strategies, and for the preventive maintenance of shared cars by monitoring characteristic changes of the parts and modules according to the mileage with the real driving data.

## 2. Driving Data Collection

### 2.1. Data Collection System

To obtain the driving data of a hybrid car, a real-time driving monitoring system has been utilized, which can monitor and save various data of the driving car on normal roads, rather than test roads, through wireless communication. With this system, the data of driving cars and driving circumstances can be collected in real-time and control signals from the car and data of the detailed parts can be obtained through the OBD-II diagnosis port.

Figure 1 shows the configuration and operation concept of the real-time driving monitoring system, which comprises three parts. The first part is a vehicle information-collecting device, which obtains the driving data of cars and parts. The second part is a real-time wireless transmission terminal, which receives the data from the vehicle information-collecting device, transmits the data to the server, and sets monitoring items and devices according to the setting received from the server. The third part is a machine-to-machine (M2M) network platform for collecting the transmitted data and indicating through a domestic mobile communication network and an operation center for saving the driving data and monitoring the driving condition of the test vehicle.



**Figure 1.** Schematic of the real-time monitoring system.

The data collected from the vehicle are transmitted to the server of the operating institutes after encryption and compression in real-time, through an M2M mobile communication network. The transmitted data are systematically saved in each database of the project according to the vehicles and provide the driving data of cars and parts and analysis of circumstances for user needs.

### 2.2. Test Vehicle and Data Specification

The following are the requirements for selecting a test vehicle. First, among commercialized hybrid cars, a vehicle's reliability and safety should already be verified through various evaluations. This means that a certain period has passed after commercialization. Second, the vehicle should have a significant number of sales for further research. Finally, a vehicle that provided diagnosis data by OBD-II was selected. Table 1 shows the specifications of a selected Toyota Prius III hybrid car, which is a five-seater passenger car of full hybrid drive type with a 60-kW electric motor and a 1.8-L gasoline engine.

Among the data collected from the test vehicles, only the diagnosis data provided by OBD-II, without separate sensors or DAQ have been utilized. In this study, only the data of physical parameters, such as speed, torque, temperature, voltage, and electric current, as well as some control target values, were significant. Thus, simple status signals, such as control signals and On/Off, were excluded. Additionally, information data, such as error codes for vehicle diagnosis and maintenance, were eliminated.

**Table 1.** Specifications of the test vehicle.

Item	Value	Unit and Type
Performance	Maximum speed	180 km/h
	Fuel efficiency	29.2 km/L
Driving Unit	Motor	60 kW
	Generator	42 kW
	Engine	73 kW
Motor	Maximum torque	207 Nm
Engine	Displacement	1798 cc
	Maximum torque	143 Nm
Battery	Type	Ni-MH
	Capacity	13 kWh

### 2.3. Driving Condition

While distribution by region is significant in a fleet test for over 100 vehicles, it is vital to set up a proper driving scenario for the one target hybrid car in this study. The driving scenarios were set up to reflect the hybrid effect through repetitive driving on highways, city roads, and combination zones. As there are differences per country between statistics such as average driving distance per day and parking time, we set up and realized a real experiment environment by achieving driving conditions that approximately met the Korean national statistics. The parking space was unified as outdoor parking. In each driving test per zone, the hybrid mode was set to the general Normal Mode, and the number of passengers was set to one identically. The test vehicle drove at a legal speed limit, and the traffic conditions, such as road congestion, were not considered apart.

To reflect seasonal effects, such as summer and winter, and weather effects, such as snow and rain, data have been collected over five years and for a total distance of 70,000 km. The average driving temperature was 32–36 °C in summer and −5–2 °C in winter. Table 2 listed the diagnosis data provided by OBD-II of a test vehicle.

**Table 2.** Definition of variables obtained from Electronic Control Unit.

Variables	Definition
Engine.Rev.Sensor	Internal combustion engine revolution (Engine RPM)
Motor.MG2.Revolution	Secondary motor-generator revolution (MG2 RPM)
Motor.Temp.No1	MG1 temperature
Auxiliary.Battery.Vol	Auxiliary battery voltage (14 V)
Engine.Coolant.Temp	Engine coolant temperature
Ambient.Temperature	Ambient temperature
Intake.Air.Temperature	Intake air temperature
Motor.MG2.Torq	Secondary motor-generator torque (MG2 Torque))
Generator.MG1.Rev	First motor-generator revolution (MG1 RPM)
Generator.MG1.Torq	First motor-generator torque (MG1 Torque)
Inverter.Temp.MG1	Inverter temperature for MG1
Inverter.Temp.MG2	Inverter temperature for MG2
Motor.Temp.No2	MG2 temperature
State.of.Charge.All.Bat	State of charge (SOC) of all battery cells
Power.Resource.VB	HV battery voltage
Power.Resource.IB	HV battery current
Auxiliary.Batt.Temperature	Auxiliary battery temperature

Figure 2 shows the main structure of the chosen hybrid electric test vehicle, where the internal combustion engine and two motors generate optimal driving torque and regenerative braking torque according to the driving situation. MG1 is a generator connected to the engine and supports MG2 through an inverter. MG2 uses the energy of the HEV battery to drive the vehicle. The engine and

MG1/MG2 motors perform an optimized hybrid driving strategy by controlling the energy flow in conjunction with the compound gear unit.

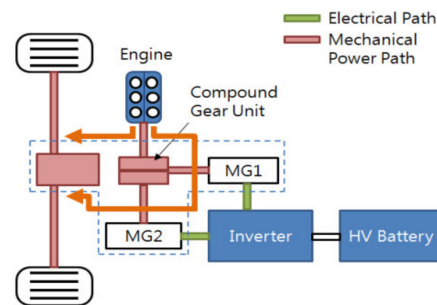


Figure 2. Schematic of hybrid electric power system.

### 3. Methodology

Here, it has been studied whether it is possible to determine the mileage interval of a HEV, depending on the fact that the driving data pattern varies according to the cumulative mileages. Through this study, a vehicle's parts whose statistical properties change drastically according to the mileage intervals can be found and a mileage interval where the characteristics of parts change considerably can be discovered. An abnormal status of a HEV can be predicted based on these. Alternatively, these can be used as a basis for comparative analysis.

#### 3.1. Data Exploration

To determine the change in the value for the main parts according to the mileage, we analyzed the characteristics of the auxiliary battery per cumulative mileage. As shown in Figure 3, as the mileage increased, the average voltage level decreased, and the width of the voltage fluctuation increased. In addition, the battery temperature per hour increased as well.

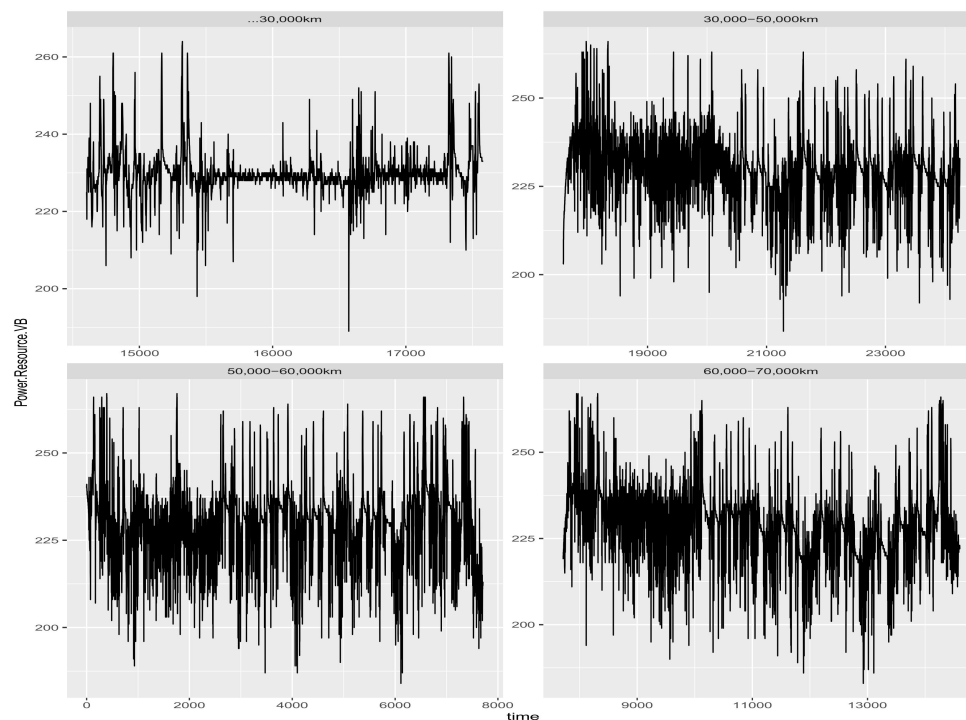


Figure 3. The time-series plot of power resource voltage of battery according to mileage intervals (1:  $\leq 30,000$ , 2: 30,000–50,000, 3: 50,000–60,000, 4: 60,000–70,000 km).

Thus, it has been found that the parts or modules of the hybrid car had different characteristic values per mileage. According to the mileage, the average value could vary, and the variability could be increased. Based on these observations, we derived features that can deduct the difference in part characteristics per mileage.

### 3.2. Features

The characteristics of data distribution are measured using factors such as centrality, variability, and normality. Centrality is evaluated from the mean, median, and mode. Variability of the data is measured from the maximum, minimum, mean, range, and variance. For determining the normality, skewness and kurtosis are used to measure whether the data follow a normal distribution. Because the aims of this study are not a prediction of distribution or statistical verification, the normality is excluded. We derive diverse feature in terms of the centrality and variability. A mean was used to measure the centrality and minimum, maximum, and standard deviation for variability. Because the vehicle parts data can be randomly reported within a specific range, the mean is more significant than a particular value with high frequency. As the range was derived from the maximum and minimum, and the variance was derived from the standard deviation, these two variables were excluded to reduce the dependency between the variables.

After the segmentation of data collected in seconds, the variability of the vehicle part characteristics has been measured within a specific interval. For obtaining the features of the data, the interval was set to 10 min. As it depends on the made by the decision maker, we set 10 min to consider the length of intervals and the quantity of data. If we set a long interval, the variability of data can be measured for a sufficient period. Instead, because only particular driving data were used, the data collected within an interval could be reduced. For example, when driving data were collected for one hour, features of six intervals were derived by a 10-min interval.

### 3.3. Classifier

In this study, an algorithm was built to derive important features to determine the mileage of a hybrid, using the random forest algorithm. This algorithm is a type of ensemble learning method and is a model extended from a decision tree model, which builds an assemble tree with sampling data and a selected feature set [18]. Each tree, with different datasets and feature sets, increases the prediction accuracy of the detection model. The random forest algorithm was adopted because it can avoid overfitting and generate feature importance from built decision trees. The random forest splits the dataset and then, builds a model from selected data and tests the model against the unselected data while generating a model. The random forest algorithm has a testing process with data unused in training and provides out-of-bag errors. This make the random forest algorithm to have a chance to avoid overfitting in the training process that the model is fine fitted to the train samples, but is not fitted to the test samples.

The importance of features is measured in terms of the Gini index. The Gini index is calculated by subtracting the sum of squared probabilities of each class, in our case, the mileage interval, from one. The Gini index gives the highest value when all classes have the same probabilities, while it gives the low value when a certain probability dominates other probabilities. The algorithm split the dataset along the attribute of a certain feature, calculated the probability of classes within each attribute, and generated the summation of the probabilities. The variable that can split data into classes using the attribute level gives the high Gini index.

$$Gini\ index = 1 - \sum p(x = k)^2 \quad (1)$$

where  $x$  is a feature and  $k$  is an attribute of  $x$ .



## 4. Experimental Results

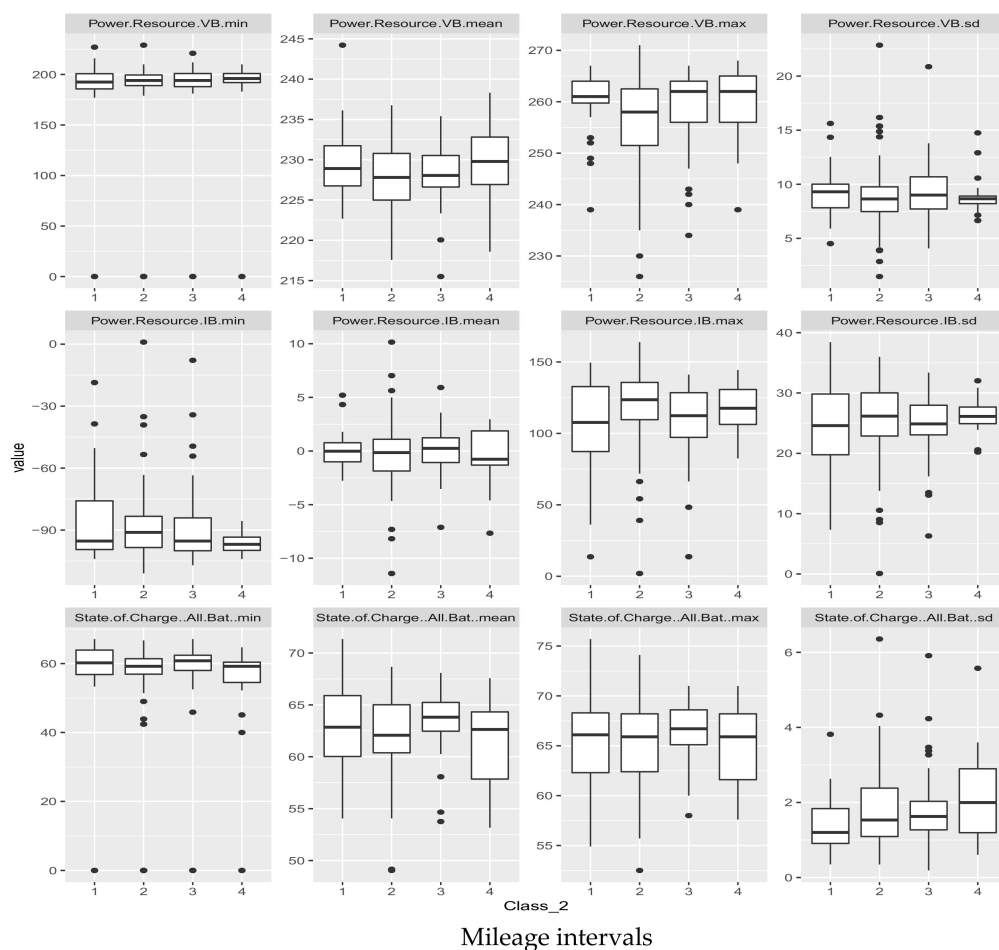
### 4.1. Feature Extraction

Table 3 shows the data collected for each mileage interval after transforming the data in seconds into a 10-min interval.

**Table 3.** Specifications of data collected by mileage intervals.

	Mileage Criteria (km)	Number of Data Points (Number of Driving Test × Hours × Data Points Per Hour)	Number of Driving Tests
Interval 1	≤30,000	32	3
Interval 2	30,000–50,000	91	7
Interval 3	50,000–60,000	41	6
Interval 4	60,000–70,000	21	2

We visualized the characteristics of the statistical values for voltage (VB), current (IB), temperature (Temp. of Batt.), and state of charge (SOC) of the battery, which are the main parts of a hybrid car. Figure 4 shows the distributions of statistics per mileage interval. As the mileage increases, the average and maximum of the statistics converge to a particular value, while the deviation increases.



**Figure 4.** Boxplot of VB, IB, SOC statistics according to mileage intervals.

In agreement with the results of the analysis, the statistics in this study had a different distribution according to the mileage interval, and thus, were suitable variables for checking the changes in the

part characteristics per mileage interval. Using data from 34 parts, we retrieved 136 features, including mean, standard deviation, maximum, and minimum, as the learning data.

#### 4.2. Algorithm Performance

The random forest provides the out-of-bag errors by testing the built model from selected data against the unselected data, while generating a model. The results of the learning model are shown in Tables 4–6. The evaluation of the learning model used the accuracy evaluation index.

**Table 4.** Performance metrics of mileage interval detection (30,000 vs. 50,000 km).

Prediction	Actual		Class Error	Accuracy
	30,000 km	50,000 km		
30,000 km	24	8	0.25	92.68%
50,000 km	1	90	0.01	

**Table 5.** Performance metrics of mileage interval detection (50,000 vs. 60,000 km).

Prediction	Actual		Class Error	Accuracy
	50,000 km	60,000 km		
50,000 km	88	3	0.03	82.58%
60,000 km	20	21	0.49	

**Table 6.** Performance metrics of mileage interval detection (60,000 vs. 70,000 km).

Prediction	Actual		Class Error	Accuracy
	60,000 km	70,000 km		
60,000 km	36	5	0.12	80.65%
70,000 km	7	14	0.33	

We first set up the mileage interval into four levels, such as 30,000, 50,000, 60,000, and 70,000, and then, set up the binary classification problem for each pair of two consecutive levels. The learning results are shown in Tables 4–6. Since, in this study, we experimented with driving until 71,000 mileage, we must collect additional driving data on the mileage, and find an interval where the characteristics of a part show differences.

Overall, the algorithm exhibited degraded performance in detecting the 60,000 km class. For three pairs of checkpoints, the algorithm exhibited good performance; however, as the mileage increased, the accuracy rather decreased. These results indicate that the characteristics of car parts change at the checkpoints.

#### 4.3. Feature Analysis

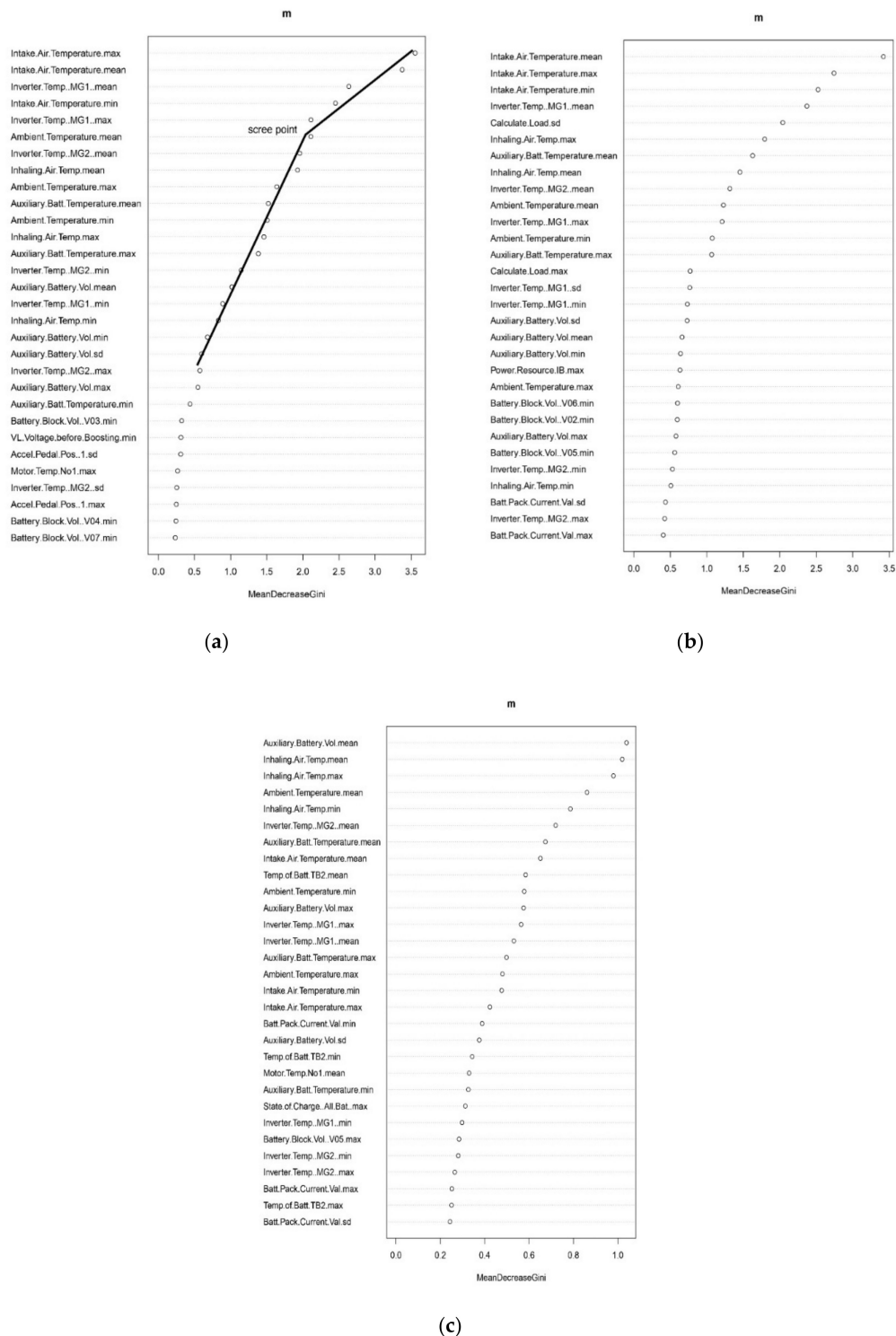
After building an algorithm that classified the mileage by the characteristic values of hybrid car parts, important variables that were used for the construction of the algorithm and the relation between them has been examined. The statistical values of the critical variables change clearly according to the mileage, and thus, should be checked per mileage interval.

##### 4.3.1. Feature Importance

The following shows the importance of the variables that worked mainly in differentiating the mileage interval. In the random forest algorithm, the importance of variables is derived by *incNodePurity*, which is measured by an impurity of nodes that is reduced after branching a tree by a specific variable. The impurity is calculated by the summation of the residual products of each node. The importance of the variables is visualized in Figure 5. The first figure represents the feature



importance in recognizing the change of mileage interval from 30,000 to 50,000 km. In this case, ‘intake air temperature mean’ has the highest value in importance. The scree point is found when the slope of importance decrease has changed as features are arranged according to the Gini index (the intercept point between the orange and the blue lines in Figure 5a). Key features from the highest Gini index to the point of this scree point are selected.



**Figure 5.** Ten minute basis feature importance of vehicle parts according to mileage interval changes: (a) 30,000 vs. 50,000 km; (b) 50,000 vs. 60,000 km; (c) 60,000 vs. 70,000 km.

Table 7 summarizes the important variables with high Gini index before the scree point. The results indicate that the characteristics of vehicle parts that significantly change depend on the mileage interval. When the mileage interval shifts from 30,000 to 50,000 km, the intake air temperature and inverter temperature for MG1 and ambient temperature, are mainly changed. For the shift of mileage interval from 50,000 to 60,000 km, ambient temperature no longer exhibits differences, but the intake air temperature and inverter temperature for MG1 show big differences between two intervals. For the shift from 60,000 to 70,000 km, the changes in the auxiliary battery voltage and the inhaling air temperature are newly observed.

**Table 7.** Important features per mileage interval change.

30,000 vs. 50,000 km	50,000 vs. 60,000 km	60,000 vs. 70,000 km
Intake.Air.Temp.mean	Intake.Air.Temp.mean	Auxillary.Battery.Vol.mean
Intake.Air.Temp.max	Intake.Air.Temp.max	Inhaling.Air.Temp.max
Inverter.Temp.MG1.mean	Intake.Air.Temp.min	Inhaling.Air.Temp.mean
Intake.Air.Temp.min	Inverter.Temp.MG1.mean	Ambient.Temp.mean
Ambient.Temp.mean	Calculated.load.max	Inverter.Temp.MG1.mean
Temp.MG1.max		Inhaling.Air.Temp.min

The analysis results show that the major temperature-related variables are closely related to the mileage interval. This is because the change in mileage interval reflects the characteristics of the driving environment, such as the atmospheric temperature according to the season. This is a self-evident result in real driving situations.

At the same time, it also clearly shows that the temperature change characteristics of the system and module by the thermal management strategy are linked to mileage. This is because the stresses of automotive parts and modules are related to temperature and the thermal management system controls the temperature in response. Since the voltage of the battery is affected by aging and thermal management of the battery, it is obvious that the battery voltage is also affected.

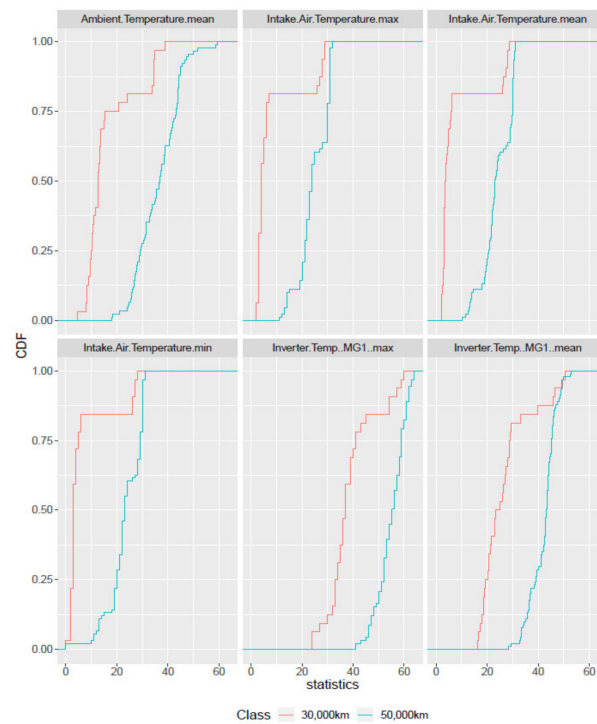
#### 4.3.2. Change in Major Part Characteristics

The cumulative distribution function (CDF) of major variables that exhibit a high Gini index so that the major changes are found within the intervals of 30,000–50,000, 50,000–60,000, and 60,000–70,000 km is displayed. The CDF plot illustrates how the values of a variable compose the entire dataset by displaying the cumulative density probability. Through the CDF plots, it can be observed how the features form the distribution. For example, the mean of the auxiliary battery voltages (denoted as Auxiliary.Battery.Vol.mean) has various values for 30,000 km, but a limited value near 14.4 for 50,000 km, which is higher than the value for 30,000 km. When comparing 30,000 and 50,000 km, major features exhibit much different distributions. The intake air temperature, ambient temperature, and inverter temperature show high skewness for 30,000 km, but the range of values increases for 50,000 km, as shown in Figure 6a.

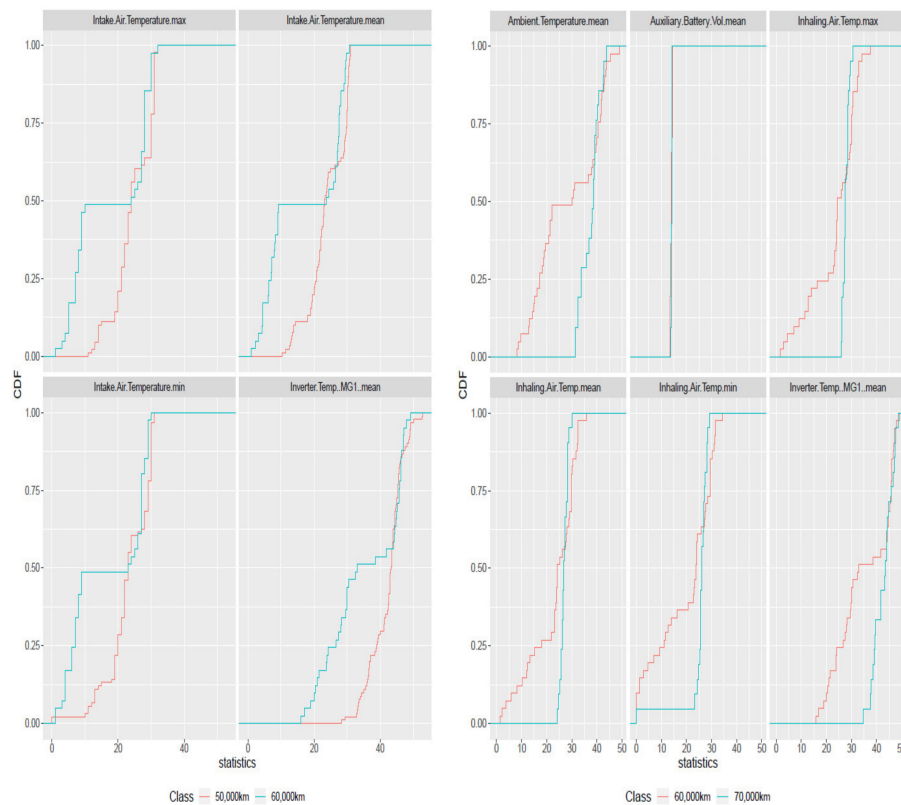
Figure 6b,c show the CDF plots of important variables that have been changed significantly in the 50,000–60,000 km and 60,000–70,000 km mileage intervals.

In this interval, a change in several parts occurred compared to the other intervals. The most important variables included auxiliary battery voltage and temperature, engine intake temperature, motor inverter temperature, and battery intake temperature.

Considering the principal parts whose characteristics vary according to the mileage interval, changes in the battery intake temperature and auxiliary battery temperature, followed by the voltage and temperature of the battery pack can be found.



(a)



(b)

(c)

**Figure 6.** CDF plot of major features of vehicle parts showing many differences: (a) from 30,000 to 50,000 km; (b) from 50,000 to 60,000 km; (c) from 60,000 to 70,000 km.

## 5. Conclusions

To trace major changes in the parts of a hybrid car with increasing mileage, which has not been investigated enough, we collected real driving data over 70,000 km with various paths under various conditions. We collected data provided by OBD-II injected into CAN. Among the collected data, we selected significant data of physical parameters and analyzed them. After aggregating CAN data collected in seconds into those in ten minutes, we measured centrality and variability, and verified whether these statistic features vary according to the mileage interval of a hybrid car.

We set up the checkpoints as 30,000, 50,000, 60,000, and 70,000 km, and performed a pair-wise comparison using a machine learning algorithm. The statistical properties are classified by the mileage interval with accuracy of 92.68%, 82.58%, and 80.65%, respectively. The high accuracy means that the correlation with the mileage section can be estimated by analyzing the driving data for a certain time (10 min) of the hybrid vehicle.

In addition, we found that the statistics of the data per part do not increase or decrease consistently in a mileage interval. Contrary to the fact that the mean and maximum of part data converge to a specific value, the deviation seems to increase. For the battery, which is the main part of a hybrid car, it is found that the voltage (VB), current (IB), temperature (Temp of Batt), and amount of charge (state of charge) of the battery decrease at the mileage interval from 30,000 to 50,000 and increase again at that over 50,000 km. However, as the mileage increases, the deviation in values increases further and gives a low accuracy in classifying the mileage interval.

By utilizing these analysis methods and the results, it can be used for maintenance strategies according to the prediction of aging of vehicles or parts. In addition, it is expected to be used for stress factor analysis and PHM to improve reliability through mileage accumulation and correlation analysis with major physical variables.

In this study, we performed a fundamental analysis through the analysis of data from a vehicle. To develop various service models based on the exact changes in car parts according to the driving mileage, the data analysis should be extended to multiple vehicles so that a reference model is developed based on that data. In the future, we will expand the analysis against other types of hybrid vehicles such as plug-in hybrid. With continuous research on a reference model, it is expected to be used for a customized maintenance per vehicle according to the driving mileages.

**Author Contributions:** Conceptualization, J.W. and I.Y.; methodology, J.W.; software, J.W.; resources, I.Y.; data curation, I.Y.; writing—original draft preparation, J.W.; writing—review and editing, C.P. and I.Y.; visualization, J.W.; supervision, I.Y.; project administration, I.Y.; funding acquisition, I.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2018R1D1A3B07048580) and also supported by the Soonchunhyang University Research Fund.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Advanced Vehicle Testing Activity. Available online: <https://avt.inl.gov/> (accessed on 19 July 2019).
2. Hydrogen Program. Available online: <https://www.hydrogen.energy.gov/> (accessed on 15 August 2019).
3. HyFLEET:CUTE Project. Available online: <https://cordis.europa.eu/project/id/19991> (accessed on 10 July 2019).
4. Rezvanizani, S.M.; Liu, Z.; Chen, Y.; Lee, J. Review and recent advances in battery health monitoring and prognostics technologies for electric vehicle (EV) safety and mobility. *J. Power Sources* **2014**, *256*, 110–124. [CrossRef]
5. You, G.; Park, S.; Oh, D. Real-time state-of-health estimation for electric vehicle batteries: A data-driven approach. *Appl. Energy* **2016**, *176*, 92–103. [CrossRef]
6. Guo, P.; Cheng, Z.; Yang, L. A data-driven remaining capacity estimation approach for lithium-ion batteries based on charging health feature extraction. *J. Power Sources* **2019**, *412*, 442–450. [CrossRef]

7. Chen, Z.; Mi, C.C.; Fu, Y.; Xu, Z.; Gong, X. Online battery state of health estimation based on Genetic Algorithm for electric and hybrid vehicle applications. *J. Power Sources* **2013**, *240*, 184–192. [[CrossRef](#)]
8. Laayouj, N.; Jamouli, H. Remaining useful life prediction of lithium-ion battery degradation for a hybrid electric vehicle. *GARJETI* **2015**, *4*, 16–23.
9. Wakita, T.; Ozawa, K.; Miyajima, C.; Igarashi, K.; Katunobu, I.; Takeda, K.; Itakura, F. Driver identification using driving behavior signals. *IEICE Trans. Inf. Syst.* **2006**, *89*, 1188–1194. [[CrossRef](#)]
10. Miyajima, C.; Nishiwaki, Y.; Ozawa, K.; Wakita, T.; Itou, K.; Takeda, K.; Itakura, F. Driver modeling based on driving behavior and its evaluation in driver identification. *Proc. IEEE* **2007**, *95*, 427–437. [[CrossRef](#)]
11. Nishiwaki, Y.; Ozawa, K.; Wakita, T.; Miyajima, C.; Itou, K.; Takeda, K. Driver identification based on spectral analysis of driving behavioral signals. In *Advances for In-Vehicle and Mobile Systems*; Abut, H., Hansen, J.H.L., Takeda, K., Eds.; Springer: Boston, MA, USA, 2007.
12. Kwak, B.I.; Woo, J.Y.; Kim, H.K. Know your master: Driver profiling-based anti-theft method. In Proceedings of the 14th Annual Conference on Privacy, Security and Thrust, Auckland, New Zealand, 12–14 September 2016; Institute of Electrical and Electronics Engineers Inc.: Auckland, New Zealand, 2016; pp. 211–218.
13. Meng, X.; Lee, K.K.; Xu, Y. Human driving behavior recognition based on hidden markov models. In Proceedings of the IEEE International Conference on Robotics and Biomimetics 2006, Kunming, China, 17–20 December 2006; pp. 274–279.
14. Choi, S.; Kim, J.; Kwak, D.; Angkititrakul, P.; Hansen, J.H. Analysis and classification of driver behavior using in-vehicle can-bus information. In Proceedings of the Biennial Workshop on DSP for In-Vehicle and Mobile Systems, Istanbul, Turkey, 17–19 June 2007; pp. 17–19.
15. Wahab, A.; Quek, C.; Tan, C.K.; Takeda, K. Driving profile modeling and recognition based on soft computing approach. *IEEE Trans. Neural Netw.* **2009**, *20*, 563–582. [[CrossRef](#)] [[PubMed](#)]
16. Kedar-Dongarkar, G.; Das, M. Driver classification for optimization of energy usage in a vehicle. *Procedia Comput. Sci.* **2012**, *8*, 388–393. [[CrossRef](#)]
17. Enev, M.; Takakuwa, A.; Koscher, K.; Kohno, T. Automobile driver fingerprinting. *Proc. Priv. Enhancing Technol.* **2016**, *1*, 34–51. [[CrossRef](#)]
18. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).