IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 2, 2019, accepted April 20, 2019, date of publication April 30, 2019, date of current version May 20, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2914188

Hybrid VARMA and LSTM Method for Lithium-ion Battery State-of-Charge and Output Voltage Forecasting in Electric Motorcycle Applications

ANGELA C. CALIWAG, (Student Member, IEEE), AND WANSU LIM, (Member, IEEE)
Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi 39177, South Korea

Corresponding author: Wansu Lim (wansu.lim@kumoh.ac.kr)

This work was supported in part by the National Research Foundation of Korea (NRF) under Grant 2017R1C1B5016837, in part by the ITRC Program under Grant IITP-2019-2014-1-00639, and in part by the Global Excellent Technology Innovation Program under Grant 10063078.

ABSTRACT Electric vehicles (EVs) have gained attention owing to their effectiveness in reducing oil demands and gas emissions. Of the electric components of an EV, a battery is considered as the major bottleneck. Among the various types of battery, lithium-ion batteries are widely employed to power EVs. To ensure the safe application of batteries in EVs, monitoring and control are performed using state estimation. The state of a battery includes the state-of-charge (SoC), state-of-health (SoH), state-of-power (SoP), and state-of-life (SoL). The SoC of a battery is the remaining usable percentage of its capacity. This mainly depends on variations of the operating condition of the EV in which the battery is applied. The SoC of a battery is reflected by its output voltage. That is, the SoC is considered to be zero when the output voltage of a battery drops below a cut-off voltage. This study proposes an SoC and output voltage forecasting method using a hybrid of the vector autoregressive moving average (VARMA) and long short-term memory (LSTM). This approach aims to estimate and forecast the SoC and output voltage of a battery when an EV is driven under the CVS-40 drive cycle. Forecasting using the hybrid VARMA and LSTM method achieves a lower root-mean-square error (RMSE) than forecasting with only VARMA or LSTM individually.

INDEX TERMS Battery output voltage, lithium-ion battery, neural network, state-of-charge, VARMA.

I. INTRODUCTION

Electric vehicles (EV) have gained attention owing to their effectiveness in reducing oil demands and gas emissions. Research on EVs addresses aims such as reducing costs and greenhouse gas emissions, and increasing the energy conversion efficiency [1]–[4]. Among the electric components of an EV, the battery is considered the major bottleneck [5]. The battery supplies the power required by an electric motor to drive an EV at a certain speed. The power required by the electric motor increases as the speed of the EV increases [6].

The various types of utilized battery include lead-acid, nickel–cadmium (Ni-Cd), nickel–metal hydride (Ni-MH), and lithium-ion (Li-ion) batteries. Among the different types of battery, it is known that Li-ion batteries have a longer cycle life, higher energy efficiency, higher power density, broader range of temperature operation, faster charging capability, lower self-discharge rate, higher voltage efficiency, and lower memory effect [4], [7]–[9]. Owing to the advantages of Li-ion batteries over other types, they have been widely employed to power electric devices such as EVs, air vehicles, automated and remote-controlled systems, bio-implanted devices, medical instrumentations, and other power tools and machines [8], [10]–[12].

The state estimation of batteries is essential to ensure the safe application of batteries in EVs [5]. This includes estimating the states of batteries such as the state-of-charge (SoC), state-of-health (SoH), state-of-power (SoP), and state-of-life (SoL). This study focuses on a battery's SoC. The SoC of a battery is the remaining usable percentage of the capacity of a battery. Thus, a 100% SoC indicates that the maximum available capacity can be used, while a 0% SoC indicates that no more available capacity can be used [5].

In addition to providing information on a battery's usable capacity, the SoC also provides information on the reliability, efficiency, and safety of an EV [5]. However, a battery's

The associate editor coordinating the review of this manuscript and approving it for publication was Ahmad Elkhateb.

SoC cannot be directly measured [6]. Some researchers have dedicated effort to SoC estimation to obtain the SoC of a battery. SoC estimation methods are classified into five types: lookup table-based, ampere-hour integral, model-based estimation, data-driven, and data–model fusion methods [5]. Lookup table-based methods use the relationship between the open circuit voltage (OCV) and SoC of a battery. With a known OCV and OCV curve, the SoC of a battery can be estimated directly. The ampere-hour integral method employs the relationship between the measured output current and SoC of a battery. With a known initial SoC, output current from the initial time to a certain time, and coulomb efficiency, the current SoC of a battery can be accurately calculated. Model-based estimation methods combine a lookup table-based method with the ampere-hour integral method. This approach employs nonlinear state estimation algorithms and adaptive filters to estimate the SoC of a battery. Data-driven methods exploit the nonlinear relationship between the measurable parameters and SoC of a battery. The measurable parameters of a battery include the current, voltage, and temperature. Data–model fusion methods combine data-driven and model-based methods. In this approach, a data-driven method is employed to identify the system parameters needed to improve the SoC estimation accuracy in the model-based method [3], [5]–[7].

Hannan M.A. et al. presented a review of state-of-charge estimation algorithms and methods in the different literatures. The result of their study shows that the conventional methods, such as the lookup table-based, ampere-hour integral-based, and model-based methods, are easy to implement but are affected by the battery aging, operating temperature, and external disturbances. Their study highlights the data-driven method which has good precision and high accuracy but has poor robustness and expensive computational cost. Hannan M.A. et al. further divides the data-driven method into three according to the algorithm used: adaptive filter algorithms learning algorithm, and nonlinear observer. They concluded that first, the adaptive filter algorithms yield good precision and high efficiency in predicting a non-linear dynamic battery SoC. However, using this algorithm has poor robustness and computationally expensive. Among the adaptive algorithms, the extended Kalman filter (EKF) used in estimating the SoC (Jiang et al.) has the least average error (less than 1%). Then, they also elaborated that the learning algorithm has also a good performance in predicting a non-linear dynamic battery SoC. Its advantage over the adaptive filter algorithm is the robustness in considering the effect of aging, temperature and noise. However, learning algorithm is more computationally expensive which require large memory unit to store and train data. Among the learning algorithms, the genetic algorithm (GA) used in estimating the SoC (Zheng) has the least average error (less than 2%). Finally, according to Hannan M.A. et al., the non-linear observer has highest accuracy and the greatest robustness against the disturbances at low computational cost. However, it is the most complex algorithm. Among the nonlinear observer algorithms, the proportional-integral observer (PIO) used in estimating the SoC (Xu et al.) has the least average error (less than 1%).

Battery SoC estimation and forecasting helps to prevent deep discharging and deterioration of a battery, and ensure reliable operation of an EV. SoC forecasting could help reduce ''range anxiety,'' and provide assurance that a charging station can be reached before the electric power of the battery runs out [13]. ''Range anxiety'' is defined as the fear of losing power in the middle of a long-distance drive [13]–[15].

In an EV, the SoC of a battery depends mainly on variations in the operating conditions. The rate of charge consumption is affected by varying mechanical forces, road conditions, and traffic flow dynamics [1], [6], [14]. Aside from the rate of charge consumption, variations in the operating conditions also affect the output voltage of the battery [3], and in practice the output voltage reflects the SoC of the battery. The SoC of a battery is said to be zero when the output voltage of a battery drops below a cutoff voltage $V_{cut-off}$ [14].

Forecasting the output voltage of a battery could provide an estimate of the time at which the output voltage of the battery will drop below the cutoff voltage $V_{cut-off}$. In EV applications, the variation in the output voltage of the battery depends on the driving speed of the EV [6], and variations in the driving speed cause variations in the load current.

Forecasting approaches are categorized into three types: (i) physical, (ii) statistical, and (iii) artificial-intelligence. Statistical and artificial-intelligence approaches, forecasting can be performed without the need for knowledge of the underlying data generation process. A hybrid of statistical and artificial-intelligence approach has been utilized in different fields to improve the forecasting accuracy. For instance, Zhang G. demonstrated the effectiveness of using hybrid autoregressive integrated moving average (ARIMA) model and neural network (NN) in various fields by forecasting sunspot activity, number of lynxes trapped in a year and exchange rate between British pound and US dollar. Results of the study conducted by G. Zhang shows higher accuracy of the forecasting the sunspot activity, number of lynxes and exchange rates compared to ARIMA model and NN model. Another study on hybrid statistical and artificial-intelligence approaches is performed by Alencar D. B. et al. They proposed a hybrid seasonal autoregressive moving average (SARIMA) and NN to forecast the windspeed in Brazil. Compared with SARIMA model, hybrid SARIMA and wavelet model, and NN model; hybrid SARIMA and NN model yields higher forecasting accuracy. To the best of authors' knowledge, the forecasting of battery voltage response in an electric motorcycle is not explored in previous studies.

This study proposes an SoC and output voltage forecasting method, using a hybrid of the vector autoregressive moving average (VARMA) and long short-term memory (LSTM). To consider the variations and uncertainties in operating conditions, we use real data from an electric motorcycle (discussed further in Section II). We also consider the driving speed of an EV, to consider the dynamic state of the battery

under real driving conditions. Thus, in this study we aim to estimate and forecast the SoC and output voltage of a battery when an EV is driven under the CVS-40 drive cycle.

The contributions of this work are summarized as follows: (i) A hybrid forecasting approach based on VARMA and LSTM is proposed to forecast the battery output voltage and SoC; (2) a thorough analysis of the dynamic behavior of a Li-ion battery in an electric motorcycle at 0 and 25°C is performed using real data; and (3) an analysis of the effects of driving an electric motorcycle at various speeds on the battery output voltage is presented, as only explored by few previous studies to the best of the authors knowledge. This study is limited to considering the CVS-40 driving cycle in South Korea.

The remainder of this paper is organized as follows. Section II presents the dataset utilized in this work. Section III presents the proposed forecasting approach. Section IV presents and discusses the results. Finally, Section V summarizes the paper and states the conclusions.

II. ELECTRIC MOTORCYCLE DATASET AND CVS-40 DRIVING CYCLE

To reduce the complexity and scope of the study, we assume that a vehicle runs under the CVS-40 drive cycle of South Korea, as shown in Fig. 1. A driving cycle depicts a variation in the vehicle driving speed with time. It is used to simulate and evaluate the performance of vehicles, such as energy consumption and range. In this study, we utilized CVS-40 driving cycle produced by South Korea. CVS-40 drive cycle consists of three speed categories: low, medium, and high. A summary of the statistics for each speed category is presented in Table 1. Each CVS-40 cycle consists of six repetitions of consecutive low, medium, and high speeds, with rests in between.

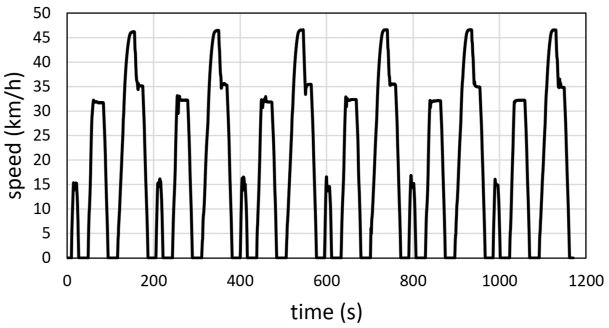


FIGURE 1. CVS-40 drive cycle of South Korea.

The first, second and third columns in table 1 contains the minimum (Min), mean, and maximum (Max) value of each speed category in all of the CVS-40 drive cycles at both 25°C and 0°C. This shows that the low speed varies from $14.95 - 18.28\ km/h$; the medium speed varies from $31.39 - 35.35\ km/h$; and the high speed varies from $36.36 - 46.61\ km/h$. These variations are reflected by the SD and Var in the fourth and fifth columns. The SD and Var values in

TABLE 1. Speed type summary statistics (units: km/h).

speed type	Min	Mean	Max	SD	Var
low speed	14.95	16.11	18.28	0.60	0.36
medium speed	31.39	32.51	35.35	0.76	0.58
high speed	36.36	43.32	46.61	4.82	23.23

table 1 shows low variations in the values at low and medium speed (SD of 0.60 and 0.76 for low and medium speeds respectively and Var of 0.36 and 0.58 for low and medium speeds respectively) but a high variation in the values at high speed (SD of 4.82 and Var of 23.23).

Data was collected while driving an electric motorcycle under the CVS-40 drive cycle in accordance to International Standard of battery-electric mopeds and motorcycles – performance (ISO13064-1). After completing each cycle, the electric motorcycle was placed at rest for 10 min. The cycles continued until the battery ran out of charge.

Two different operating temperatures were considered in this work: 0 and 25 °C. As illustrated in Fig. 2, the discharging time at 25 degrees is longer than that at 0 degrees. The dataset for 25 degrees consists of eight CVS-40 cycles until the battery was fully discharged, while that for 0 degrees consists of only six CVS-40 cycles before the battery was fully discharged. The summary statistics of the battery output voltage for each CVS-40 cycle at 0 and 25 °C are presented in Tables 2 and 3, respectively.

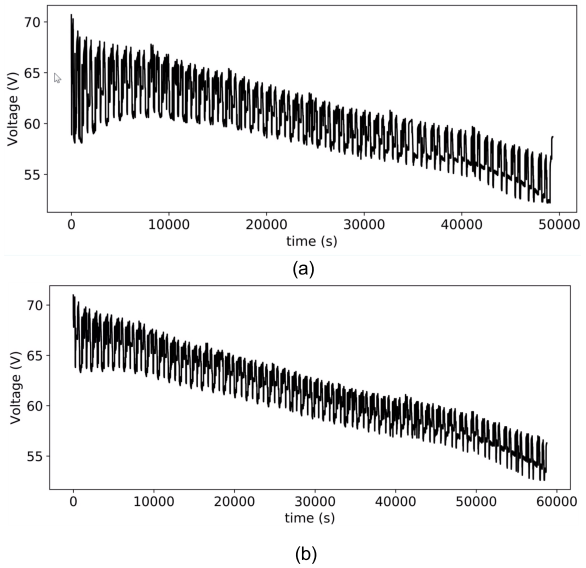


FIGURE 2. Battery output voltage for an electric motorcycle driven under the CVS-40 driving cycle, measured from fully charged to fully discharged at (a) 25 and (b) 0 °C.

In tables 2 and 3, the first, second and third columns contain the Min, mean, and Max values of the battery output voltage in each cycle at 25 and 0°C respectively. The minimum battery voltage corresponds to the voltage response of the battery when the EV is driven at the highest speed

TABLE 2. Battery output voltage (V) summary statistics: 25 °C.

Cycle number	Min	Mean	Max	SD	Var
First cycle	63.20	67.35	71.00	1.97	3.88
Second cycle	61.50	65.48	68.50	1.64	2.69
Third cycle	59.90	63.68	66.30	1.49	2.44
Fourth cycle	58.20	61.81	64.20	1.44	2.07
Fifth cycle	56.90	60.06	62.20	1.30	1.69
Sixth cycle	55.40	58.76	61.10	1.27	1.61
Seventh cycle	52.70	56.77	59.40	1.45	2.10
Eighth cycle	52.60	55.35	57.00	1.45	2.10

TABLE 3. Battery output voltage (V) summary statistics: 0 °C.

Cycle number	Min	Mean	Max	SD	Var
First cycle	58.10	64.88	70.70	3.12	9.73
Second cycle	59.70	64.08	67.80	2.21	4.88
Third cycle	57.90	62.06	65.40	1.92	3.69
Fourth cycle	56.30	59.94	63.00	1.72	2.96
Fifth cycle	54.80	58.41	61.30	1.59	2.53
Sixth cycle	52.20	56.24	59.70	1.97	3.88

(approximately $40 - 45\ km/h$) while the maximum voltage corresponds to the voltage level of the battery at the beginning of each drive cycle. Decreasing value of minimum voltage is affected by the decreasing value of the maximum voltage. That is, the voltage level of the battery decreases but it still has to supply the power need to drive the EV at the same driving cycle. The fourth and third column in tables 2 and 3 corresponds to the SD and Var of the battery voltage at 25 and 0°C. SD and Var measure of the deviations of the measured battery output voltage from the average battery output voltage. A decreasing SD and Var implies a decreasing voltage response of the battery to the same driving cycle used to drive the EV.

Aside from driving speed, the datasets for both 0 and 25 °C also include the battery output voltage, battery output current, motor input power, and motor output torque. The data contains measurements taken every 0.1 s. The electric motorcycle driving speed, battery output voltage, battery output current, motor input power, and motor output torque are considered as potential explanatory variables, which can be used to forecast the battery's output voltage, output current, and SoC. The summary statistics of the potential explanatory variables at 0 and 25 °C are presented in Tables 4 and 5, respectively.

III. BATTERY SOC AND OUTPUT VOLTAGE FORECASTING

Forecasting approaches are categorized into three types: (i) physical, (ii) statistical, and (iii) artificial-intelligence [16]–[18]. Physical approaches employ physical data to build a physical model. This requires a detailed description of the physical aspects of the system, such as the motor power, battery type and size, weight, and so on [13], [19] Statistical approaches employ historical data to build a statistical model [9], [17]. Meanwhile, artificial-intelligence approaches also employ features from historical data, to train an artificial neural network (ANN). ANN is an artificial

TABLE 4. Explanatory variable summary statistics: 25 °C.

Variable	Min	Mean	Max	SD	Var
Speed (km/h)	0	17.58	46.61	15.67	245.55
Voltage (V)	52.6	61.86	71.00	3.89	15.13
Current (A)	0.35	10.18	35.79	11.19	126.22
Power (W)	-20.00	616.17	2300.00	673.01	452942.46
Torque (N·m)	-4.90	10.02	22.41	7.87	61.94

TABLE 5. Explanatory variable summary statistics: 0 °C.

Variable	Min	Mean	Max	SD	Var
Speed (km/h)	0	16.93	43.52	15.08	227.41
Voltage (V)	52.20	60.93	70.67	3.74	13.99
Current (A)	-1.17	11.46	35.90	12.36	152.77
Power (W)	-80.00	675.32	2200.00	724.08	524291.85
Torque (N·m)	-3.17	9.86	20.89	7.44	55.35

intelligence modelling approach that imitates the information processing of the human brain [23]. In both statistical and artificial-intelligence approaches, forecasting can be performed without the need for knowledge of the underlying data generation process [23].

In this study, we propose a hybrid statistical and artificial-intelligence approach to forecast the battery output voltage and SoC. We combine VARMA, a statistical approach, with LSTM, an artificial-intelligent approach.

VARMA represents the application of the ARMA model to multivariate time series. It forecasts the next step in each time series using the past values of the same time series and other correlated time series [21], [22]. VARMA can represent other models, such as pure vector autoregressive (VAR), pure moving average (VMA), and combined VAR and VMA (VARMA) models. Similar to ARMA, VARMA is limited to the assumption that there is a linear correlation among the time series values. Thus, VARMA is only capable of capturing the linear patters in the data [23].

LSTM is a type of recurrent neural network (RNN), which is capable of learning long-term dependencies [8]. LSTM is a flexible computing framework, used to model nonlinear problems. Unlike VARMA, LSTM is capable of capturing the nonlinear patterns in the data. [23] However, it requires a huge amount of data to learn the nonlinearity, which could lead to overfitting. Thus, LSTM alone is not capable of handling both linear and nonlinear patterns concurrently.

To overcome the disadvantages of utilizing VARMA or LSTM independently, a hybrid VARMA and LSTM is adopted in this work. The battery output voltage and SoC forecasting is divided into three sections: (i) selection of the explanatory variables, (ii) battery output voltage forecasting, and (iii) battery SoC forecasting.

A. SELECTION OF EXPLANATORY VARIABLES

To account for the effects of variations in the operating conditions, five potential explanatory variables are considered in this study: the battery output voltage, battery output current,

motor speed, motor input power, and motor output torque. These variables can be obtained by direct measurement.

To determine which of the potential explanatory variables affect the battery output voltage and SoC forecasting, we compute the Pearson correlation between the potential explanatory variables with the battery output voltage and output current.

After selecting the explanatory variables, we also determine the effects of removing the explanatory variables with weaker relationships with the battery output voltage and output current, and three models are utilized. One considers all explanatory variables, both influential and non-influential. The second only considers the influential explanatory variables. The third only considers speed as the influential explanatory variable. The accuracies of both models are compared.

B. BATTERY OUTPUT VOLTAGE FORECASTING

In forecasting the battery output voltage, only the speed is considered as an explanatory variable.

In this study, we forecast the battery output voltage using a hybrid VARMA and LSTM model. The historical data of the battery output voltage is decomposed into its linear and nonlinear components. That is,

$$V_O(t) = V_{OL}(t) + V_{ON}(t) \tag{1}$$

where $V_O(t)$ is the battery output voltage time series, and $V_{OL}(t)$ and $V_{ON}(t)$ are its linear and nonlinear components, respectively.

The forecasted battery output voltage time series is then expressed as

$$\hat{V}_O(t) = \hat{V}_{OL}(t) + \hat{V}_{ON}(t) \tag{2}$$

where $\hat{V}_{OL}(t)$ and $\hat{V}_{ON}(t)$ are the forecasted linear and nonlinear components of the battery output voltage time series, respectively.

First, VARMA is utilized to forecast the linear component of the battery output voltage $\hat{V}_{OL}(t)$. Then, the nonlinear component $\hat{V}_{ON}(t)$ is obtained by subtracting the forecasted linear component of the battery output voltage from the original time series. That is,

$$V_{ON}(t) = V_O(t) - \hat{V}_{OL}(t) + e_L(t) \tag{3}$$

where $e_L(t)$ is the error between the actual and predicted linear components of the batter output voltage time series. $V_{ON}(t)$ also denotes the residual of the linear model, or the error when only VARMA is used in forecasting the battery output voltage.

Then, LSTM is employed to forecast the nonlinear component of the battery output voltage $\hat{V}_{ON}(t)$. In other words, LSTM is used to forecast the residual of the linear model, or the error when only VARMA is used in forecasting the battery output voltage.

Finally, the forecasted battery output voltage time series $\hat{V}_O(t)$ is obtained using (1).

C. BATTERY SOC FORECASTING

In forecasting the battery SoC, both the battery output voltage and speed are considered as explanatory variables.

We first forecast the battery output current in response to the motor speed, with the assumption that the EV is driven under CVS-40. Then, we compute the charge consumed by the EV at each point in time, and the total charge consumed by the EV from a fully charged to fully discharged state. Finally, the SoC of the battery is computed as

$$SoC(t) = 1 - \frac{Q(t)}{Q_{all}} \tag{4}$$

where $Q(t)$ is the charge consumed by the EV at each point in time, and Q_{all} is the total charge consumed by the EV from the fully charged to fully discharged state.

Similar to the battery output voltage forecasting, we forecast the battery SoC using the hybrid VARMA and LSTM method. The historical data of the battery output current is decomposed into its linear and nonlinear components. That is,

$$I_O(t) = I_{OL}(t) + I_{ON}(t) \tag{5}$$

where $I_O(t)$ is the battery output current time series, and $I_{OL}(t)$ and $I_{ON}(t)$ are its linear and nonlinear components, respectively.

The forecasted battery output current time series is then expressed as

$$\hat{I}_O(t) = \hat{I}_{OL}(t) + \hat{I}_{ON}(t) \tag{6}$$

where $\hat{I}_{OL}(t)$ and $\hat{I}_{ON}(t)$ are the forecasted linear and nonlinear components of the battery output current time series, respectively.

First, VARMA is utilized to forecast the linear component of the battery output current $\hat{I}_{OL}(t)$. Then, the nonlinear component of the battery output current $\hat{I}_{ON}(t)$ is obtained by subtracting the forecasted linear component of the battery output current from the original time series. That is,

$$I_{ON}(t) = I_O(t) - \hat{I}_{OL}(t) + e_L(t) \tag{7}$$

where $e_L(t)$ is the error between the actual and predicted linear components of the batter output voltage time series. $I_{ON}(t)$ also denotes the residual of the linear model, or the error when only VARMA is utilized in forecasting the battery output voltage.

Then, LSTM is employed to forecast the nonlinear component of the battery output current $\hat{I}_{ON}(t)$. In other words, LSTM is used to forecast the residual of the linear model, or the error when only VARMA is used in forecasting the battery output current.

Finally, the forecasted battery output current time series $\hat{I}_O(t)$ is obtained using (2).

After obtaining the forecasted battery output current time series $\hat{I}_O(t)$, the SoC is obtained using (4). The charge consumed by the EV at each point in time, $\mathrm{Q}(t)$, is obtained as

$$\mathrm{Q}(t) = \int_0^t \hat{I}_O(t)\, dt \tag{8}$$

where t is the current time. The charge consumed by the EV from the fully charged to fully discharged state, Q_{all}, is obtained as

$$Q_{all} = \int_0^T \hat{I}_O(t)\, dt \tag{9}$$

where T is the time at which the battery reaches its fully discharged state. The values obtained in (8) and (9) are used to compute the SoC using (4).

IV. RESULTS

A. SELECTION OF EXPLANATORY VARIABLES

The Pearson correlation between the five potential explanatory variables at 25 and 0 °C is presented in Tables 6 and 7, respectively. The explanatory variables at both 25 and 0 °C exhibit weak correlations between the current, speed, power, and torque with the voltage (−0.33, −0.20, −0.26, and −0.24), and strong correlations between the speed, power, and torque with the current (0.64, 1, and 0.78). This may be due to the fact that the voltage level of the battery decreases over time while the load it has to supply remains constant. In obtaining the correlations, the behavior of the variables when the battery is at fully-charged state to fully-discharged state is considered.

TABLE 6. Correlation matrix of explanatory variables summary statistics: 25 °C

Variable	Voltage	Current	Speed	Power	Torque
Voltage	1				
Current	-0.33	1			
Speed	-0.20	0.64	1		
Power	-0.26	1	0.64	1	
Torque	-0.24	0.78	0.79	0.78	1

TABLE 7. Correlation matrix of explanatory variables summary statistics: 0 °C

Variable	Voltage	Current	Speed	Power	Torque
Voltage	1				
Current	-0.48	1			
Speed	-0.34	0.68	1		
Power	-0.42	1	0.68	1	
Torque	-0.35	0.80	0.81	0.81	1

The results shown in Tables 6 and 7 suggest that the explanatory variables have little effect on the output voltage of a battery, and that considering these variables may not efficiently increase the accuracy of the voltage forecasting. On the other hand, the results also suggest that the explanatory variables have a considerable effect on the output current of a battery, and that considering these variables may increase the SoC estimation and forecasting accuracy.

Among the explanatory variables, the speed has the lowest correlation with the output voltage of the battery. This contradicts our hypothesis that the driving speed of an electric motorcycle should affect the output voltage of the battery.

TABLE 8. Average RMSEs of three models.

Variable	Average RMSE at 25 °C	Average RMSE at 0 °C
Model 1	0.5827	1.0496
Model 2	0.6829	0.9774
Model 3	0.5454	0.8857

We hypothesize that the output voltage of the battery varies to supply the power required by the electric motor to drive the electric motorcycle at a certain speed. To further investigate this issue, we performed output voltage level forecasting using three models: (i) Model 1, developed using five explanatory variables; (ii) Model 2, developed using four explanatory variables; and (iii) Model 3, developed using only the electric motorcycle driving speed as an explanatory variable.

The battery output voltage time series forecasting for 100 steps ahead was performed for seven and five iterations using the data for 25 and 0 °C, respectively. The numbers of iterations were chosen based on the total number of CVS-40 driving cycles performed before the battery was fully discharged. On the first iteration, the models were trained using data from the first CVS-40 driving cycle of each temperature type. On the second iteration, the models were trained using the data from the first and second CVS-40 driving cycles of each temperature type, and so on. On the final iteration, the models were trained using the data from the second to the last CVS-40 driving cycle for each temperature type. The root-mean-square errors (RMSEs) obtained in forecasting the battery output voltage using the three models are depicted in Fig. 3. The x-axis denotes the number of CVS-40 cycles used as training data, and the y-axis denotes the RMSE for each forecast iteration.

From Fig. 3 (a), three observations can be drawn: (i) the RMSE of Model 2 is always higher than those of Models 1 and 3, except with one and four cycles; (ii) the RMSE of Model 3 is the lowest among the models, except with two and four cycles; and (iii) the difference between the RMSEs of Models 1 and 3 is small except with seven cycles, where a difference of over 0.2 is observed. These observations imply that the Model 3 is best. Model 3 has the lowest RMSE for most of the numbers of cycles, and the lowest computational cost, because only one explanatory variable is considered.

In Fig. 3 (b), the order of the models based on their RMSEs differs for each number of cycles. That is, for one cycle Model 1 exhibits the highest RMSE, while Model 3 exhibits the lowest; for two cycles Model 3 has the highest RMSE while Model 2 has the lowest; for three cycles, Model 1 has the highest RMSE while Model 3 has the lowest; and so on. These observations do not support the implications of Fig. 3 (a). To provide more points of comparison, we took the average RMSE, as shown in Table 8.

In Table 8, the average RMSE of Model 3 is the lowest for both 25 and 0 °C. The results in Table 8 support the implications of Fig. 3 (a). The ambiguous observation from

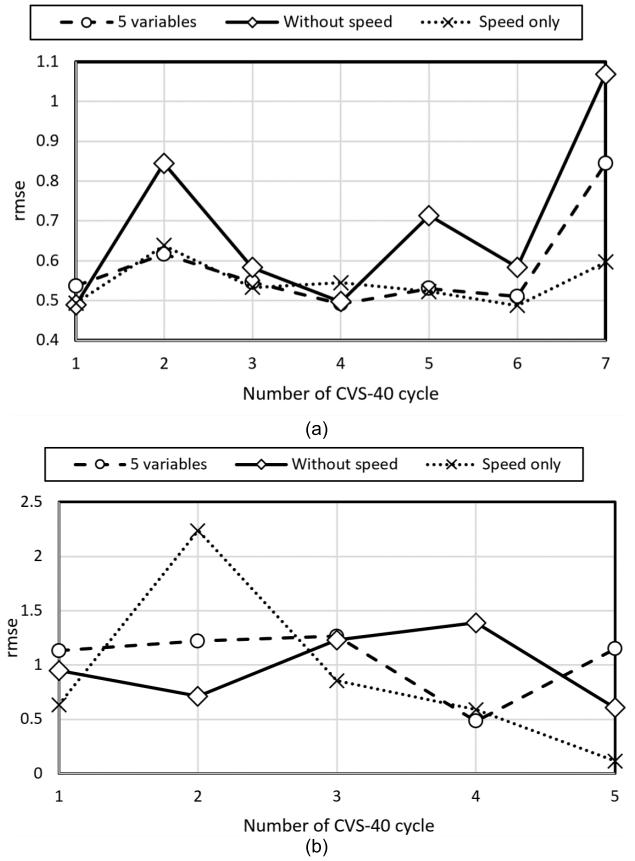


FIGURE 3. RMSE for predicting 100 steps ahead with varying numbers of CVS-40 cycles as training data at (a) 25 and (b) 0°C.

Fig. 3 (b) could be a result of the effect of driving the EV at a very low temperature.

B. BATTERY OUTPUT VOLTAGE FORECASTING

Based on the results shown in Table 8, only the speed is utilized as an explanatory variable for forecasting the battery output voltage.

We performed one cycle ahead forecasting for six and four iterations using the data for 25 and 0 °C, respectively. The results are presented in Tables 9 and 10 for the data at 25 and 0 °C, respectively.

Table 9 lists the RMSEs obtained in one cycle ahead battery output voltage forecasting using only VARMA, only LSTM, and hybrid VARMA and LSTM at 25 °C. The number of training samples differs in each iteration. That is, in the first iteration data measured during the first and second CVS-40 driving cycles are used as training data, in the second iteration data measured during the third cycle are additionally used for training, and so on. Thus, in the last iteration data measured from the first to the seventh CVS-40 driving cycles are used as training samples to predict the last cycle before the battery reaches the fully discharged state. At 25 °C, for dataset we utilized the electric motorcycle was driven for eight CVS-40 driving cycles before its battery reached a fully discharged state.

TABLE 9. One cycle ahead forecasting error (RMSE): 0 °C.

Training sample (cycles)	VARMA	LSTM	VARMA + LSTM
1st - 2nd	2.660	0.249	0.179
1st - 3rd	3.591	0.274	0.227
1st - 4th	4.128	0.357	0.143
1st - 5th	5.576	0.280	0.221
average	3.989	0.290	0.193

Overall, it is clear that the RMSE obtained using the hybrid VARMA and LSTM method is the lowest. In addition, it is notable that the RMSEs obtained using only LSTM and hybrid VARMA and LSTM are relatively close, while that obtained using only VARMA is far from the others. Looking more closely at Table 9, one can observe that the RMSE obtained using the hybrid VARMA and LSTM approach decreases with the number of training iterations. The only exception was for the last iteration, where the RMSE increased from 0.142 to 0.201. On the one hand, the decrease in the RMSE with the number of training iterations may be a result of the increase in the number of training samples at each iteration. On the other hand, the increase in the RMSE for the last iteration is due to the failure of the forecasting model to predict the erroneous behavior of the battery as it approaches a fully discharged state. Moreover, the RMSE obtained using only LSTM also decreases with the number of training iterations except at the last iteration, where the RMSE increased from 0.0.273 to 0.410. In contrast, the RMSE obtained using only VARMA increased with the number of training iterations except at the last iteration, where the RMSE decreased from 5.5081 to 2.2148. The increase in the RMSE with the number of training iterations may be a result of the failure of the forecasting model to learn and predict the nonlinear behavior of the battery as it approached a fully discharged state. On the other hand, further investigation of the cause of the decrease in the RMSE for the final iteration reveals that the battery reached a fully discharged state while driving the electric motorcycle approximately one third through the CVS-40 drive cycle. Thus, only one third of the CVS-40 drive cycle is forecasted by the VARMA-only approach in the last iteration, causing its RMSE decrease compared to the previous iteration.

TABLE 10. One cycle ahead forecasting error (RMSE): 25 °C.

Training sample (cycles)	VARMA	LSTM	VARMA + LSTM
1st - 2nd	2.869	0.904	0.159
1st - 3rd	3.692	0.319	0.158
1st - 4th	4.242	0.310	0.157
1st - 5th	4.489	0.203	0.149
1st - 6th	5.508	0.273	0.142
1st - 7th	2.215	0.410	0.201
average	3.836	0.4032	0.161

Table 10 presents the RMSEs obtained for one cycle ahead battery output voltage forecasting using only VARMA, only

LSTM, and the hybrid VARMA and LSTM method at 0 °C. Similar to the observations from Table 9, it is observed that the RMSE obtained using hybrid VARMA and LSTM is the lowest. Further, the RMSEs obtained using only LSTM and hybrid VARMA and LSTM are relatively close, while that obtained using only VARMA is far from the others. In comparing the results in Tables 9 and 10, one can observe that the trend of the RMSE in Table 9 differs from that in Table 10. The difference in the trend of the RMSE is assumed to be a result of the difference in the operating temperature. Operating the EV at a lower temperature tends to result in erratic behavior of the battery.

C. BATTERY SOC FORECASTING

For the battery output current forecasting, we performed multiple cycles ahead forecasting for six and four iterations with the data for 25 and 0 °C, respectively. Unlike for battery output voltage forecasting, where only one cycle ahead was forecasted, in the battery output current forecasting the remaining cycles before the battery reached its fully discharged state were forecasted. The results are presented in Tables 11 and 12. Table 11 presents the RMSEs obtained in multiple cycles ahead battery output current forecasting using only VARMA, only LSTM, and hybrid VARMA and LSTM at 25 °C. The number of training samples differs in each iteration. That is, in the first iteration data measured during the first and second CVS-40 driving cycles is used as training data, in the second iteration data measured during the third CVS-40 driving cycle is additionally used, and so on. Thus, in the last iteration data measured from the first to the seventh CVS-40 driving cycles are utilized as training samples to predict the last cycle before the battery reaches the fully discharged state. At 25 °C, for the dataset we utilized the electric motorcycle was driven for eight CVS-40 driving cycles before its battery reached a fully discharged state.

TABLE 11. Battery output current forecasting error (RMSE): 25 °C.

Training sample (cycles)	VARMA	LSTM	VARMA + LSTM
1st - 2nd	13.468	1.353	1.342
1st - 3rd	13.377	1.306	1.326
1st - 4th	12.944	1.293	1.287
1st - 5th	12.516	1.256	1.226
1st - 6th	11.918	1.184	1.110
1st - 7th	11.201	1.046	1.058
average	12.571	1.240	1.225

Overall, it is clear that the RMSE obtained using the hybrid VARMA and LSTM approach is the lowest. Further, the RMSEs obtained using only LSTM and hybrid VARMA and LSTM are relatively close, while that obtained using only VARMA is distinct from the others. Looking more closely at Table 11, one can observe that the RMSEs obtained using only VARMA, only LSTM, and hybrid VARMA and LSTM decrease with the number of training iterations. The number of training samples increases in each iteration, while the number of remaining cycles before the battery reaches its fully discharged state decreases. This results in a decreasing RMSE for each iteration.

In comparing Tables 9 and 11, it is clear that the RMSEs in Table 9 is lower than in Table 11. This may be a result of forecasting for different numbers of steps ahead. In Table 9, one cycle ahead forecasting was implemented. In Table 10, multiple cycle ahead forecasting was implemented. The training samples in both Tables 9 and 11 are the same in each iteration. Thus, a greater RMSE is expected for multiple cycle ahead forecasting than in one cycle ahead forecasting.

TABLE 12. Battery output current forecasting error (RMSE): 25 °C.

Training sample (cycles)	VARMA	LSTM	VARMA + LSTM
1st - 2nd	14.920	1.634	1.604
1st - 3rd	14.630	1.671	1.660
1st - 4th	13.986	1.494	1.495
1st - 5th	13.606	1.811	1.744
average	14.286	1.653	1.626

Table 12 presents the RMSEs obtained in multiple cycle ahead battery output current forecasting using only VARMA, only LSTM, and hybrid VARMA and LSTM at 0 °C. Similar to the observations in Table 11, it is observed that the RMSE obtained using hybrid VARMA and LSTM is the lowest. Further, the RMSEs obtained using only LSTM and hybrid VARMA and LSTM are relatively close, while the RMSE obtained using only VARMA is far from the others. In comparing the results in Tables 11 and 12, one can observe that the trend of the RMSE in Table 11 differs from that in Table 12. The difference in the trend of the RMSE is assumed to be a result of the difference in the operating temperature. Operating the EV at a lower temperature tends to lead to erratic behavior of the battery.

The small difference in the RMSE obtained using LSTM only and VARMA + LSTM model is due to the high RMSE in using VARMA in the latter model. Still, the result shows that LSTM is effective in improving the forecasting result of VARMA only model. Optimization in the VARMA model can be performed to increase its accuracy and reduce the RMSE obtained in using the hybrid VARMA and LSTM model. In addition to that, integration of VARMA to the LSTM model reduces the risk of overfitting wherein LSTM uses the value at time $t-1$ as a forecasted value at time t.

Based on Tables 9, 10, 11, and 12, it is clear that the RMSE using hybrid VARMA and LSTM is the lowest. As observed, LSTM is effective in compensating for the error in forecasting with only VARMA. On the other hand, VARMA is also effective in compensating for the error in forecasting with only LSTM, although the difference between forecasting with only LSTM and hybrid VARMA and LSTM is small.

Following the battery output current forecasting, the battery SoC was predicted using (4), (8), and (9). The predicted and actual battery SoCs are compared and depicted in Figs. 4 and 5 for 25 and 0 °C, respectively.

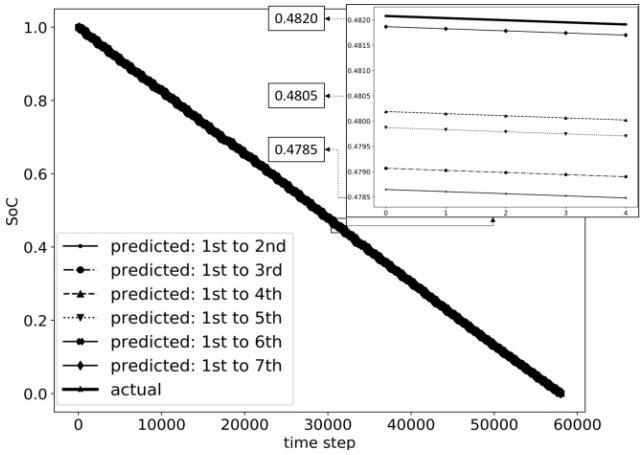


FIGURE 4. Predicted and actual battery SoCs at 25 °C.

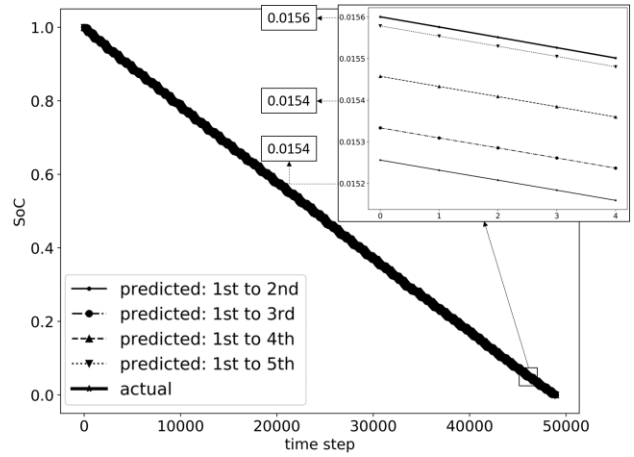


FIGURE 5. Predicted and actual battery SoCs at 0 °C.

Figure 4 shows the predicted and actual battery SoCs per time step at 25 °C, where an SoC of 1.0 corresponds to a 100% SoC, or a fully charged state, and an SoC of 0.0 corresponds to a 0% SoC, or a fully discharged state. As shown in Fig. 4, the SoC of the battery reduces from 1.0 to 0.0 after 60,000 time steps, where each timestep has a duration of 0.1 s. It is clear that the SoC decreases linearly over time. Looking more closely, the sub-figure of Fig. 4 clearly shows that the battery SoC predicted using the battery output current forecasted using the first to seventh drive cycles yields the closest prediction to the actual battery SoC. Moreover, the battery SoC predicted using the battery output current forecasted with the first to second drive cycles yields the furthest prediction to the actual battery SoC. In comparing the predicted values, it is clear that the distance from the actual value decreases as the number of training samples or the number of drive cycles increases.

Figure 5 shows the predicted and actual SoC per time step at 0 °C. Similar to Fig. 5, an SoC of 1.0 corresponds to a 100% SoC and an SoC of 0.0 corresponds to a 0% SoC. Unlike in Fig. 4, the SoC of the battery reduces from 1.0 to 0.0 after

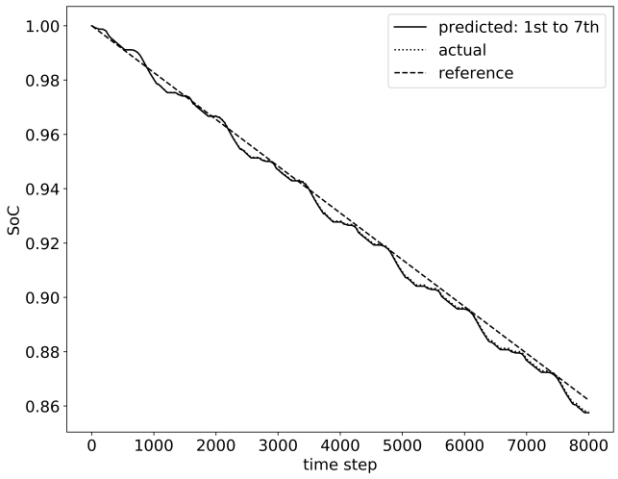


FIGURE 6. Predicted and actual battery SoCs at 25 °C, and a linear reference line between 0 and 8,000 timesteps.

50,000 time steps, where each timestep has a duration of 0.1 s. It is again clear that the SoC decreases linearly over time. However, as shown in Figs. 4 and 5, the battery at 0 °C reached 0% SoC earlier than that at 25 °C.

Figure 6 shows the predicted and actual battery SoCs at 25 °C, with a linear line for reference. This figure illustrates the effect of the variation in the EV driving speed on the SoC of the battery. A variation in the SoC is caused by the variation in the driving speed of the EV. Higher variations in driving speed cause higher variations in SoC. Overall, the trend is decreasing. However, if we look more closely the SoC curve deviates away from the linear reference line over time. Thus, the SoC is consumed more quickly over time.

V. CONCLUSION

In this paper, we presented real data gathered by driving an electric motorcycle under the CVS-40 driving cycle, described the CVS-40 driving cycle of South Korea, derived factors that affect the battery output voltage and SoC, proposed a hybrid forecasting approach based on VARMA and LSTM to forecast the battery output voltage and SoC, and presented an analysis of the effect of driving an EV at various speeds on the battery output voltage.

An electric motorcycle was driven under the CVS-40 driving cycle of South Korea at 25 and 0 °C. The results of the experiment show that the electric motorcycle's driving speed could be utilized as an explanatory variable in forecasting the battery output voltage and SoC. To forecast the battery output voltage and SoC, we employed a hybrid VARMA and LSTM method. The hybrid VARMA and LSTM approach was capable of capturing both the linear and nonlinear features of the battery voltage and battery SoC. First, the battery output voltage is forecasted at one cycle ahead. Forecasting using hybrid VARMA and LSTM achieved RMSEs of 0.161 and 0.193 for data at 25 and 0 °C, respectively. Then, the battery SoC was forecasted at multiple cycles ahead. Forecasting at

one cycle ahead achieved the lowest error. The error increased as the number of cycles to be forecasted increased. The results of the experiment demonstrate that the variation in the electric motorcycle speed causes a variation in the SoC of its battery.

REFERENCES

[1] C. Liu, K. Chau, D. Wu, and S. Gao, "Opportunities and Challenges of Vehicle-to-Home, Vehicle-to-Vehicle, and Vehicle-to-Grid Technologies," *IEEE J. Mag.*, vol. 101, no. 11, pp. 2409–2427, Nov. 2013. Accessed: Feb. 27, 2019. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6571224

[2] M. Vojtisek-Lom, "Anticipated effects of gradual replacement of internal combustion engines with electric drives on vehicle exhaust emissions in prague," *IEEE J. Mag.*, vol. 5, no. 4, pp. 136–145, Jul. 2013. Accessed: Feb. 27, 2019. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6646326

[3] F. Sun and R. Xiong, "A novel dual-scale cell state-of-charge estimation approach for series-connected battery pack used in electric vehicles," *J. Power Sour.*, vol. 274, pp. 582–594, Jan. 2015.

[4] M. A. Hannan, M. M. Hoque, A. Hussain, Y. Yusof, and P. J. Ker, "State-of-the-art and energy management system of lithium-ion batteries in electric vehicle applications: Issues and recommendations," *IEEE Access*, vol. 6, pp. 19362–19378, 2018.

[5] R. Xiong, J. Cao, Q. Yu, H. He, and F. Sun, "Critical review on the battery state of charge estimation methods for electric vehicles," *IEEE J. Mag.*, vol. 6, no. 8, pp. 1832–1843, Aug. 2017. Accessed: Feb. 25, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8168251

[6] M. Jafari, A. Gauchia, S. Zhao, K. Zhang, and L. Gauchia, "Electric vehicle battery cycle aging evaluation in real-world daily driving and vehicle-to-grid services," *IEEE J. Mag.*, vol. 4, no. 1, pp. 122–134, Mar. 2018. Accessed: Feb. 25, 2019. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8070984

[7] Y. Jiang, J. Jiang, C. Zhang, W. Zhang, Y. Gao, and N. Li, "State of health estimation of second-life $LiFePO_4$ batteries for energy storage applications," *J. Cleaner Prod.*, vol. 205, pp. 754–762, Dec. 2018.

[8] Y. Zhang, R. Xiong, H. He, and M. G. Pecht, "Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries," *IEEE Trans. Veh. Technol.*, vol. 67, no. 7, pp. 5695–5705, Jul. 2018.

[9] J. Larminie and J. Lowry, *Electric Vehicle Technlogy Explained*. New York, NY, USA: Wiley, 2003.

[10] D. Liu, J. Zhou, and Y. Peng, "Data-driven prognostics and remaining useful life estimation for lithium-ion battery: A review," *Instrumentation*, vol. 1, no. 1, pp. 59–70,2014.

[11] L. Ren, L. Zhao, S. Hong, S. Zhao, H. Wang, and L. Zhang, "Remaining useful life prediction for lithium-ion battery: A deep learning approach," *IEEE Access*, vol. 6, pp. 50587–50598, 2018.

[12] L. Zhang, Z. Mu, and C. Sun, "Remaining useful life prediction for lithium-ion batteries based on exponential model and particle filter," *IEEE Access*, vol. 6, pp. 17729–17740, 2018. Accessed: Feb. 25, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8318570

[13] C.-H. Lee and C.-H. Wu, "A novel big data modeling method for improving driving range estimation of EVs," *IEEE Access*, vol. 3, pp. 1980–1993, 2015.

[14] D. T. Lee, S. J. Shiah, C. M. Lee, and Y. C. Wang, "State-of-charge estimation for electric scooters by using learning mechanisms," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 544–556, Mar. 2007.

[15] E. Chemali, P. J. Kollmeyer, M. Preindl, R. Ahmed, and A. Emadi, "Long short-term memory networks for accurate state-of-charge estimation of Li-ion batteries," *IEEE Trans. Ind. Electron.*, vol. 65, no. 8, pp. 6730–6739, Aug. 2018.

[16] J. Ma, M. Yang, X. Han, and Z. Li, "Ultra-short-term wind generation forecast based on multivariate empirical dynamic modeling," *IEEE Trans. Ind. Appl.*, vol. 54, no. 2, pp. 1029–1038, Mar. 2017.

[17] J. Jung and R. P. Broadwater, "Current status and future advances for wind speed and power forecasting," *Renew. Sustain. Energy Rev.*, vol. 31, pp. 762–777, Mar. 2014.

[18] M. Ozkan and P. Karagoz, "A novel wind power forecast model: Statistical hybrid wind power forecast technique (SHWIP)," *IEEE Trans. Ind. Inform.*, vol. 11, no. 2, pp. 375–387, Apr. 2015.

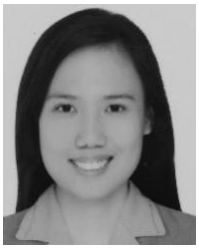
[19] Z. Asus, E.-H. Aglzim, D. Chrenko, Z.-H. C. Daud, and L. L. Moyne, "Dynamic modeling and driving cycle prediction for a racing series hybrid car," *IEEE J. Emerging Sel. Topics Power Electron.*, vol. 2, no. 3, pp. 541–551, Sep. 2014.

[20] D. B. Alencar, C. M. Affonso, R. C. L. Oliveira, and J. C. R. Filho, "Hybrid approach combining SARIMA and neural networks for multi-step ahead wind speed forecasting in Brazil," *IEEE Access*, vol. 6, pp. 55986–55994, 2018.

[21] M. Simionescu, "The use of Varma models in forecasting macroeconomic indicators," *Econ. Sociol.*, vol. 6, no. 2, pp. 94–102, 2013.

[22] W. W. S. Wei, *Time Series Analysis Univariate and Multivariate Methods*. Boston, NA, USA: Pearson Education, 2019.

[23] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, Jan. 2003.



ANGELA C. CALIWAG was born in Bulacan, Philippines, in 1996. She received the B.S. degree in electrical engineering from the Mapúa Institute of Technology, Manila, Philippines, in 2017. She is currently pursuing the M.S. degree in IT convergence engineering with the Kumoh National Institute of Technology, Gumi, South Korea, where she was a Research Assistant with the Future Communications Systems Laboratory, from 2017 to 2019. Her research interests include the analysis and improvement of battery management systems, signal processing, and time-series forecasting and sign language recognition with machine learning.

WANSU LIM received the Ph.D. degree from the Gwangju Institute of Science and Technology (GIST), South Korea, in 2010. From 2010 to 2014, he was a Research Fellow (2010–2013) of the University of Hertfordshire, U.K., and then a Postdoctoral Researcher (2013–2014) with the Institut National de la Recherche Scientifique (INRS), Canada. Since 2014, he has been an Assistant Professor with the Kumoh National Institute of Technology (KIT), South Korea. His research interests include statistical analysis, machine learning, and optimization.