2023 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)
June 28-30, 2023. Seattle, Washington, USA

An Efficient Hybrid Deep Learning Approach for Accurate Remaining EV Range Prediction

Magdy Abdullah Eissa, and Pingen Chen*

Abstract— **The accurate remaining Electric Vehicle (EV) range estimation is necessary to overcome EV users' range anxiety and infrastructure limitations. However, the traditional methods of EV Remaining Driving Range (RDR) estimation assumes the vehicle speed and energy consumption are consistent with the profiles in the recent history. But in the real world, the driving mode changes rather dynamically according to the user's speed profile, which significantly impacts RDR. Thus, the key question to be addressed in this work is how to accurately predict RDR considering the variation of the user speed profile during the driving trip. So, this work proposed a hybrid deep learning approach for accurate RDR estimation, where the future speed is then updated according to the average speed predicted in a 15-min prediction window. The deep learning approach combines a convolutional neural network (CNN) with Long Short-Term Memory (LSTM) to predict the remaining range of EVs based on historical EV speed data. The proposed CNN-LSTM-hybrid model is trained by exploiting the historical driving data of about 50 users in a two-week test-drive period. The test performance of the proposed EV range estimator is validated using real-world driving data that shows the high accuracy of RDR prediction with an average error of 3.762 km in a testing time window of 7.5 hours. The test results demonstrate the effectiveness of the proposed approach in the EV speed profile prediction, and thus RDR estimation with a high accuracy.**

I. Introduction

Global warming, poor air quality, and high dependence on fossil fuels are the primary factors driving the shift from conventional cars to electric vehicles (EVs). A study from Stanford University [1], found that, to mitigate these issues, 139 countries could achieve 100% clean and renewable energy by 2050, but it would require a significant, immediate overhaul of the world's energy infrastructure.

One of the major obstacles to the transition to EVs is the limited charging infrastructure, particularly in rural areas. Many people use cars daily, and these limitations can cause anxiety for EV users and hinder the shift from conventional vehicles to EVs [2] [3] [4]. In the short term, an accurate prediction of remaining EV range is necessary to alleviate range anxiety and overcome infrastructure limitations. In the long term, it can also prevent overloading of the distribution network during EV charging [4].

Various research groups are developing algorithms to address the problem of estimating remaining driving range (RDR) for EVs. These existing algorithms can be mainly classified into two categories: model-based approaches (e.g., [2]), and data-driven approaches (e.g., [3], and [4]). In [5], a hybrid model-based RDR estimation approach using a particle filter with Markov chains was reported. An improved RDR estimation framework based on a physics-based EV model was introduced in [2]. The authors in [6] developed an RDR estimation model that combines Kernel Principal Component feature parameters and a fuzzy C clustering algorithm for identifying and forecasting driving cycles. On the other hand, data-driven RDR estimation methods using extreme and light gradient boosting regression trees were proposed in [7]. In [3], the authors used real-world EV data to develop nonlinear RDR estimation models using a data-driven method. The study [4] proposed a clustering algorithm to classify driving patterns, and then developed a multi-mode RDR estimator based on predefined clusters.

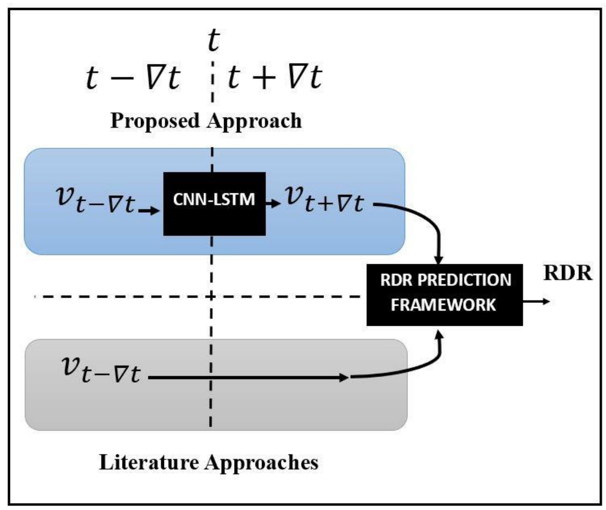

Figure 1 Schematic diagram of proposed RDR Prediction approach vs the literature approaches

In most of the existing literature, instead of predicting future driving patterns, it is assumed that the driving pattern will not change in the updated time window of RDR (∇t). This assumption may only be valid for very short time horizons (e.g., a few seconds) but can potentially result in significant prediction errors in longer time horizons, and thus lead to confusion in RDR for EV users. To balance EV user comfort (by providing a more accurate estimation in a reasonable time

*Research supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under the Vehicle Technologies Office (VTO), Award Number DE-EE0008888 and DE-EE0009862. The views expressed herein do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

Pingen Chen is with the Department of Mechanical Engineering at Tennessee Technological University, Cookeville, TN 38505 USA (phone: 931-372-3310; fax: 931-372-6340; e-mail: pchen@tntech.edu).
Magdy Abdullah Eissa is with the Department of Mechanical Engineering at Tennessee Technological University, Cookeville, TN 38505 USA.

horizon) with sensitivity to changes in driving patterns (i.e., higher RDR accuracy), this work proposes an efficient deep learning approach for estimating RDR. This approach updates the prediction using the predicted future value of EV speed within a 15-minute time horizon ($i.e., \nabla t = 15\, min$), as shown in Figure 1.

The proposed hybrid model combines a convolutional neural network (CNN) with Long Short-Term Memory (LSTM) to predict the remaining range of EVs based on historical data of EV speed. The main advantage of using a CNN is its ability to automatically extract features from time series data, while LSTM has been shown to be effective in predicting values in sequence-to-sequence series. At each time step, the CNN extracts key features from the sequence, while the LSTM is trained to predict the future value at the next time step. The training and parameters of the hybrid network were fine-tuned using Bayesian optimization.

The paper's contributions can be summarized as follows:

1. Existing literature assumes unchanged driving patterns in updated time windows, leading to prediction errors for longer time horizons and confusion for EV users.
2. The paper proposes an efficient deep-learning approach for estimating the remaining driving range (RDR) of EVs, considering both accuracy and sensitivity to changing driving patterns.
3. The approach combines CNN and LSTM to leverage the strengths of feature extraction and sequence-to-sequence prediction.
4. The CNN extracts key features from the historical EV speed data, while the LSTM predicts future values at each time step.
5. The hybrid model is fine-tuned using Bayesian optimization to optimize training and parameter settings.
6. By integrating CNN and LSTM, the proposed approach enhances RDR estimation accuracy within a 15-minute time horizon, providing EV users with reliable predictions for informed decision-making.

The remaining sections of this paper are organized as follows: Section II introduces real-world EV driving data. Section III describes the proposed EV RDR prediction framework. Section IV explains the proposed CNN-LSTM-based prediction approach. Section V presents the results and evaluation performance metrics of the RDR prediction, and Section VI provides the conclusions and future work directions.

II. DATA COLLECTION AND PRE-PROCESSING

A. Description of EV

In this study, a 2019 Nissan Leaf EV was used as the research platform. The vehicle is featured with a 40-kWh battery and an EPA-rated range of 149 miles per full charge. The vehicle was operated by different volunteers in an EV test-drive program in Upper Cumberland region in TN which covered 14 counties. Each participant operated the EV for about two weeks for their daily applications.

B. Data Collection

The real-world EV data used in this work were collected using the aforementioned Nissan Leaf SV. Key EV powertrain data were collected including battery state of charge (SOC), battery current and voltage, battery temperature, motor torque and speed, vehicle GPS location, ambient temperature, vehicle speed, accelerator pedal position, and others. The data logging was achieved using Hem Data's On-Board Diagnosis (OBD) Mini Logger, which was installed on the test vehicle, with a data sampling rate of 1 Hz. The second-by-second vehicle velocity profile and day-to-day battery SOC profile are illustrated in Figure 2. and Figure 3, respectively.

C. Data Pre-processing

Snapshot-based methods [8] organize time series data by converting it into a two-dimensional matrix, enabling structured analysis and modeling. This format captures temporal information along one axis and attributes along the other, facilitating pattern identification and interpretation. The input sequence size is determined by an n-window (15 minutes) and the LSTM is expected to receive an input of n-window cases to predict the next window, one step ahead. The data were normalized to improve the convergence process using a normalization algorithm. Then, the lag matrix function was used to create "look back" rolling windows, as shown in Figure 4.

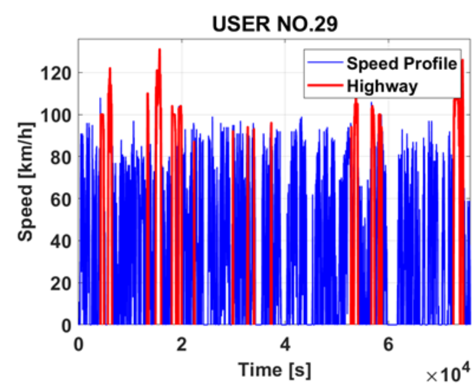

Figure 2: Vehicle velocity profile for user No.29. (Blue: vehicle speed profile. Red: highway speed.)

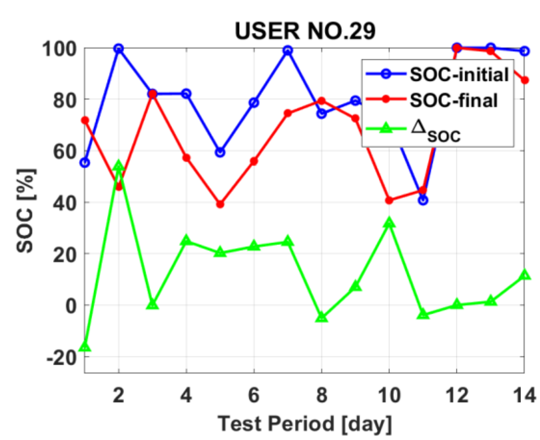

Figure 3:Day-to-day battery SOC profile for user No.29 (blue: initial SOC of the day; red: final SOC of the day; green: change of SOC)

III. RDR Prediction Framework

The EV RDR depends on two factors including the energy available in the battery pack and the energy consumption of the EV [4]. These two factors are described below in detail.

A. *Energy Available in the Battery Pack*

The remaining energy in the battery pack can be described in (1).

$$E_{battery} = \Delta \mathrm{SOC} \cdot \mathrm{C} \tag{1}$$

where $E_{battery}$ is the remaining energy (Wh) in the EV battery. ΔSOC denotes the difference between the current battery energy level (SOC) and the minimum SOC (SOC_{min}) at which the vehicle can still function appropriately. In this work, SOC_{min} is 4%. C is the energy capacity of the EV battery pack (Wh).

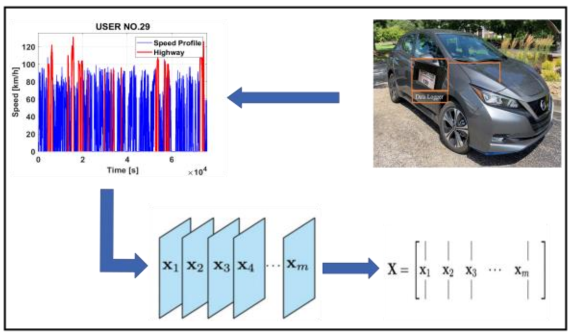

Figure 4:Data collection and preparation process.

B. *EV Energy Consumption (E_C)*

E_C is the average energy consumption value of the user, corresponding to the trip, as calculated in [2].

$$E_C = \int_0^{t_f} P_{elec} dt \ , \tag{2}$$

where, t_f denotes the travel time; P_{elec} is the instantaneous electric power drawn from or fed into the battery pack and is calculated in (3) [2].

$$P_{elec} = \begin{cases} \frac{P_d}{\eta_{prop}} & a \geq 0 \\ \eta_{regen} P_d & a < 0 \end{cases}, \tag{3}$$

where η_{prop} and η_{regen} are the overall powertrain efficiencies in the propulsion mode and the regenerative braking mode, respectively; a denotes the acceleration/deceleration rate; P_d denotes the vehicle's driving power. The vehicle model in [2] is applied to compute the E_C.

C. *RDR Estimation*

The RDR is calculated according to the following formula:

$$RDR(t) = \frac{E_{battery}(t)}{E_C(t)} = \frac{\Delta \mathrm{SOC}(t) \cdot \mathrm{C}.\eta_{mech}}{F_g + F_r + 0.5\,\rho\, C_d A \bar{V}^2 + m\bar{a}} \ , \tag{4}$$

$$F_g = m\ g\ sin\ (\emptyset)\ , \tag{5}$$

$$F_r\ = m\ g\ C_r\ cos\ (\emptyset)\ , \tag{6}$$

where C_r is the rolling coefficient and $\emptyset$ represents the grade angle of the road; C_d is the aerodynamic drag coefficient; A is the frontal area of the vehicle; ρ is the air density; m is the vehicle mass. The values of the vehicle parameters mentioned above are listed in Table 1.

Note that E_C should consider different driving modes. According to the literature [4][9], driving modes vary according to the user's speed profile, which has a significant impact on RDR. The prediction of energy consumption using the predicted speed trajectory and vehicle model [2] is the foundation for the RDR estimation of EVs. Therefore, the CNN-LSTM network used to predict vehicle velocity accounts for variations in the user's speed profile during the driving trip. As a result, the RDR is updated based on the future average predicted velocity, which is fed into the RDR estimation algorithm in a prediction time window of 15 minutes, as shown in Figure 5. This leads to an updated RDR estimation in (7).

$$RDR(w) = \frac{\Delta \mathrm{SOC} \cdot \mathrm{C}.\eta_{mech}}{F_g(w) + F_r(w) + 0.5\,\rho\, C_d A V_P^2(w) + m a_p(w)} \tag{7}$$

where V_P and a_p are the predictive velocity and acceleration. w is the prediction time window.

IV. CNN-LSTM-Based Prediction Approach

Time series prediction is a significant area of deep learning in recent years and has been widely researched for various applications, such as EVs [10][11]. From an application perspective, it is challenging to establish an accurate mathematical model for a complex dynamic system like an EV, and accurately predict future speed profiles while considering diverse driving patterns and different levels of experience. Consequently, data-driven methods can be used to predict future behaviour and performance based on historical vehicle data for EV users.

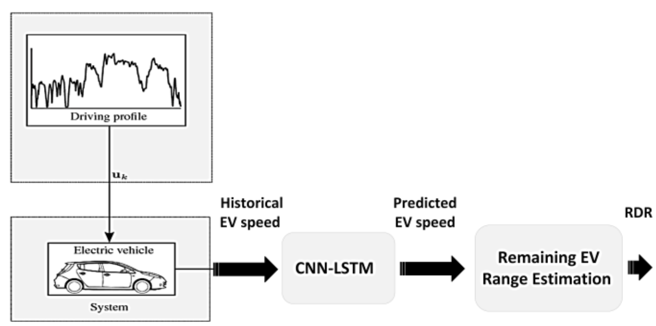

Figure 5 Deep learning-based RDR Estimation framework.

In this work, an accurate prediction system is proposed based on a hybrid deep-learning approach. This approach consists of three parts. The first part involves preparing the dataset for training and testing, which was collected from 50 EV users during a two-week test-drive program. In the second part, a CNN is chosen to automatically extract features from the time series data of vehicle velocity. In the third part, LSTM is used to complete the prediction system. Under this context, this proposed approach was applied to predict the future value of velocity every 15 minutes and fed to the RDR estimation framework discussed in the previous section. Figure 6 presents the proposed CNN-LSTM-based framework for predicting future speed profiles.

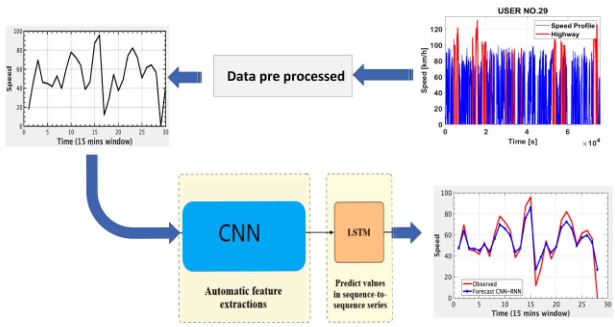

Figure 6: CNN-LSTM-based framework for predicting future speed profiles

A. The CNN-LSTM model workflow

The proposed model focuses on using deep learning to formulate and solve prediction problems. The steps of the workflow for the proposed model are:

TABLE 1: THE TEST VEHICLE PARAMETERS

Parameters	Value	Unit
ρ	1.28	kg/m^3
C_d	0.01	–
C_r	0.28	-
m	1619	kg
A	2.576	m^2

1) Training and Model Parameters Tuning

Fine-tuning of the training and model parameters was achieved by using the Bayesian optimization technique. The selected values of the parameters are listed in Table 2.

TABLE 2: PARAMETERS OF THE CNN-LSTM MODEL

Parameters	Value	Description
Lag	1	How many mins to look back
MiniBatchSize	64	Minimum Batch size
MaxEpochs	60	Maximum number of Epochs
Learning rate	0.00611	Learning rate

2) Training and Testing Data Preparation

The data was split into training and testing sets, with 90% of the data used for training and 10% for testing. To improve the convergence process, the data was normalized. After that, the lagged time series data in the "look back" rolling window, was created using the lag matrix function in MATLAB.

3) Create a Hybrid CNN-LSTM Network Architecture

In this work, the sequence look-back lag is set to 15 minutes, so the CNN learns to identify features in windows of 15 minutes. The features are passed to the network, and a regression layer with one neuron predicts the driving efficiency for the next 15 minutes (one step ahead). The architecture of the proposed hybrid network consists of three stages with 23 layers. Figure 7 presents the architecture of the full hybrid CNN-LSTM network.

Predicting the Testing Data

The network expects a sequence of lag values, and a rolling back window, to predict the cases for the next 15 minutes. The lag features of the first input set are fed to the trained model, and the prediction output contains the one-step-ahead prediction of the sequence.

Validation of Velocity Prediction Performance

In the validation process, the data sets from three new users, who were not used during the training and testing of the deep learning model, were selected for performance evaluation. Each user has different lengths of operation and

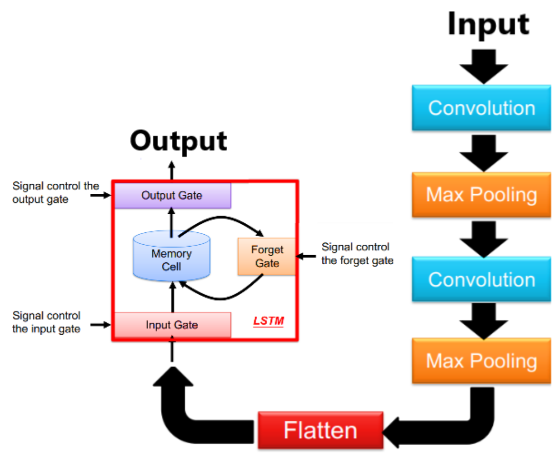

Figure 7: The architecture of the full hybrid CNN-LSTM network.

different road feature distributions. The future speed profile was predicted every 15 minutes for each user. To systematically evaluate the speed prediction results, two assessment criteria were selected: Root Mean Square Error (RMSE) in (8) and Correlation Factor (CF) in (9).

$$RMAE = \sqrt{\frac{1}{M}\sum_{i=1}^{M}|\hat{y}_i - y_i|}, \tag{8}$$

$$CF = \frac{\sum_{i=1}^{M}(\hat{y}_i-\bar{\hat{y}}_i)(y_i-\bar{y}_i)}{\sqrt{\sum_{i=1}^{M}(\hat{y}_i-\bar{\hat{y}}_i)^2(y_i-\bar{y}_i)^2}}, \tag{9}$$

where M is the size of observation data; $\hat{y}_i$ and y_i are the predicted and observed speed values respectively.

Figure 8, Figure 9, and Figure 10 present the results for velocity predictions for three unique EV users. As can be seen from Figure 8, while the average vehicle velocity changed rather dynamically during the operation for User 1, the predictions of the average vehicle velocity in each interval match well with the observed (measured) values, with the CF = 0.9964 and RMSE = 0.1098 for User 1. The good agreement between the predicted average vehicle velocity and the actual values can be observed in Figure 9 and Figure 10. It was found that CF = 0.99235 and RMSE = 0.1188 for User 2. For User 3, CF = 0.98788 and RMSE = 0.1030.

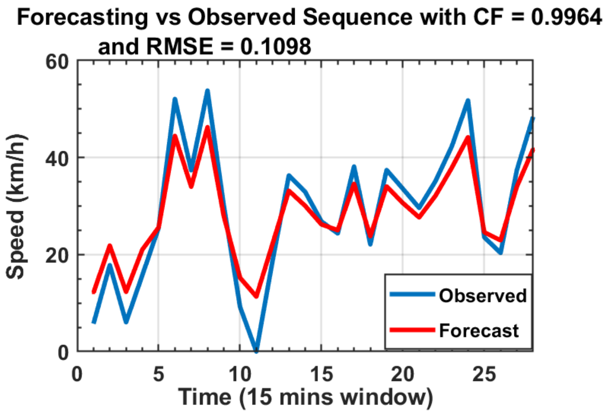

Figure 8: User 1: Velocity prediction of three EV users. The red curve is the prediction result. The blue line is the observed (measured) data.

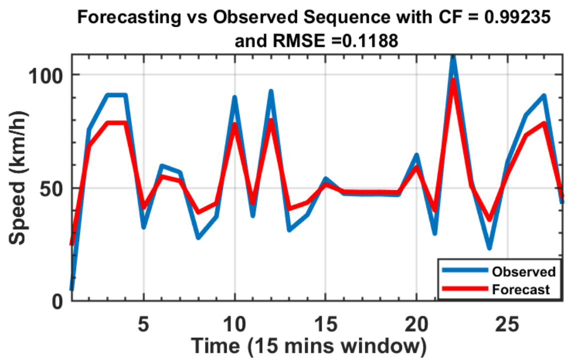

Figure 9: User 2: Velocity prediction of three EV users. The red curve is the prediction result. The blue line is the observed (measured) data.

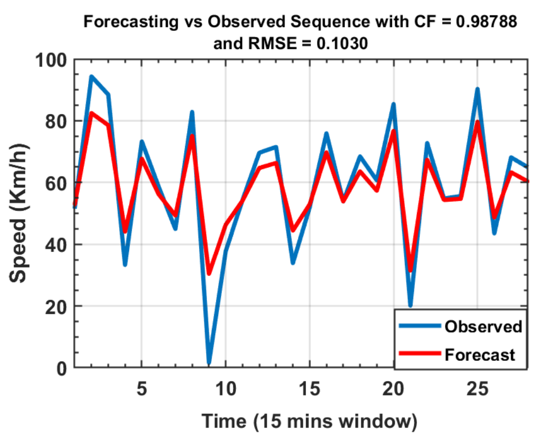

Figure 10: User 3: Velocity prediction of three EV users. The red curve is the prediction result. The blue line is the observed (measured) data.

B. Validation of RDR Estimation Algorithm

Based on the prediction of future speed profiles for EV users, the proposed RDR prediction framework was validated. RDR is calculated at each time interval and its value is updated every 15 minutes using the predicted velocity value. Figure 11 and Figure 12 present the SOC profile and results of RDR predictions for User 1, respectively. Figure 13 and Figure 14 present the SOC profile and results of RDR predictions for User 2, respectively. Figure 15 and Figure 16 present the SOC profile and results of RDR predictions for User 3, respectively.

To evaluate the RDR prediction results, the absolute error is quantified in (10).

$$E = \frac{1}{M}\sum_{i=1}^{M}\left|\widehat{RDR}_i - RDR_i\right|, \quad (10)$$

where $\widehat{RDR}_i$ and RDR_i The predicted and observed RDR respectively. RDR_i is calculated based on (4) using the measured average vehicle speed at the time interval of 15 mins. $\widehat{RDR}_i$ is computed using the proposed RDR estimation algorithm.

The average absolute errors for the three users in the evaluation process are presented in Table 3. As can be seen from Table 3, The estimation errors are 1.07 km, 4.88 km, and 5.34 km for Users 1-3, respectively, with an average estimation error of 3.76 km. The observed prediction errors in RDR are small and within an acceptable range. This finding suggests that the proposed framework has the potential to serve as a valuable tool for EV manufacturers to predict the RDR and enhance their vehicle's performance.

The observed variability in RDR prediction errors, from 1.07 km to 5.34 km, among the three users suggests that the accuracy of RDR prediction may be influenced by individual driving behaviors. This variability could be attributed to different driving patterns, such as variations in speed, acceleration, and deceleration, as well as variations in the number of stops and starts. Additionally, driving behaviors can affect the battery's efficiency and energy consumption rate, which can ultimately influence RDR predictions.

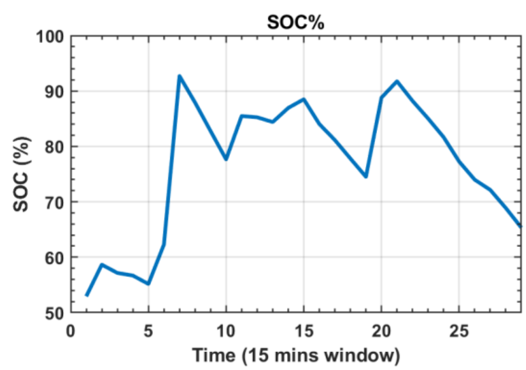

Figure 11: SOC profile for User 1.

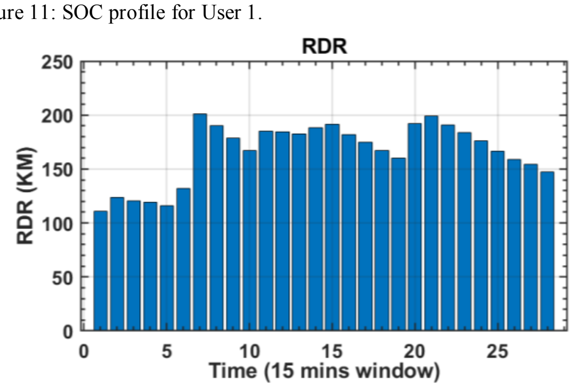

Figure 12: RDR prediction for User 1.

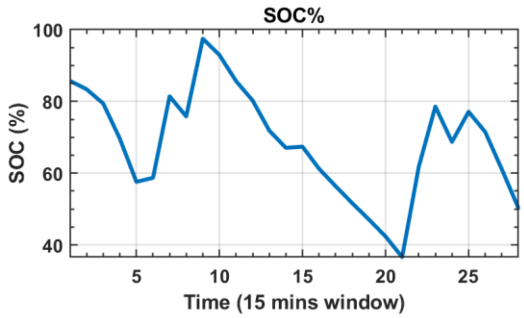

Figure 13 SOC profile for User 2.

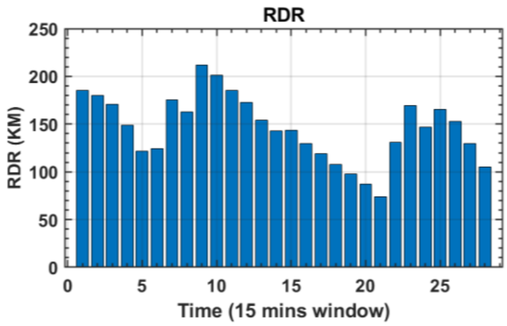

Figure 14: RDR prediction for User 2.

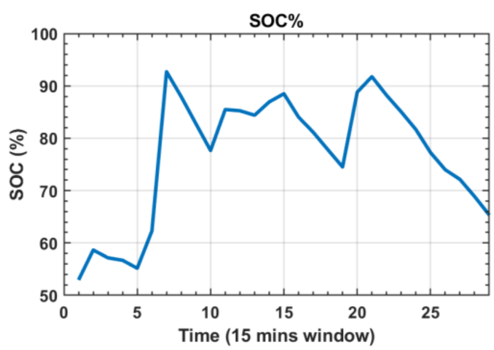

Figure 15 SOC profile for User 3.

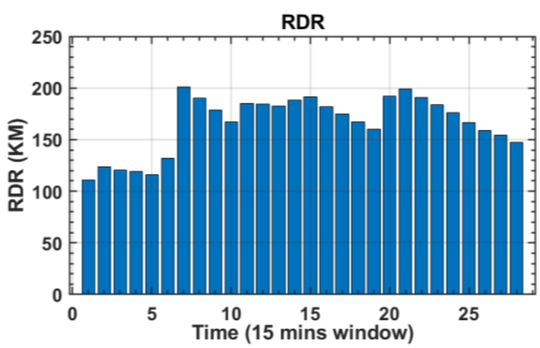

Figure 16: RDR prediction for User 3.

TABLE 3: RDR ESTIMATION ERRORS IN VALIDATION

User No	E	Unit
user 1	1.07	km
user 2	4.88	km
user3	5.34	km
Averge	3.76	km

V. CONCLUSION AND FUTURE WORK

This paper aims to fill the gap between RDR estimation methods that have unsatisfactory estimation accuracy under complex conditions. An efficient hybrid deep learning approach for RDR estimation was introduced in this study to overcome these issues. The proposed approach is divided into two stages: vehicle velocity prediction and RDR prediction. In the first stage, a CNN-LSTM-based framework was implemented for predicting EV speed profile based on real-world data from new EV users. Based on the predicted speed profiles, the RDR estimation framework was constructed in the second stage. The validation results demonstrate that the proposed predictive approach has a high accuracy in EV speed profile prediction and RDR estimation. Future work will include combing the prediction of remaining discharge energy as a function of future load prediction, and the RDR estimation method based on updated future average power and speed prediction.

ACKNOWLEDGMENTS

This research is based upon the work supported by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy under the Vehicle Technologies Office (VTO), Award Number DE-EE0008888 and DE-EE0009862.

REFERENCES

[1] M.Z. Jacobson, "Roadmaps to transition countries to 100% clean, renewable energy for all purposes to curtail global warming, air pollution, and energy risk," Earth's Future, vol. 5, pp. 948-952, 2017.

[2] M. Lamantia, Z. Su, and P. Chen, "Remaining driving range estimation framework for electric vehicles in platooning applications," in 2021 American Control Conference (ACC), pp. 424-429, 2021.

[3] J. Bi, Y. Wang, Q. Sai, and C. Ding, "Estimating remaining driving range of battery electric vehicles based on real-world data: A case study of Beijing, China," Energy, vol. 169, pp. 833-843, 2019.

[4] L. Mao, A. Fotouhi, N. Shateri, and N. Ewin, "A multi-mode electric vehicle range estimator based on driving pattern recognition," Proc. Inst. Mech. Eng. Part C, vol. 236, pp. 2677-2697, 2022.

[5] J.A. Oliva, C. Weihrauch and T. Bertram, "Model-based remaining driving range prediction in electric vehicles by using particle filtering and Markov chains," in 2013 World Electric Vehicle Symposium and Exhibition (EVS27), pp. 1-10, 2013.

[6] C. Pan, W. Dai, L. Chen, L. Chen and L. Wang, "Driving range estimation for electric vehicles based on driving condition identification and forecast," AIP Advances, vol. 7, pp. 105206, 2017

[7] L. Zhao, W. Yao, Y. Wang, and J. Hu, "Machine learning-based method for remaining range prediction of electric vehicles," IEEE Access, vol. 8, pp. 212423-212441, 2020.

[8] J.N. Kutz, S.L. Brunton, B.W. Brunton, and J.L. Proctor, Dynamic mode decomposition: data-driven modeling of complex systems, SIAM, 2016,

[9] C. Chen, X. Zhao, Y. Yao, Y. Zhang, J. Rong, and X. Liu, "Driver's eco-driving behavior evaluation modeling based on driving events," J. Adv. Transport., vol. 2018, 2018.

[10] X. Lai, Y. Huang, H. Gu, X. Han, X. Feng, H. Dai, Y. Zheng, and M. Ouyang, "Remaining discharge energy estimation for lithium-ion batteries based on future load prediction considering temperature and ageing effects," Energy, vol. 238, pp. 121754, 2022.

[11] H. Shen, Z. Wang, X. Zhou, M. Lamantia, K. Yang, P. Chen and J. Wang, "Electric Vehicle Velocity and Energy Consumption Predictions Using Transformer and Markov-Chain Monte Carlo," IEEE Transactions on Transportation Electrification, vol. 8, pp. 3836-3847, 2022.