Indonesian Journal of Electrical Engineering and Computer Science
Vol. 32, No. 2, November 2023, pp. 1004~1013
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v32.i2.pp1004-1013

Deep learning-based Ipoh driving cycle prediction

Nurru Anida Ibrahim[1], Arunkumar Subramaniam[1], Siti Norbakyah Jabar[1,2,3], Salisa Abdul Rahman[1,2,3]
[1]Department of Electronics and Instrumentation, Faculty of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu, Terengganu, Malaysia
[2]Energy Storage Research Group (ESRG), Universiti Malaysia Terengganu, Terengganu, Malaysia
[3]Renewable Energy and Power Research Interest Group (REPRIG), Universiti Malaysia Terengganu, Terengganu, Malaysia

Article Info

Article history:

Received May 30, 2023
Revised Jul 9, 2023
Accepted Jul 30, 2023

Keywords:

Driving cycle
Emissions
Fuel economy
Hybrid electric vehicles
K-means clustering
Predictive modelling
Recurrent neural network

ABSTRACT

The driving cycle is a series of driving behaviours, such as acceleration, braking, and cruising, that occur over a set length of time. Predicting the driving cycle can help to improve vehicle performance or anticipate the range of an electric car. Based on prior data, long short-term memory (LSTM) networks can be used to forecast a vehicle's driving cycle. This paper studies a driving cycle prediction based on LSTM by recurrent neural network (RNN) using developed driving cycle data. The objectives of this paper are; to develop an Ipoh driving cycle (IDC), to develop a prediction of future IDC and to analyze the prediction of IDC. Firstly, the driving data is collected in three different routes in Ipoh city at back-from-work times. Then the data is divided into micro-trips and the driving features are extracted. The features are used to develop a driving cycle using k-means clustering approach. The prediction is developed after the training of neural networks by using LSTM network approach with root mean square error (RSME) of 6.2252%.

Corresponding Author:

Salisa Abdul Rahman
Department of Electronics and Instrumentation, Faculty of Ocean Engineering Technology and Informatics
Universiti Malaysia Terengganu
21030 Kuala Nerus, Terengganu, Malaysia
Email: salisa@umt.edu.my

1. INTRODUCTION

A driving cycle is a common test process used to assess a vehicle's performance and emissions under particular driving circumstances [1]–[3]. The cycle includes a set order of acceleration, deceleration, and steady-state driving that simulates a typical driving style for a specific vehicle category, such as driving on a highway or in an urban area [4]. Regulatory organizations, including the environmental protection agency (EPA), frequently use the driving cycle to calculate a vehicle's fuel economy and emissions ratings. For consumers and governments to make educated judgements about vehicle purchases and environmental legislation, these ratings are crucial [5], [6]. The federal test procedure 75 (FTP-75) used in the United States and the new European driving cycle (NEDC) used in Europe are only two examples of the various standardized driving cycles that may be employed in different areas or nations [7]. These cycles include particular durations, repetitions, and speed and acceleration profiles that are intended to simulate actual driving circumstances in the relevant regions.

The development of driving cycles in Malaysia is important for ensuring that vehicles meet local emissions and fuel efficiency standards. It also helps to provide consumers with accurate information about the performance of different vehicles under typical driving conditions in the country [8]. While predicting the driving cycle can serve several purposes in various contexts. Predicting the driving cycle is beneficial in improving a vehicle's performance, particularly in terms of pollutants and fuel economy. Driving patterns can be reliably predicted so that car systems can be modified in real-time to function best under anticipated driving circumstances, improving efficiency and lowering environmental impact [9]–[11].

Predicting the driving cycle is also essential for efficient energy management in electric vehicles (EVs) [12]. The battery usage can be managed to maintain enough range while reducing wasteful energy consumption by anticipating future driving habits. Utilising this knowledge will enable the implementation of clever energy management techniques, such as optimising regenerative braking or changing powertrain parameters. The driving cycle in Malaysia has been established for a few states like Kuala Lumpur, Kuala Terengganu, and Ipoh itself. However, none of the research ever discussed about the prediction of the driving cycle in Malaysia.

There are several methods and techniques that can be used for driving cycle prediction which are; statistical method, machine learning, hidden Markov models (HMM), rule-based approach, hybrid approach, sensor fusion, and many more. To model and forecast upcoming driving trends, statistical techniques use historical driving data. Time series analysis, regression models, and statistical forecasting techniques are a few examples of these approaches. Future driving cycles can be predicted by identifying patterns and trends in the data and extrapolating them. This is done by analyzing the statistical features of the data. While, driving cycle prediction can also be done using machine learning approaches, such as different kinds of neural networks including long short-term memory (LSTM) networks [13]. These techniques produce forecasts using past driving data to identify trends and linkages. Algorithms that use machine learning are capable of handling complex and nonlinear data interactions as well as changing driving situations. On the other hand, HMMs are probabilistic representations of sequences of events that can be used to forecast and depict them, such as driving patterns. The fundamental system states and the likelihoods of changing between them are captured by HMMs. HMMs can calculate the most likely order of states and forecast the driving cycle by examining driving data [14].

In rule-based techniques, explicit rules or algorithms are created based on professional expertise or domain-specific norms [15]. To anticipate the driving cycle, these rules take into account a variety of input factors, including vehicle speed, acceleration, and traffic circumstances. Rule-based approaches can be useful for identifying particular driving behaviours, but they might necessitate extensive domain expertise. In order to predict the driving cycle, sensor fusion integrates data from many sensors, such as global positioning system (GPS), accelerometers, and vehicle CAN bus data [16]. A more complete and precise description of the driving behaviours can be obtained by merging data from several sensors, improving forecast accuracy. However, the choice of approach is influenced by various elements, including the data at hand, the intricacy of the driving patterns, processing power, and the particular needs of the application. It is usual practise to investigate and contrast many approaches in order to decide which one is best for driving cycle prediction in a particular situation.

In this paper, a new driving cycle was developed for Ipoh city in Perak, Malaysia. Then, the future IDC is predicted by using LSTM network. LSTM networks are a type of recurrent neural network (RNN) capable of learning order dependence in sequence prediction problems. In this research, the combination of machine learning (k-means clustering) and deep learning (RNN) will be used to maximize the accuracy, to increase the capability of handling big driving data, and to improve the efficiency and feasibility of the IDC prediction. The objectives of this paper are to develop a driving cycle of Ipoh city at back-from-work times along the main routes, to develop the prediction of the future IDC via LSTM, and to analyse the prediction of IDC. In section 2, the methodology including the route selection, data collection, driving cycle development which is by using micro-trips and k-means approach, and lastly, the prediction of IDC via LSTM will be discussed. Later, in section 3, the results and analysis of the IDC and IDC prediction will be presented.

2. METHOD

Figure 1 shows the flow chart and research activities on how to develop an IDC along the main route for the Ipoh citizens back to work at 5.00 p.m. The inputs of IDC are second-by-second speed. The data are collected at back-from-work (BFW) time which is 5.00 p.m. with 5 runs of data along three different routes. The routes have been chosen as selected routes based on its traffic volume in Ipoh. In this research, the on-board measurement method will be used using a GPS installed in the driving cycle tracking device (DC-TRAD). DC-TRAD is a device constructed mainly to collect accurate speed and time series data. The data gathered then will be stored in MySQL and managed in MATLAB. Then the parameters for each run will be extracted for the development of the driving cycle. The driving cycle is developed by using k-means clustering approach. Then, the prediction of the future driving cycle will be taking place by using LSTM network of deep learning.

An LSTM network is trained on previous driving data, which includes characteristics like speed, acceleration, and brake pedal position. The network is then used to anticipate the driving cycle. The expected driving pattern is represented by the output sequence that the LSTM network generates after processing the input sequence of characteristics over time [17].

Based on the input data and the desired output, the network learns to modify the weights of the gates in the LSTM cells during training. In order to forecast the driving cycle, the network learns to spot patterns in the input data. The driving cycle of a new car can be predicted using real-time sensor data once the network has been trained. The projected driving cycle is created by the LSTM network after processing the sensor data over time.

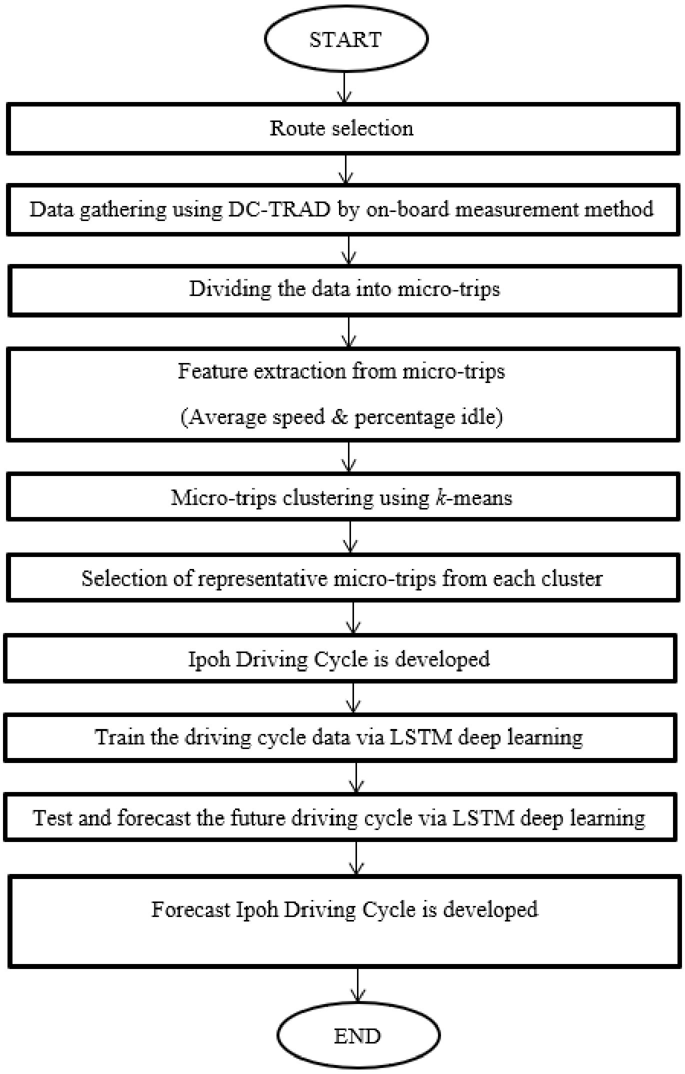
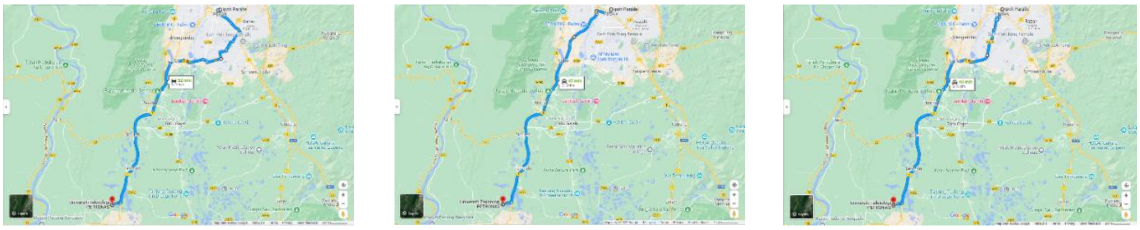


Figure 1. Flow chart of the research activities

2.1. Ipoh driving data collection

Data collection was done on back-from-work route of Ipoh city in which it begins at the Majlis Bandaraya Ipoh (MBI) building and ends at Universiti Teknologi Petronas (UTP). This route was selected as this is categorized as the busiest road in the city and this route is mostly used by the citizens of Ipoh city to go and back from work [18]. There are 3 different routes, as shown in Figure 2, namely route A, route B, and route C, from MBI to UTP, as per Figures 2(a)-2(c), respectively. Data was collected by using driving cycle tracking device (DC-TRAD) [19].



(a) (b) (c)

Figure 2. Three different routes for IDC development namely (a) route A, (b) route B, and (c) route C

2.2. IDC development

Micro trips are the foundation of a drive cycle development. A micro-trip is a journey between two timespans separated by zero vehicle velocity [20]. Every micro journey has an idle phase at the beginning and a decreasing deceleration phase at the end. The total amount of data must be divided into several micro-trips. After this procedure, a significant number of micro-trips can be acquired for all the data that was gathered. The micro-trips are then divided into a number of groups based on the traffic conditions, such as heavy, medium, and light traffic flow. The micro-trips will be clustered using the k-means technique.

It is necessary to extract driving features before clustering the micro-trips. The micro-trips can be used to extract a variety of driving characteristics, including average speed, average running speed, average acceleration and deceleration, and root mean square of acceleration. Additionally, there are four driving modes: cruising mode, accelerating mode, and decelerating mode [21]. However, just two features average speed and percentage of idle will be employed for this purpose. These two characteristics were chosen since they will have the biggest impact on emission [22]. The micro-trips are divided into three groups in Figure 3 according to the k-means clustering algorithm. Each group has its own traits and represents a distinct traffic situation; heavy, medium, and light traffic flow.

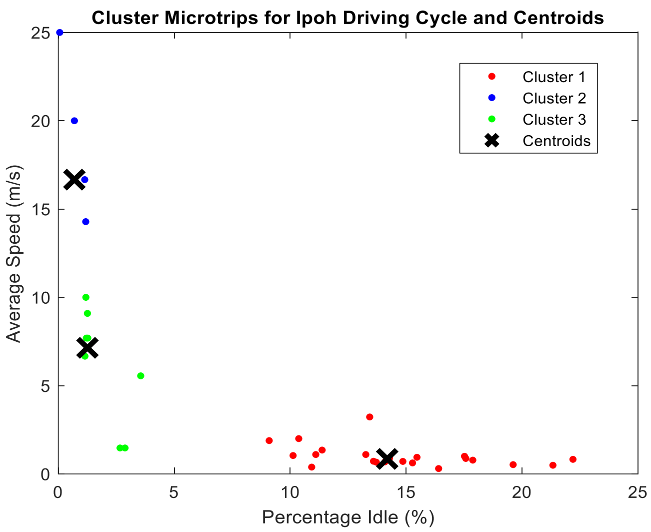


Figure 3. Clustering of the micro-trips

The driving cycle for each cluster is then created by selecting the micro-trip representatives. The representative micro-trips will be the ones that are most nearby the cluster center. The micro-trips will then be merged to create the IDC along the three major routes at 1,700 hours.

2.3. IDC prediction

A deep learning model called an RNN is made specifically to handle sequential data. It is a kind of neural network that processes sequential data, such as time series, natural language, and audio, since the architecture of the network includes loops, which allow information to persist over time. An RNN's ability to keep internal memory or hidden state, which allows it to handle inputs of any length and capture temporal dependencies within the sequence, is its important characteristic. While a form of short-term memory, this hidden state enables the network to recall and take into account previous data while it processes present inputs [23].

A form of RNN architecture called LSTM was created to solve the issue of disappearing gradients in conventional RNNs. In the vanishing gradient problem, information can be lost across lengthy input sequences when gradients get too small during backpropagation through time. By including memory cells that can retain data for a long time, LSTM networks deal with this problem. Three gates the input gate, the output gate, and the forget gate are in charge of managing these memory cells. The amount of fresh information supplied to the memory cell is governed by the input gate, the amount of information produced from the memory cell is governed by the output gate, and the amount of information preserved or forgotten in the memory cell is governed by the forget gate. The LSTM gains the ability to modify the weights of the gates based on the input data and the intended output during training. This enables the network to delete unnecessary information and store and retrieve information from the memory cells selectively. The end result is a network that can handle lengthy input data sequences while keeping track of important information and avoiding the vanishing gradient problem.

Let *N* be the number of LSTM cells. One LSTM cell has a hidden state h_t, a memory cell c_t, and three different gates, namely, input gate i_t, forget gate f_t, and an output gate o_t. Each gate uses the sigmoid as a transfer function, resulting in a value in the frequency [0, 1]. In (1) shows the formula to calculate the i_t, f_t, and o_t. In theory, the LSTM architecture takes the driving vector $v_t \epsilon R^D$, the prior hidden state $h_{t-1} \epsilon R^M$, and the previous memory cell vector c_{t-1} as inputs at time step *t*. The (2) is how it calculates g_t, h_t and c_t:

$$\begin{aligned} i_t &= \delta(W_i v_t + U_i h_{t-1} + b_i) \\ f_t &= \delta(W_f v_t + U_f h_{t-1} + b_f) \\ o_t &= \delta(W_o v_t + U_o h_{t-1} + b_o) \end{aligned} \tag{1}$$

where δ(.) represents the sigmoid function. W_i, W_f, W_o, U_i, U_f, U_o are the weight matrices for the gates. b_i, b_f, b_o are the bias weights:

$$\begin{aligned} g_t &= tan\ h\ (W_g v_t + \ U_g h_{t-1} + b_g) \\ c_t &= f_t \odot c_{t-1} + i_t \odot g_t \\ h_t &= o_t \odot tan\ h(c_t) \end{aligned} \tag{2}$$

where the notation $\odot$ denotes the element-wise multiplication, *tanh*(.) refers to the hyperbolic tangent function, W_g and U_g are the weight matrices, b_g is the bias term.

The input for the prediction procedure is the driving data load. For a better fit and to prevent the training from diverging, the training data needs to be normalised to have a zero mean and unit variance. At the time of prediction, the test data must be standardised using the same parameters as the training data. The replies must be given as training sequences with values altered by just one step in order to anticipate the values of succeeding time steps in a series. As a result, the LSTM network learns to anticipate the value of the subsequent time step at each time step in the input sequence. Predictors are exercise routines that skip the most recent time step. The next step is to build an LSTM regression network [24].

Figure 4 shows the flow chart of the prediction of the IDC. The LSTM layer in this project has 200 hidden units, and it trains the LSTM network using the predetermined training settings. Following the training phase, the predicted future time steps will occur. The root means square error (RMSE), which compares the forecast and test data, is calculated after the forecasted data has been created. The accuracy of a predictive model or the calibre of an estimator is frequently assessed using the RMSE statistic. It measures the typical discrepancy between a model's projected values and the actual observed values. The fact that RMSE is expressed using the same units as the dependent variable makes it straightforward to understand and compare to the actual observed values. Since it shows smaller average deviations between expected and actual values, a lower RMSE indicates greater predictive performance [25]. The formula for RMSE is shows in the (3).

$$RMSE = \sqrt{}(1/n) * \Sigma(y - \hat{y})^2 \tag{3}$$

where:
- n is the total number of observations.
- Σ denotes the summing operation.
- (y - ŷ) represents the difference between the actual value (y) and the predicted value (ŷ).

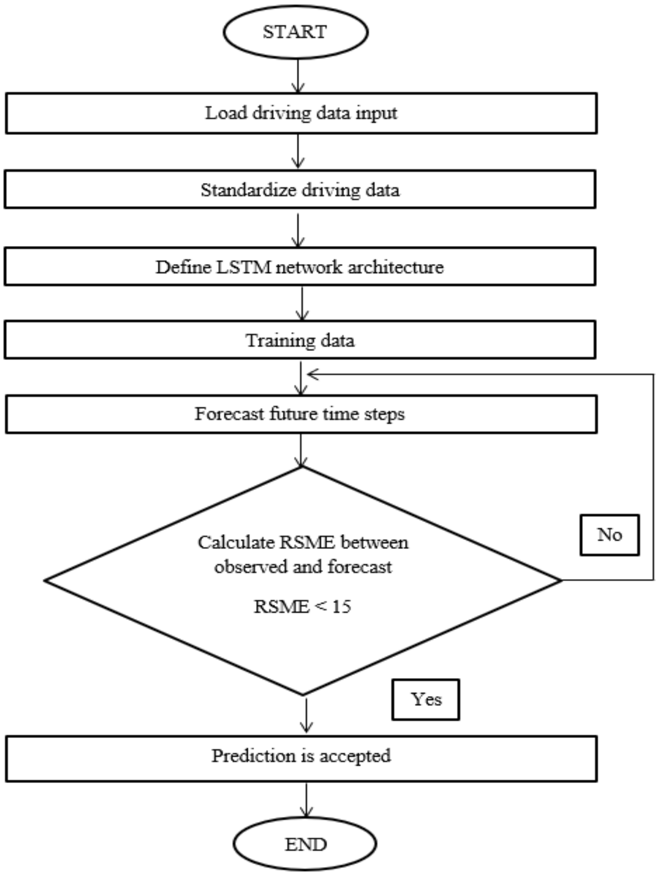


Figure 4. Flow chart of the IDC prediction

Figure 5 shows the training process for the developed IDC. The gradient threshold is set to 1 to avoid the gradients from exploding during the 250-epoch training option for this project. After 125 epochs, the learn rate is reduced by a factor of 0.2 after being set at 0.005 for the initial epoch.

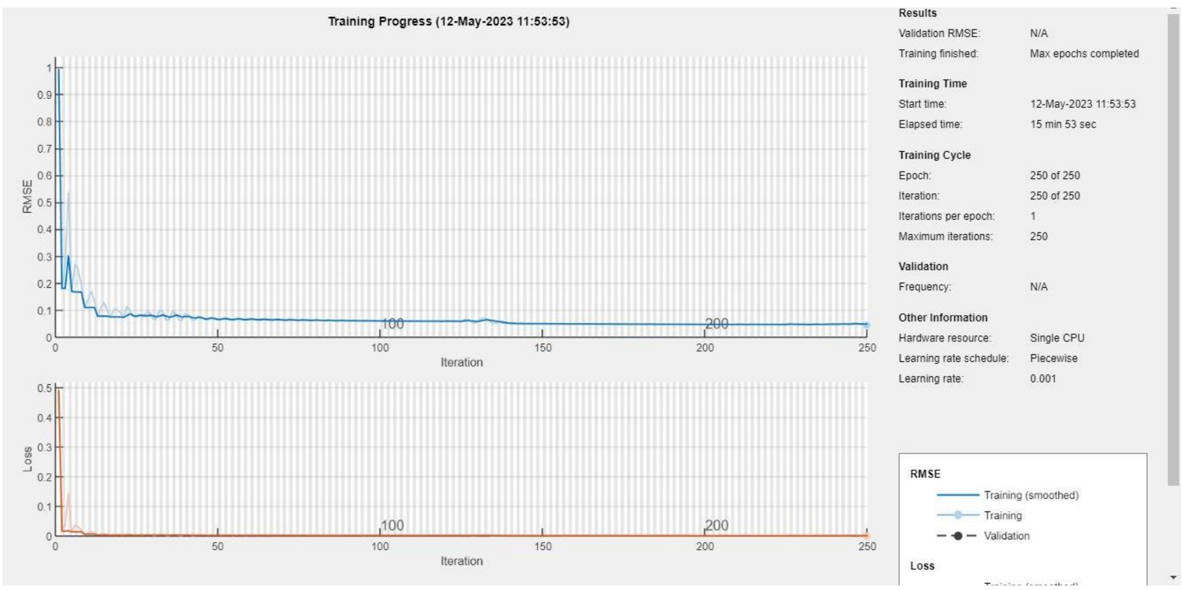


Figure 5. Training progress for IDC

3. RESULTS AND DISCUSSION

3.1. IDC analysis

The final IDC is depicted in Figure 6. Table 1 lists the IDC characteristics in terms of the nine assessment criteria. The total distance for the final IDC is 32.36 km. The table demonstrates that the prevalent speed range was above 40 km/h. This demonstrates that the Ipoh city routes are typically under medium- and clear-traffic conditions. Additionally, the produced IDC's average speed of 56.88 km/h indicates that the vehicles are going at a medium speed and that more micro-trips are being found below the average speed. As a result of the frequent stops made along the way, there is increased fuel consumption and pollution during that time.

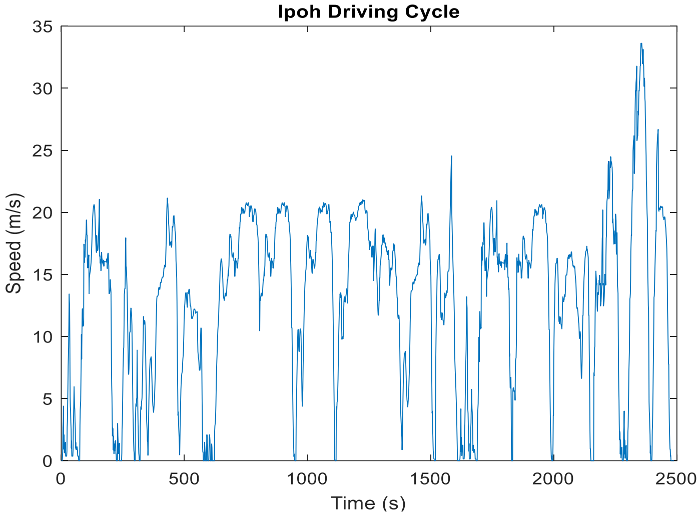


Figure 6. IDC

Table 1. The parameters of the developed Ipoh driving cycle

Parameters	Ipoh driving cycle
Distance (km)	32.36
Time (s)	2487
Average speed (km/h)	46.82
Average running speed (km/h)	49.61
Average acceleration (m/s^2)	0.56
Average deceleration (m/s^2)	0.58
RMS (m/s^2)	0.89
Percentage idle (%)	4.46
Percentage acceleration (%)	46.30
Percentage deceleration (%)	44.41
Percentage cruise (%)	4.74

3.2. IDC prediction analysis

Figure 7 shows the plot of forecast and observed data of IDC after the training process. The forecast IDC is start after 2487 seconds right after the previous IDC. Figure 8 in the other hand shows the root square mean error (RSME) between speed and time of the forecast and test data. From the figure, it shows that the RSME is 6.2252%. A RMSE of 6.2252% often denotes very strong predictive performance for the predictive model. Achieving an even lower value implies that the model is making accurate predictions relative to the scale of the target variable. An RMSE of 6.2252% denotes that, on average, the model's predictions are only 6% off from the actual value. For these applications and domain, this level of accuracy may be sufficient. It is frequently very desirable.

Real-world driving data may be noisy and inconsistent due to a number of issues, including measurement flaws or erratic traffic circumstances. LSTM networks are able to handle noisy data well and may be trained to identify patterns in the noise. The LSTM's capacity to eliminate unimportant noise and concentrate on crucial aspects aids in better driving cycle prediction. Moreover, depending on the length of the trip or the traffic, driving cycles can vary in length. LSTM networks are adaptable to handling various driving cycle durations without requiring set input sizes since they can process sequences of varying lengths. This versatility is especially useful when working with driving data from the real world, where route lengths can vary greatly.

Also, without major architectural changes, LSTM networks may adjust to shifting driving conditions or patterns. They learn from historical data and effectively grasp the underlying dynamics of the driving cycle, allowing them to generalize well to previously unimagined driving conditions. Due to its versatility, the LSTM technique can make accurate predictions even in circumstances that weren't present during training.

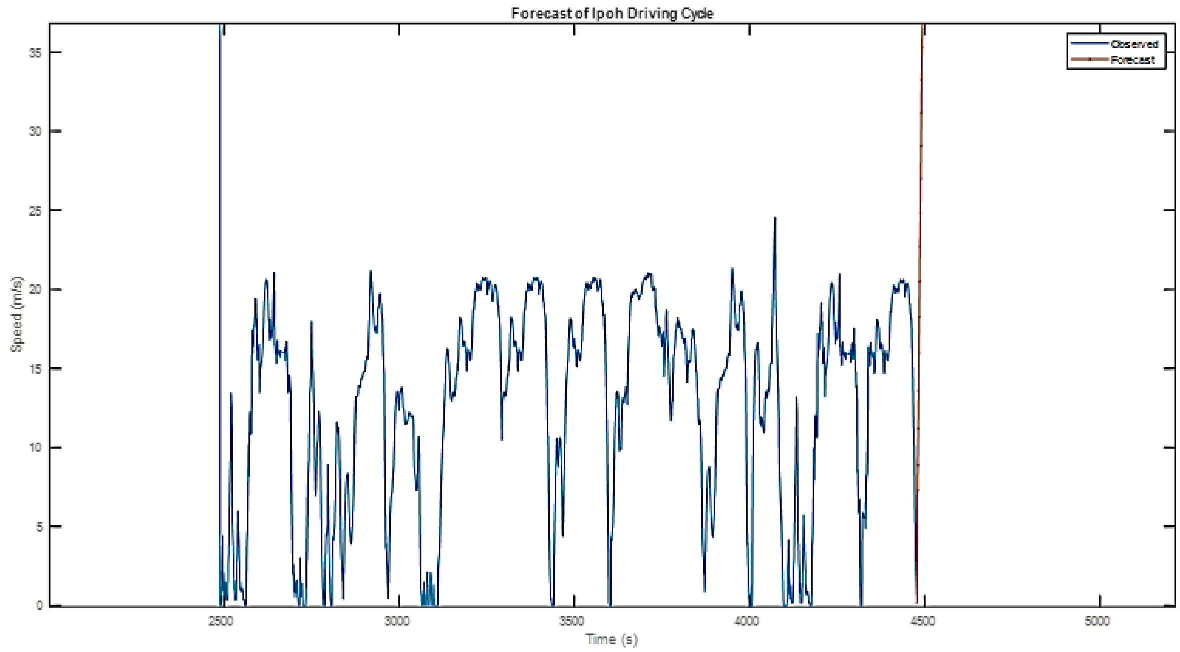


Figure 7. Prediction of IDC

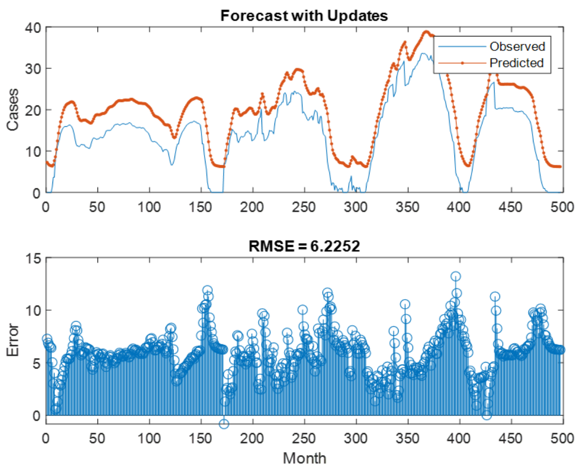


Figure 8. Comparison between forecast and test data

4. CONCLUSION

To conclude, the sampling data had been successfully collected along the three main routes in Ipoh city at the peak hour (1,700 hours) using DC-TRAD. The construction of IDC is done using clusters groups of micro-trips by k-means method. The prediction also has been successfully done by using LSTM by RNN deep learning with the RMSE is 6.2252%. All things considered, the LSTM RNN technique has advantages in capturing long-term relationships, handling variable-length sequences, maintaining temporal information, dealing with noisy and irregular data, permitting nonlinear mappings, and offering adaptability and generalisation. Because of these advantages, LSTM networks are useful for precise and reliable driving cycle prediction. For the future work, the further analysis of the IDC prediction regarding the energy consumption and emission an be discussed.

ACKNOWLEDGEMENTS

The authors would like to be obliged to Universiti Malaysia Terengganu for providing financial assistance under Matching Grant 2022 (MG1+3/ 2022/ 53495/ UMT) and Faculty of Ocean Engineering Technology and Informatics, UMT for all their technical and research support for this work to be successfully completed.

REFERENCES

[1] X. Zhao, X. Zhao, Q. Yu, Y. Ye, and M. Yu, "Development of a representative urban driving cycle construction methodology for electric vehicles: a case study in Xi'an," *Transportation Research Part D: Transport and Environment*, vol. 81, p. 102279, Apr. 2020, doi: 10.1016/j.trd.2020.102279.

[2] H. Kaymaz, H. Korkmaz, and H. Erdal, "Development of a driving cycle for Istanbul bus rapid transit based on real-world data using stratified sampling method," *Transportation Research Part D: Transport and Environment*, vol. 75, pp. 123–135, Oct. 2019, doi: 10.1016/j.trd.2019.08.023.

[3] B. Borlaug *et al.*, "Estimating region-specific fuel economy in the United States from real-world driving cycles," *Transportation Research Part D: Transport and Environment*, vol. 86, p. 102448, Sep. 2020, doi: 10.1016/j.trd.2020.102448.

[4] N. H. Arun, S. Mahesh, G. Ramadurai, and S. M. S. Nagendra, "Development of driving cycles for passenger cars and motorcycles in Chennai, India," *Sustainable Cities and Society*, vol. 32, pp. 508–512, Jul. 2017, doi: 10.1016/j.scs.2017.05.001.

[5] X. Zhao, Q. Yu, J. Ma, Y. Wu, M. Yu, and Y. Ye, "Development of a representative EV urban driving cycle based on a k-means and SVM hybrid clustering algorithm," *Journal of Advanced Transportation*, vol. 2018, pp. 1–18, Nov. 2018, doi: 10.1155/2018/1890753.

[6] H. Y. Tong, "Development of a driving cycle for a supercapacitor electric bus route in Hong Kong," *Sustainable Cities and Society*, vol. 48, p. 101588, Jul. 2019, doi: 10.1016/j.scs.2019.101588.

[7] Y. Cui, H. Xu, F. Zou, Z. Chen, and K. Gong, "Optimization based method to develop representative driving cycle for real-world fuel consumption estimation," *Energy*, vol. 235, p. 121434, Nov. 2021, doi: 10.1016/j.energy.2021.121434.

[8] M. A. Abas, S. Rajoo, and S. F. Z. Abidin, "Development of Malaysian urban drive cycle using vehicle and engine parameters," *Transportation Research Part D: Transport and Environment*, vol. 63, pp. 388–403, Aug. 2018, doi: 10.1016/j.trd.2018.05.015.

[9] H. He, J. Cao, and J. Peng, "Online prediction with variable horizon for vehicle's future driving-cycle," *Energy Procedia*, vol. 105, pp. 2348–2353, May 2017, doi: 10.1016/j.egypro.2017.03.674.

[10] J. Rios-Torres, J. Liu, and A. Khattak, "Fuel consumption for various driving styles in conventional and hybrid electric vehicles: Integrating driving cycle predictions with fuel consumption optimization," *International Journal of Sustainable Transportation*, vol. 13, no. 2, pp. 123–137, Jun. 2019, doi: 10.1080/15568318.2018.1445321.

[11] F. A. Bender, H. Uzuner, and O. Sawodny, "An adaptive driver model for driving cycle prediction in the intelligent truck," *IFAC Proceedings Volumes (IFAC-PapersOnline)*, vol. 19, no. 3, pp. 10682–10687, 2014, doi: 10.3182/20140824-6-za-1003.00285.

[12] J. J. Valera, B. Heriz, G. Lux, J. Caus, and B. Bader, "Driving cycle and road grade on-board predictions for the optimal energy management in EV-PHEVs," *International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium (EVS27)*, Nov. 2014, doi: 10.1109/EVS.2013.6914763.

[13] Y. Wu *et al.*, "Spatial–temporal data-driven full driving cycle prediction for optimal energy management of battery/supercapacitor electric vehicles," *Energy Conversion and Management*, vol. 277, p. 116619, Feb. 2023, doi: 10.1016/j.enconman.2022.116619.

[14] Y. Li, J. Peng, H. He, and S. Xie, "The study on multi-scale prediction of future driving cycle based on Markov chain," *Energy Procedia*, vol. 105, pp. 3219–3224, May 2017, doi: 10.1016/j.egypro.2017.03.709.

[15] S. Shishegar, S. Duchesne, and G. Pelletier, "An integrated optimization and rule-based approach for predictive real time control of urban stormwater management systems," *Journal of Hydrology*, vol. 577, p. 124000, Oct. 2019, doi: 10.1016/j.jhydrol.2019.124000.

[16] K. D. Huang, M. K. Nguyen, and P. T. Chen, "A rule-based control strategy of driver demand to enhance energy efficiency of hybrid electric vehicles," *Applied Sciences (Switzerland)*, vol. 12, no. 17, p. 8507, Aug. 2022, doi: 10.3390/app12178507.

[17] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, Mar. 2020, doi: 10.1016/j.physd.2019.132306.

[18] A. R. Pandi, "2014 road traffic volume Malaysia (RTVM)," *Highway Planning Division, Ministry of Works Malaysia*, 2015.

[19] S. K. Arun, I. N. Anida, P. Walker, J. S. Norbakyah, and A. R. Salisa, "A comparative study of conceptual design and prototype for DC-trad using EV powertrain for RTW DC in KT city," *International Journal of Energy Production and Management*, vol. 7, no. 1, pp. 59–72, Mar. 2022, doi: 10.2495/EQ-V7-N1-59-72.

[20] Q. Wang, H. Huo, K. He, Z. Yao, and Q. Zhang, "Characterization of vehicle driving patterns and development of driving cycles in Chinese cities," *Transportation Research Part D: Transport and Environment*, vol. 13, no. 5, pp. 289–297, Jul. 2008, doi: 10.1016/j.trd.2008.03.003.

[21] I. N. Anida and A. R. Salisa, "Driving cycle development for Kuala Terengganu city using k-means method," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 9, no. 3, pp. 1780–1787, Jun. 2019, doi: 10.11591/ijece.v9i3.pp1780-1787.

[22] A. Fotouhi and M. Montazeri-Gh, "Tehran driving cycle development using the k-means clustering method," *Scientia Iranica, Transactions A: Civil Engineering*, vol. 20, no. 2, pp. 286–293, 2013, doi: 10.1016/j.scient.2013.04.001.

[23] B. A. Hammou, A. A. Lahcen, and S. Mouline, "Towards a real-time processing framework based on improved distributed recurrent neural network variants with fastText for social big data analytics," *Information Processing and Management*, vol. 57, no. 1, p. 102122, Jan. 2020, doi: 10.1016/j.ipm.2019.102122.

[24] A. M. Howlader, D. Patel, and R. Gammariello, "Data-driven approach for instantaneous vehicle emission predicting using integrated deep neural network," *Transportation Research Part D: Transport and Environment*, vol. 116, p. 103654, Mar. 2023, doi: 10.1016/j.trd.2023.103654.

[25] S. Glen, "RMSE: root mean square error 2022," StatisticsHowTo.com: Elementary Statistics for the rest of us!, 2023 [Online]. Available: *https://www.statisticshowto.com/probability-and-statistics/regression-analysis/rmse-root-mean-square-error/*.

BIOGRAPHIES OF AUTHORS

Nurru Anida Ibrahim received the Bachelor of Engineering (Mechatronics) from International Islamic University Malaysia, Gombak in 2015 her and Master Degree in Physics (Renewable Energy) from Universiti Malaysia Terengganu in 2019. She currently pursuing her Ph.D. in Engineering (Energy and Electrical) Her main research interest is plug in hybrid vehicle, driving cycle, energy rate, emissions, and implementation of IoT in big data management. She can be contacted at email: P4620@pps.umt.edu.my.

Ts. Arunkumar Subramaniam received the Bachelor of Applied Sciences (Electronic and Instrumentation Physics) and Master Degree in Automation and Control from Universiti Malaysia Terengganu in 2019 and 2022 respectively and currently pursuing his Ph.D. in Automation and Control in Universiti Malaysia Terengganu. He is also currently working as Process Engineer (Measurement and Automation Technology) at KURZ Production Sdn. Bhd. His main research interest is in instrumentation and automation systems, control systems, communication protocols, energy management, driving cycle and implementation of IoT in big data management. He can be contacted at email: P4374@pps.umt.edu.my.

Dr. Siti Norbakyah Jabar received the Bachelor of Applied Science (Physics Electronic and Instrumentation) and M.Sc. and Ph.D. from Universiti Malaysia Terengganu, Terengganu, in 2009, 2014 and 2019 respectively. She is currently working as a lecturer at Universiti Malaysia Terengganu, Malaysia. Her main research interest is electric vehicle, hybrid electric vehicle, plug in hybrid vehicle, modelling and simulation, and renewable energy. She can be contacted at email: bakyahjabar@umt.edu.my.

AP. Ts. Dr. Salisa Abdul Rahman received the B.E. and M.E. in Electrical and Electronics Engineering from University of Technology Petronas, Perak, Malaysia in 2004 and 2006, respectively while Ph.D. in Optimal Energy Management Strategy for the University of Technology Sydney Plug-In Hybrid Electric Vehicles from University of Technology Sydney (UTS), Australia. She is currently working as a senior lecturer at University Malaysia Terengganu, Malaysia. Her research interests are in hybrid electric vehicles, innovation powertrain, simulation and modeling, energy management strategy, driving cycles, fuel economy, emissions, and optimization. She can be contacted at email: salisa@umt.edu.my.