

Article

Neural Network-Based Prediction of Vehicle Fuel Consumption Based on Driving Cycle Data

Jakov Topić *, Branimir Škugor and Joško Deur

Department of Robotics and Automation of Manufacturing Systems, Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, 10002 Zagreb, Croatia; branimir.skugor@fsb.hr (B.Š.); josko.deur@fsb.hr (J.D.)

* Correspondence: jakov.topic@fsb.hr; Tel.: +385-1-6168-555

Abstract: This paper deals with fuel consumption prediction based on vehicle velocity, acceleration, and road slope time series inputs. Several data-driven models are considered for this purpose, including linear regression models and neural network-based ones. The emphasis is on accounting for the road slope impact when forming the model inputs, in order to improve the prediction accuracy. A particular focus is devoted to conversion of length-varying driving cycles into fixed dimension inputs suitable for neural networks. The proposed prediction algorithms are parameterized and tested based on GPS- and CAN-based tracking data recorded on a number of city buses during their regular operation. The test results demonstrate that a proposed neural network-based approach provides a favorable prediction accuracy and reasonable execution speed, thus making it suitable for various applications such as vehicle routing optimization, synthetic driving cycle validation, transport planning and similar.

Keywords: driving cycle; data processing; feedforward neural networks; city buses; fuel consumption; prediction



Citation: Topić, J.; Škugor, B.; Deur, J. Neural Network-Based Prediction of Vehicle Fuel Consumption Based on Driving Cycle Data. *Sustainability* **2022**, *14*, 744. <https://doi.org/10.3390/su14020744>

Academic Editors:
Elżbieta Macioszek, Anna Granà,
Raffaele Mauro and
Margarida Coelho

Received: 8 November 2021

Accepted: 5 January 2022

Published: 10 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent decades, enhancements of information and communication technology have enabled the online collection of large quantities of driving data for the purpose of monitoring vehicle fleet operation for improved fleet management performance [1]. An essential component of fleet management systems corresponds to solving vehicle routing problems (VRP) [2], which is aimed at assigning vehicles and routes for accomplishing driving/delivery missions with a minimum number of vehicles and fuel cost. A VRP optimization algorithm requires the prediction of vehicle fuel consumption for the given route, period of day, driver, etc. [3]. The fuel consumption prediction is based on a model, which should account for both driving behaviors and traffic conditions. Since some variables, which affect vehicle fuel consumption, are not measured on-board and/or are not broadcasted (e.g., vehicle weight and ambient conditions), significant efforts are made to model fuel consumption based only on a subset of key, standardly available variables, such as vehicle velocity and road slope [4]. In general, fuel consumption models can be divided into two main categories: (i) first-principle (physics-based) models that describe vehicle dynamics at each time step using a set of mathematical equations corresponding to different vehicle subsystems and components [5], and (ii) data-driven machine learning (ML) models that represent an abstract mapping of a set of input/explanatory variables into an output space defined by target variable(s) [6].

The physics-based approaches can provide high prediction accuracy [7,8], but at the cost of low computing efficiency. Another disadvantage of physics-based models is that they require knowledge of many vehicle dynamics, powertrain parameters and multi-dimensional maps, which are usually not available. Therefore, instead of relying on a detailed microscopic model, there is a need to develop a fast, macroscopic model that predicts

fuel consumption for an entire driving cycle at once to make VRP feasible. This problem can be efficiently solved by utilizing state-of-the-art ML modelling approaches [9–11], which are characterized by automatic pattern learning from available data. Most commonly used models for these purposes are artificial neural networks (NNs), because they are universal approximators that can represent nonlinear characteristics of a complex system by using a nonlinear activation function [12,13]. The NNs can also readily be re-parameterized for different types of vehicle, which is not the case with physical models. The main challenge is to determine the NN architecture and inputs. The former is usually determined by the “trial and error” or “grid search” procedure [14], while the latter varies between studies. For example, in [15], the parameterization of models for the prediction of fuel consumption are realized on a subset of empirically selected statistical features, extracted from historical driving data. Another approach considers the application of a certain feature-selection technique such as a filter-based method (e.g., Pearson correlation) or an embedded method (e.g., LASSO regularization) to find the subset of the most relevant features for a given prediction task [10,16]. The disadvantages of these approaches are a need to extract most relevant input features instead of using full driving cycle information and a reliance on the automatic feature extraction property of NNs.

Recently, the success of NNs in regression and classification tasks has increased significantly with the development and application of more advanced NN types, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), etc., along with the utilization of deep learning (DL) techniques. A comprehensive analysis of different advanced ML models and DL techniques is given in [17], where the prediction accuracies of corresponding models are examined and compared in the case of a short-term electric microgrid load forecasting problem. More complex prediction tasks should consider probabilistic model output, which, instead of point predictions (i.e., related to expected values), relies on the prediction of conditional probability distributions of target/dependent variables or related statistical indices (e.g., quantiles) [18].

In this paper, we propose an NN model structured in two-stages, as outlined in Figure 1, where the first NN predicts the driving cycle features based on the given route and traffic conditions (current, historical), while the second NN converts the driving cycle features into the fuel consumption prediction information. The main advantage of this two-stage model arrangement is that the traffic and energy modeling tasks are effectively separated.

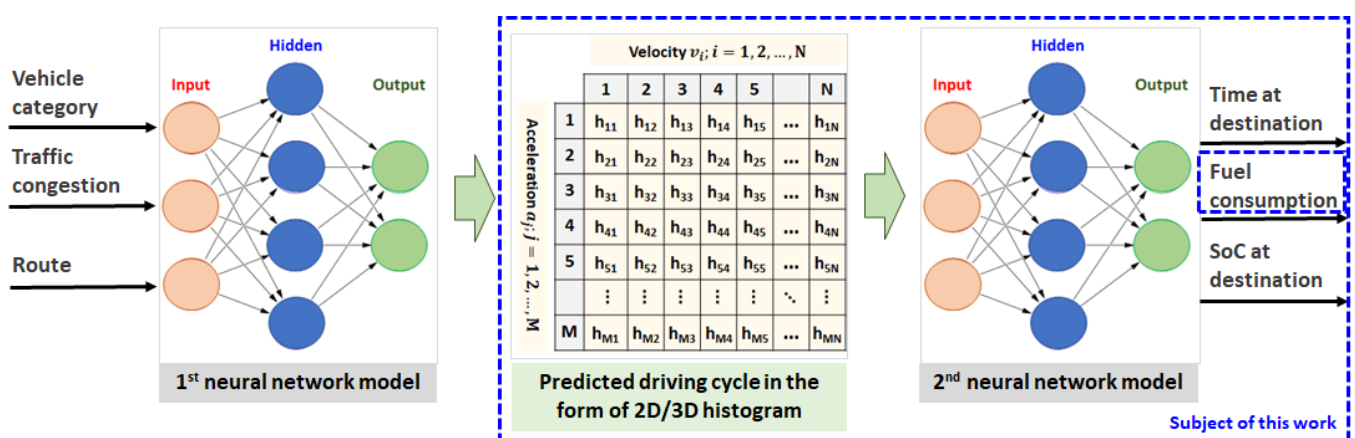


Figure 1. Illustration of a two-stage neural network (NN) model for predicting driving cycle features (1st NN model) and fuel consumption (2nd NN model).

This paper deals with the design of the second NN model based on experimentally acquired city bus GPS and CAN tracking data. By following the basic concept from [19], the recorded driving cycles of different lengths are represented by a fixed-dimension histogram, whose axes correspond to vehicle and road states (velocity, acceleration and also road slope).

For the purpose of well-conditioned NN training, which accounts for a limited resolution of fuel consumption measurement, the recorded driving cycles are pre-processed into a rich set of micro-cycles-based synthetic driving cycles. The fuel consumption prediction performance is examined in comparison with linear regression models, as well as the NN that does not account for road slope. The varying city-bus mass is not taken into account as the model input because (i) it is generally unknown and (ii) the NN model can implicitly extract a knowledge of mass from the driving cycle features [20].

Apart from proposing the two-stage structure of fuel consumption prediction NN model, as shown in Figure 1 and elaborated above, the main contributions of the paper include: (i) a method of synthesizing a rich set of representative and well-conditioned driving cycles by randomly concatenating a varying number of recorded micro-cycles corresponding to fuel consumption measurement resolution, and (ii) a systematic examination of the benefits of using the proposed NN fuel prediction consumption model when compared to simpler, linear regression and NN models, where the emphasis is on analyzing the influence of the effect of road slope.

In addition to fuel consumption prediction, the proposed NN model can be used (after retraining) to predict time-at-destination and electric vehicle (EV) battery state-of-charge (SoC), which represent crucial information for developing optimal EV charging management strategies [21]. Additionally, the NN model can be employed as a key component in the validation of synthetic driving cycles [22], which are typically used for the purpose of optimal design of EV configurations and energy management strategies, and in general fleet simulations [23].

The paper is organized as follows. Driving cycle data acquisition for a city bus is described in Section 2. The considered linear and neural network-based regression models are presented in Section 3. A detailed comparative analysis of different models in terms of fuel consumption prediction accuracy is provided in Section 4. Concluding remarks are given in Section 5.

2. Driving Cycle Processing and Analysis

2.1. Recorded Data

The driving cycle data have been recorded on 10 city buses of the same type (MAN Lion's City NL 323), operating over a single bus route in the city of Dubrovnik for the period of six months. The recording was conducted using a commercial GPS/GPRS tracking module switched to the sampling time of 1 s. The GPS data included geographical coordinates (longitude, latitude and altitude), travelled distance and vehicle velocity. In addition, the recorded data included the cumulative fuel consumption, which is acquired through the vehicle CAN bus based on the Fleet Management System (FMS) interface ver. 2.0 and the same sampling time basis of 1 s. The vehicle acceleration is determined by time differentiation of the recorded velocity. The road slope was reconstructed offline from the vertical and horizontal velocity components using high-precision GPS equipment [24]. Figure 2 shows an example of a recorded driving cycle for a single direction of the considered route.

It should be noted that for some buses the fuel consumption measurement resolution was 0.5 L, whereas on other buses this resolution was much finer and equaled 0.001 L. The methodology presentation in this paper concerns fine-resolution data (Figure 2b), while the results are given for both cases.

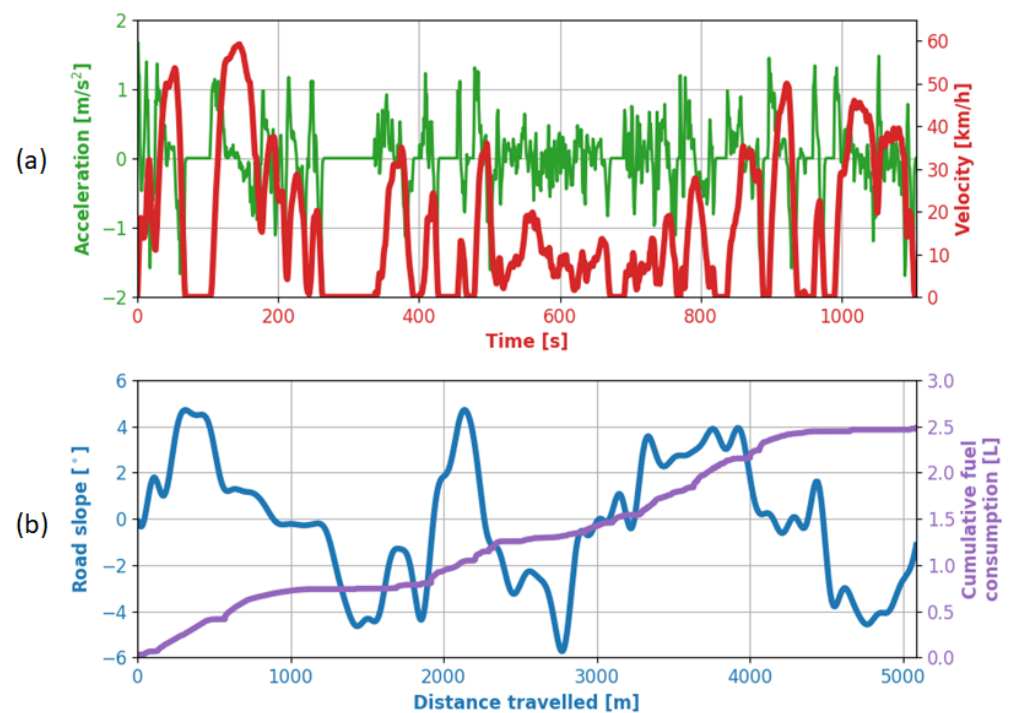


Figure 2. Example of a recorded driving cycle for city bus route Babin kuk-Pile in the city of Dubrovnik: (a) velocity and acceleration time profiles, (b) road slope and cumulative fuel consumption vs. distance travelled profile.

2.2. Data Processing

Figure 3 (orange diamonds) shows that for the given route length, the trip duration varies significantly due to traffic condition variations. When preparing the data for training and testing the fuel consumption prediction models, the recorded driving cycles are first decomposed into micro-cycles (green triangles in Figure 3), in such a way that each micro-cycle corresponds to the fuel consumption exactly equal to a (rough) measurement resolution of 0.5 L. This decomposition is applied to remove inaccuracies related to limited fuel consumption resolution. In the second data preparation step, numerous driving cycles are composed by combining different numbers of randomly chosen micro-cycles. More specifically, 20 groups of 200 combined driving cycles are generated, where each group includes a different number of micro-cycles contained within the combined cycle, starting from 1 for 1st group to 20 for 20th group. In this way, a highly dispersed dataset in terms of travelled distance (see blue circles in Figure 3) and fuel consumption is prepared, where the limited resolution-related fuel consumption errors are removed. The validity of combined driving cycles is confirmed in Figure 4, which shows that the synthesized/combined driving cycles' mean velocity and acceleration distributions resemble well the distributions of the original recorded driving cycles.

Figure 5 shows the dependence of fuel consumption on trip length and mean road slope in the case of combined driving cycles. Evidently, both trip length and mean road slope have a strong impact on fuel consumption, with very high associated correlation indices: $\rho = 0.96$ and $\rho = 0.8$, respectively (calculated by Matlab function *corr(.)*). Hence, both inputs should be included in a fuel consumption prediction model to provide high prediction accuracy.

For the sake of prediction model development, the combined driving cycles are randomly divided into three datasets: training, validation and test datasets containing 70%, 15% and 15% of total data, respectively. The training dataset is aimed at model training/parameterization, the validation dataset is used to prevent overfitting in the case of NN-based models and the test dataset is aimed at examining the model performance.

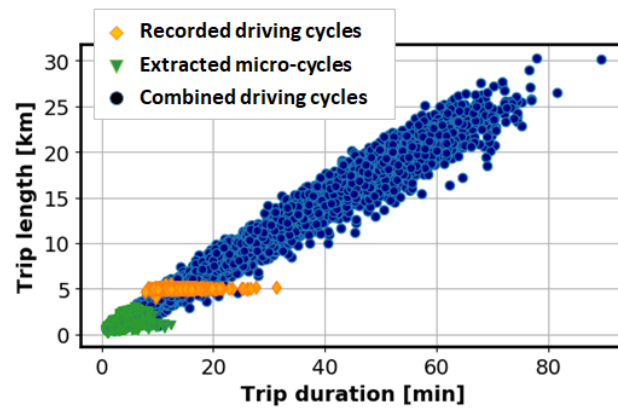


Figure 3. Trip length vs. trip duration plots for recorded driving cycles, extracted micro-cycles and combined driving cycles synthesized from micro-cycles.

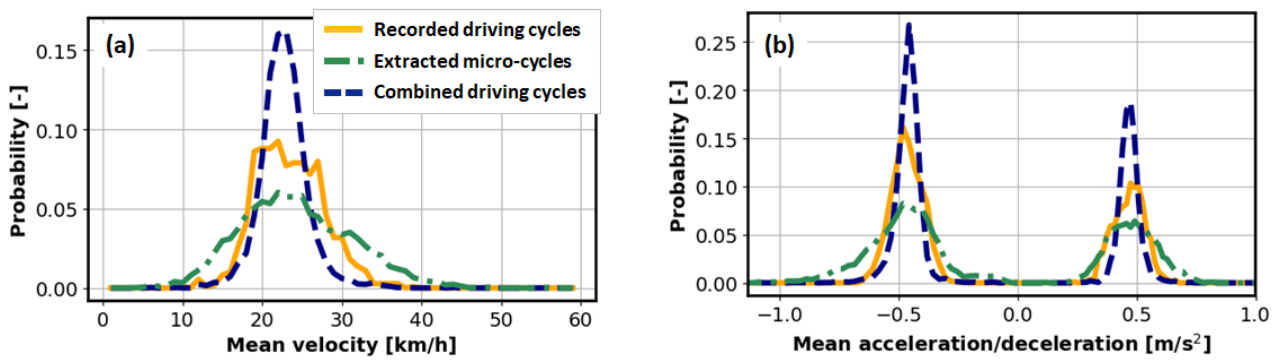


Figure 4. Distributions of mean velocity (a) and mean acceleration (b) for different driving cycles from Figure 2.

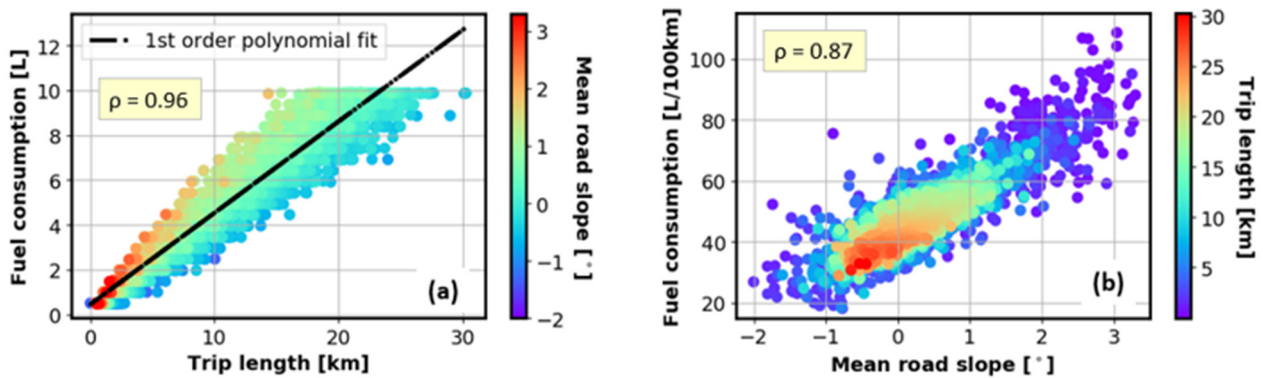


Figure 5. Fuel consumption vs. trip length (a) and mean road slope (b) for combined driving cycles.

3. Fuel Consumption Modeling

3.1. Linear Regression Models

Firstly, a simple linear model with a single, travelled distance input (d) is considered:

$$V_f = p_0 + p_1d, \tag{1}$$

where V_f is the fuel consumption and p_0 and p_1 are model parameters. The model is further referred to as $Poly_{1D}$ and is visualized in Figure 5a by a first order polynomial fit, which captures the data expectations but not their variation caused mostly by the road slope effect.

Secondly, a linear-in-parameter regression model accounting for both trip length (d) and road slope inputs (θ) is formulated as (referred to as $Poly_{2D}$ below):

$$V_f = p_{00} + p_{10}d + p_{01}\theta + p_{20}d^2 + p_{11}d\theta + p_{02}\theta^2, \quad (2)$$

where p_{ij} , $i, j = 0, 1, 2$ represents the model parameters.

The parameters of both regression models have been obtained by applying the linear least-square method to the training set.

3.2. Neural Network (NN)-Based Modeling

Two feedforward NN models are considered, with the first being represented by the architecture shown in Figure 6a and inputted by a flattened 2D histogram (H2D) defined in vehicle velocity and acceleration axes. Each matrix element stands for the number of particular (velocity, acceleration) pairs occurring in a considered combined driving cycle. Here, the vehicle velocity is discretized with the resolution of 2 km/h in the range from 0 to 80 km/h, while the acceleration is discretized with a resolution of 0.5 m/s² in the range from -3 to 3 m/s², which results in 533 possible states. By counting the discrete states (i.e., the velocity and acceleration pairs) for a given fixed sampling time, the information of distance travelled, which predominantly influences the fuel consumption, is implicitly contained within the 2D input matrix, having the same (unique) structure for driving cycles of different length [19].

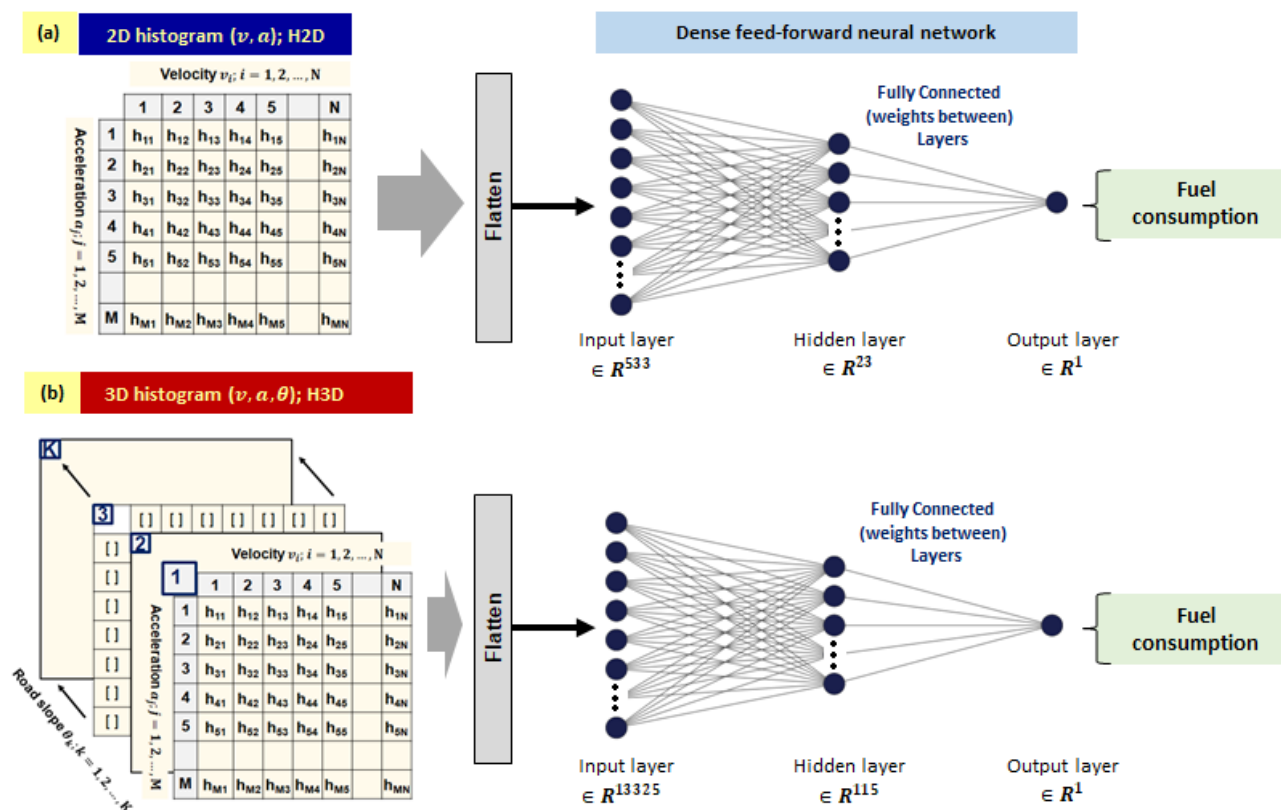


Figure 6. Proposed neural network architectures for fuel consumption prediction, based on driving cycle-related 2D (a) and 3D histogram input type (b).

The road slope information is accounted for in the second NN model in terms of 3D histogram (H3D) matrix input (Figure 6b), where the road slope is discretized with the resolution of 0.5° in the range from -6° to 6° , thus resulting in 13,325 states. The output of both NN models is represented by the cumulative fuel consumption of the entire input driving cycle (trip-based prediction), which is expressed in liters.

Training of the NN models from Figure 6 is performed within a Python environment, using the Keras module [25] with Tensorflow as a backend [26] and an in-built ADAM training algorithm [14]. The learning rate is set to $3 \cdot 10^{-3}$ because the performance of a Learning Rate Range Test [27] has indicated that the optimal learning rate for this prediction task is somewhere between 10^{-3} and 10^{-2} . The batch size and the number of epochs are set to 16 and 100, respectively. Furthermore, each hidden layer uses a Rectified Linear Unit (ReLU) activation function. The fuel consumption mean square error (MSE) is chosen as a loss function to be minimized:

$$L = \frac{1}{n} \sum_{i=1}^n \left(V_{f,model,i} - V_{f,rec,i} \right)^2, \quad (3)$$

where n denotes the total number of driving cycles included in training and $V_{f,model,i}$ and $V_{f,rec,i}$ stand for the model-predicted and recorded fuel consumption for the i th driving cycle, respectively.

The particular NN architectures shown in Figure 6 have been selected according to the geometric pyramid rule proposed in [28]. This rule states that for a three-layer network with N input and M output neurons, the hidden layer should have $\sqrt{N \cdot M}$ neurons. By subsequently performing the grid search method based on a generalization error criterion [14] (where different candidates characterized by various numbers of hidden layers and related neurons are considered), it has been found that only a negligible improvement in prediction accuracy can be obtained by increasing the NN complexity. Therefore, the initial NN architectures have been retained, i.e., the ones considering a single hidden layer with 23 and 115 hidden neurons are used for the cases of NN-H2D and NN-H3D, respectively.

The learning curves of both NN models are shown in Figure 7 to illustrate the progress of the train and validation loss functions over learning epochs. In the case of the H2D input, both the train and validation loss curves saturate at around 40 epochs, and they reach roughly the same values. This indicates the quality of training and the absence of overfitting (i.e., the validation curves does not start to increase after a certain number of epochs; the monotonicity feature). In the case of the H3D input, both training and validation losses are significantly lowered when compared to the H2D input case, thus preliminarily indicating the importance of accounting for the road slope information. The gap between these two curves increases as learning progresses, meaning that the NN model fit is getting better on the training dataset when compared to the validation dataset. Despite this gap, the validation learning curve monotonicity is preserved, i.e., there is no overfitting. The weights for which the lowest validation loss is obtained during the NN's learning process are finally adopted as parameters of corresponding NNs.

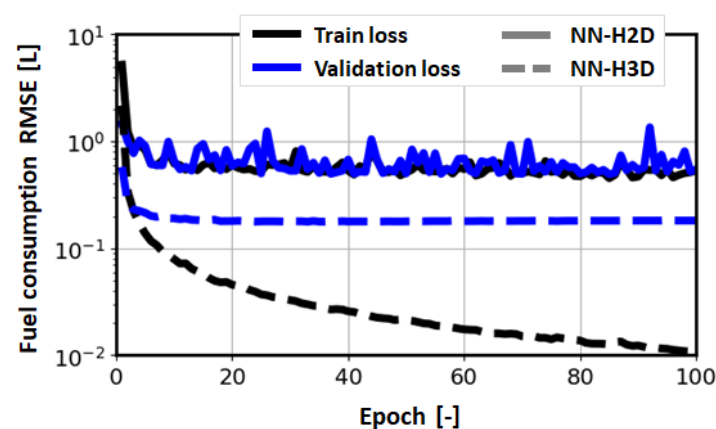


Figure 7. Train and validation learning curves for different (H2D and H3D input-related) NN architectures.

4. Comparative Results

The performance of fuel consumption prediction based on the four models established in Section 3 has been examined by using the test dataset, comprised of augmented experimental data for two different resolutions of fuel consumption measurements, separately (rough and fine; see Section 2). Figure 8 shows the corresponding plots of predicted vs. real/recorded fuel consumption. Evidently, the simplest linear model $Poly_{1D}$ exhibits the largest errors regarding fuel consumption predictions (i.e., deviation with respect to an ideal 1:1 straight line), while the H3D input-based NN model provides the most accurate predictions.

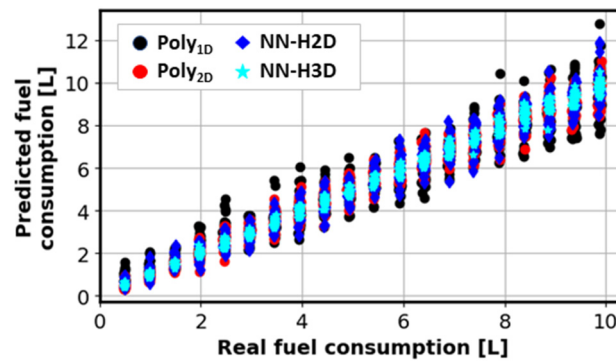


Figure 8. Predicted vs. real fuel consumption determined for different models based on test dataset.

The respective distributions of absolute and relative prediction errors/residuals are shown in Figure 9. Interestingly, the more complex linear regression model $Poly_{2D}$ has comparable error distribution to the simpler NN model (NN-H2D). Specifically, accounting for the road slope information can make the regression model competitive with the NN model with no road slope input. The ultimate NN-H3D model significantly outperforms other models, with its prediction error distributed mostly within ± 0.5 L (i.e., within $\pm 10\%$ of relative error). It should also be noted that the predictions are well balanced, with their mean error close to zero.

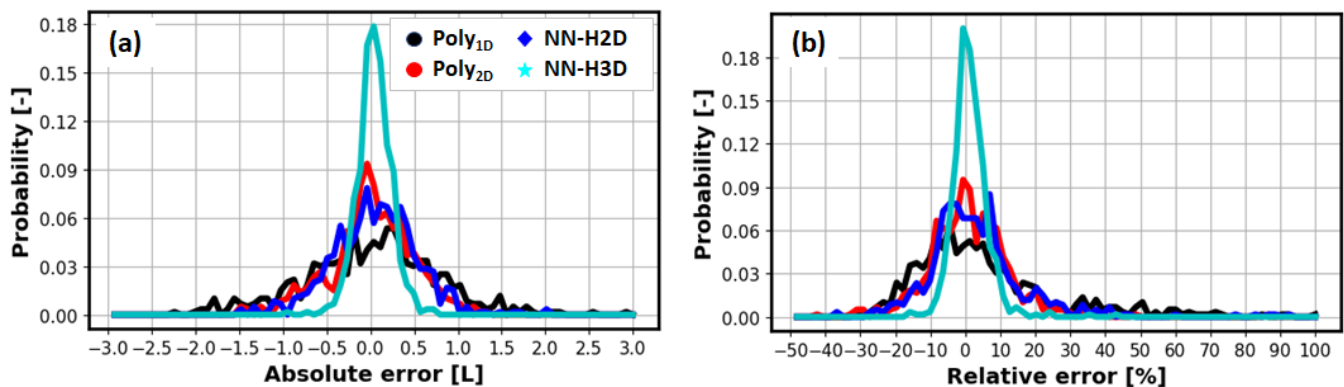


Figure 9. Distribution of prediction errors/residuals for different models and test dataset: absolute (a) and relative residuals (b).

The model's performance is further quantified by an R^2 value which is determined on the test dataset as:

$$R^2 = 1 - \frac{\sum_{i=1}^n (V_{f,rec,i} - V_{f,model,i})^2}{\sum_{i=1}^n (V_{f,rec,i} - \bar{V}_{f,rec})^2}. \quad (4)$$

The R^2 value is often interpreted as an amount of variance in recorded data, which can be explained by the considered model. It takes values between 0 and 1, where $R^2 = 1$

corresponds to an ideal fit, while $R^2 = 0$ corresponds to the case in which the model output is constant and equal to the mean value of recorded fuel consumptions (here $\bar{V}_{f,rec}$). Additionally, R^2 is expressed with respect to $Poly_{1D}$ model fuel consumption predictions, where $V_{f,Poly_{1D},i}$ is used instead of $\bar{V}_{f,rec}$ in (4). This indicator is denoted as R^2_{corr} , and can be interpreted as an additional variance in recorded data that can be explained by the considered model and that is not explained by the $Poly_{1D}$ model.

Table 1 provides standard deviations of prediction residuals, corresponding R^2 and R^2_{corr} indicators for each model and average execution times \bar{T}_{exec} (for a single prediction), all calculated on the test dataset. The standard deviation of residuals of the NN-H3D model equals 0.19 L, and is reduced approximately by the factors of 4 and 2.5 when compared to the $Poly_{1D}$ model and the $Poly_{2D}$ and NN-H2D models, respectively. Note that an interval which contains 95% of model predictions can be calculated as approximately $\pm 2 \cdot \sigma_{res}$, which in the case of the NN-H3D model equals to around 0.35 L. All models have large R^2 values (>0.9), where the NN-H3D model approaches the almost ideal value of 1. The relative relations of different models are better described when using the R^2_{corr} indicator, which shows that the $Poly_{2D}$ and NN-H2D models explain around 60% more of the variance in the recorded data compared to the $Poly_{1D}$ model, while in the case of the NN-H3D model this share exceeds 90% (i.e., again nearing the ideal value of 1). The average execution time \bar{T}_{exec} for a single prediction falls in the range from 1.5 to 1420 μs , depending on the model complexity. Moreover, the average execution time \bar{T}_{exec} for the NN-H2D model is 35% lower with respect to the NN-H3D model, in spite of significant differences in input size (i.e., vector with 13,325 elements for the NN-H3D model vs. 533 elements for the H2D model; see Section 3).

Table 1. Fuel consumption prediction performance indicators for different models and test dataset.

Model	Standard Deviation of Residuals, σ_{res} [L]	R^2 [-]	R^2_{corr} [-]	\bar{T}_{exec} ² [μs]
<i>Poly_{1D}</i>	0.757 (0.697) ¹	0.931 (0.943)	0.0 (0.0)	1.47
<i>Poly_{2D}</i>	0.472 (0.398)	0.973 (0.981)	0.611 (0.674)	3.75
NN-H2D	0.484 (0.470)	0.972 (0.973)	0.593 (0.534)	920
NN-H3D	0.194 (0.133)	0.995 (0.998)	0.934 (0.964)	1420

¹ Numbers in brackets correspond to the case of using low-resolution fuel consumption tracking data. ² Correspond to average execution time on a workstation having 16 GB RAM and Intel® Xeon® Processor E5-1620 v3 @ 3.50GHz.

Table 1 also provides the results related to the dataset with a rough fuel consumption resolution (equal to 0.5 L; see Section 2). These results confirm that a similar level of performance can be achieved regardless of whether low-resolution or high-resolution fuel consumption tracking data are available.

5. Conclusions

Different approaches to fuel consumption prediction based on driving cycle data-fed linear regression and NN models have been presented. For the needs of well-conditioned model training, an originally recorded driving cycle dataset has been augmented by generating a rich set of combined driving cycles in terms of randomly concatenating a different number of recorded micro-cycles, where each micro-cycle is determined by cumulative fuel consumption equal to the measurement resolution of 0.5 L. In the case of the NN model, the driving cycles of different lengths are faithfully represented by a fixed-dimension histogram/matrix, where each cell value is obtained by counting the driving cycle discrete-amplitude state information related to vehicle velocity and acceleration, as well as road slope.

It has been shown that adding the road slope input significantly increases the prediction accuracy, i.e., the standard deviation of residuals reduces from 0.8 L to 0.5 L in the case of linear regression models, and from 0.5 L to 0.2 L in the case of neural network models. The ultimate performance of the NN model based on 3D histogram matrix input is manifested in its ability to capture variance in the recorded data to an extent that is 95% higher when compared to the basic linear regression model. By proper synthesis of driving cycles used for NN training, the favorable accuracy of NN-based fuel consumption prediction is preserved in the case of using more common, low-resolution fuel consumption measurement (R_{corr}^2 of 0.534 and 0.964 for the NN-H2D and NN-H3D models, respectively). The NN models have also shown to be rather computationally efficient (\bar{T}_{exec} between 920 and 1420 μ s), thus confirming their suitability for vehicle routing problem (VRP) applications.

The main advantage of the proposed NN-based fuel consumption prediction approach when compared to the use of physical model-based microsimulations is its high accuracy with no demand on a number of physical city bus parameters. However, the approach can have deteriorated accuracy in extrapolation regions, i.e., when facing unseen input data. An additional disadvantage of the proposed NN model is that it does not take into account the impact of auxiliary devices such as heating, ventilation and air conditioning (HVAC) systems, which are not directly related to the driving cycle. Similarly, although implicitly captured through the driving cycle, the absence of an explicit passenger mass input can affect the prediction accuracy. Additionally, the presented study relates to a city bus fleet consisting of the same vehicle/engine type, and the NN designed may be inaccurate in the case of a mixed-vehicle fleet. Finally, the NN is not directly applicable in real-time applications, because it requires knowledge of the driving cycle in advance, which is usually not available.

Therefore, the main effort of future research should be directed towards the design of the first NN from Figure 1, which would provide a driving cycle histogram for the NN designed in this paper, based on the knowledge of vehicle route and current and historical traffic condition data, as well as vehicle-specific features (e.g., curb weight, vehicle category, engine displacement or similar). In addition, the impact of HVAC consumption should be modeled through a separate regression model trained on the same experimental dataset, which would take into account the weather condition data (ambient air temperature and solar irradiance) as primary inputs, as well as vehicle velocity (reflecting the HVAC radiator inlet air mass flow) and passenger load (related to passengers' metabolic heat). To further improve the prediction accuracy, vehicle mass data could be added as an input to the NN model. A possible practical realization of such a prediction model could be based on increasing each element of the flattened input histogram/matrix by the amount of vehicle mass averaged over the entire driving cycle. Similarly, information on vehicle or engine type should be fed to the NN to account for mixed fleet conditions. Finally, using a geographically and chronologically broader input dataset and/or applying certain regularization techniques should be considered to overcome possible extrapolation issues.

Author Contributions: Conceptualization, J.T., B.Š. and J.D.; methodology, J.T. and B.Š.; software, J.T.; validation, J.T., B.Š. and J.D.; writing—original draft preparation, J.T. and B.Š.; writing—review and editing, B.Š. and J.D.; visualization, J.T.; supervision, J.D. All authors have read and agreed to the published version of the manuscript.

Funding: It is gratefully acknowledged that this work has been done within the project ACHIEVE (“Adaptive and Predictive Control of Plug-in Hybrid Electric Vehicles”; web site: <http://achieve.fsb.hr/> (accessed on 7 November 2021)), supported by the Croatian Science Foundation under the Grant agreement No. IP-2018-01-8323. In addition, the research work of the first author was supported by the European Regional Development Fund under the grant KK.01.1.1.01.0009 (DATACROSS).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data are not publicly available due to privacy restrictions of related transport company.

Acknowledgments: We thank Libertas d.o.o. for city bus tracking data support.

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

References

1. Yao, Y.; Zhao, X.; Liu, C.; Rong, J.; Zhang, Y.; Dong, Z.; Su, Y. Vehicle Fuel Consumption Prediction Method Based on Driving Behavior Data Collected from Smartphones. *J. Adv. Transp.* **2020**, *2020*, 1–11. [[CrossRef](#)]
2. Kedia, R.K.; Naick, B.K. Review of vehicle route optimisation. In Proceedings of the 2017 2nd IEEE International Conference on Intelligent Transportation Engineering (ICITE), Singapore, 1–3 September 2017; pp. 57–61.
3. Hu, L.; Zhong, Y.; Hao, W.; Moghimi, B.; Huang, J.; Zhang, X.; Du, R. Optimal Route Algorithm Considering Traffic Light and Energy Consumption. *IEEE Access* **2018**, *6*, 59695–59704. [[CrossRef](#)]
4. Wickramanayake, S.; Bandara, H.M.N.D. Fuel consumption prediction of fleet vehicles using Machine Learning: A comparative study. In Proceedings of the 2016 Moratuwa Engineering Research Conference (MERCon), Moratuwa, Sri Lanka, 5–6 April 2016; pp. 90–95. [[CrossRef](#)]
5. Ben-Chaim, M.; Shmerling, E.; Kuperman, A. Analytic Modeling of Vehicle Fuel Consumption. *Energies* **2013**, *6*, 117–127. [[CrossRef](#)]
6. Perrotta, F.; Parry, T.; Neves, L.C. Application of machine learning for fuel consumption modelling of trucks. In Proceedings of the 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 1–14 December 2017.
7. Suh, B.; Chang, Y.H.; Han, S.B.; Chung, Y.J. Simulation of a powertrain system for the diesel hybrid electric bus. *Int. J. Automot. Technol.* **2012**, *13*, 701–711. [[CrossRef](#)]
8. Wang, J.; Rakha, H.A. Fuel consumption model for heavy duty diesel trucks: Model development and testing. *Transp. Res. Part D Transp. Environ.* **2017**, *55*, 127–141. [[CrossRef](#)]
9. Çapraz, A.G.; Özel, P.; Sevkli, M.; Beyca, Ö.F. Fuel Consumption Models Applied to Automobiles Using Real-time Data: A Comparison of Statistical Models. *Procedia Comput. Sci.* **2016**, *83*, 774–781. [[CrossRef](#)]
10. Kim, Y.-R.; Jung, M.; Park, J.-B. Development of a Fuel Consumption Prediction Model Based on Machine Learning Using Ship In-Service Data. *J. Mar. Sci. Eng.* **2021**, *9*, 137. [[CrossRef](#)]
11. Bhoraskar, A. Prediction of Fuel Consumption of Long Haul Heavy Duty Vehicles using Machine Learning and Comparison of the Performance of Various Learning Techniques. Master's Thesis, Delft University of Technology, Delft, The Netherlands, 30 August 2019.
12. Schoen, A.; Byerly, A.; Hendrix, B.; Bagwe, R.M.; Dos Santos, E.C.; Ben Miled, Z. A Machine Learning Model for Average Fuel Consumption in Heavy Vehicles. *IEEE Trans. Veh. Technol.* **2019**, *68*, 6343–6351. [[CrossRef](#)]
13. Li, Y.; Tang, G.; Du, J.; Zhou, N.; Zhao, Y.; Wu, T. Multilayer Perceptron Method to Estimate Real-World Fuel Consumption Rate of Light Duty Vehicles. *IEEE Access* **2019**, *7*, 63395–63402. [[CrossRef](#)]
14. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
15. Delgado, O.F.; Clark, N.N.; Thompson, G.J. Modeling Transit Bus Fuel Consumption on the Basis of Cycle Properties. *J. Air Waste Manag. Assoc.* **2011**, *61*, 443–452. [[CrossRef](#)] [[PubMed](#)]
16. Bifulco, G.N.; Galante, F.; Pariota, L.; Spena, M.R. A Linear Model for the Estimation of Fuel Consumption and the Impact Evaluation of Advanced Driving Assistance Systems. *Sustainability* **2015**, *7*, 14326–14343. [[CrossRef](#)]
17. Lopez-Martin, M.; Sanchez-Esguevillas, A.; Hernandez-Callejo, L.; Arribas, J.; Carro, B. Novel Data-Driven Models Applied to Short-Term Electric Load Forecasting. *Appl. Sci.* **2021**, *11*, 5708. [[CrossRef](#)]
18. Lopez-Martin, M.; Sanchez-Esguevillas, A.; Hernandez-Callejo, L.; Arribas, J.; Carro, B. Additive Ensemble Neural Network with Constrained Weighted Quantile Loss for Probabilistic Electric-Load Forecasting. *Sensors* **2021**, *21*, 2979. [[CrossRef](#)] [[PubMed](#)]
19. Topić, J.; Škugor, B.; Deur, J. Neural Network-Based Modeling of Electric Vehicle Energy Demand and All Electric Range. *Energies* **2019**, *12*, 1396. [[CrossRef](#)]
20. Hihlik, M.; Topić, J.; Škugor, B.; Deur, J. Neural Network-based Prediction of Fuel Consumption of a Conventional Delivery Vehicle based on GPS-collected Tracking Data. In Proceedings of the 15th Conference on Sustainable Development of Energy, Water and Environment Systems (SDEWES), Cologne, Germany, 1–5 September 2020.
21. Said, D.; Mouftah, H.T. A Novel Electric Vehicles Charging/Discharging Management Protocol Based on Queuing Model. *IEEE Trans. Intell. Veh.* **2019**, *5*, 100–111. [[CrossRef](#)]
22. Topić, J.; Škugor, B.; Deur, J. Synthesis and Feature Selection-Supported Validation of Multidimensional Driving Cycles. *Sustainability* **2021**, *13*, 4704. [[CrossRef](#)]
23. Silvas, E.E.; Hofman, T.T.; Murgovski, N.N.; Etman, L.F.P.; Steinbuch, M.M. Review of Optimization Strategies for System-Level Design in Hybrid Electric Vehicles. *IEEE Trans. Veh. Technol.* **2016**, *66*, 1. [[CrossRef](#)]
24. Škugor, B.; Hrgetić, M.; Deur, J. GPS measurement-based road slope reconstruction with application to electric vehicle simulation and analysis. In Proceedings of the 8th Conference on Sustainable Development of Energy, Water and Environment Systems (SDEWES), Dubrovnik, Croatia, 27 September–2 October 2015.
25. Keras. Available online: <https://keras.io> (accessed on 7 November 2021).

26. Tensorflow. Available online: <https://tensorflow.org> (accessed on 7 November 2021).
27. Smith, L.N. Cyclical Learning Rates for Training Neural Networks. In Proceedings of the 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), Santa Rosa, CA, USA, 24–31 March 2017. [CrossRef]
28. Masters, T. Designing Feedforward Network Architectures. In *Practical Neural Network Recipes in C++*; Elsevier BV: New York, NY, USA, 1993; pp. 173–185.