EVS38 Symposium
Göteborg, Sweden, June 15-18, 2025

Investigation on Machine Learning Models for Forecasting Auxiliary Energy Consumption of HD BEVs

Zhenkan Wang[1], Yuantao Fan[2], Henrik Ydreskog[1], Slawomir Nowaczyk[2]

[1] *Volvo Group Trucks Technology, zhenkan.wang@volvo.com*
[2] *Center for Applied Intelligent System Research, Halmstad University*

Executive Summary

Forecasting the energy consumption of a battery electric vehicle (BEV) is crucial for both customers and sales teams in selecting the optimal configuration of the vehicle, as well as for route planning and charging strategy development. A dedicated physics-based model has been developed for propulsion systems at Volvo for years. However, the usage of vehicles, especially heavy-duty (HD) trucks, is not always dependent upon propulsion systems but also upon other auxiliary systems. The auxiliary energy consumption in an HD truck can normally be more than 10% of the total energy consumption or even more than 20% in some extreme conditions. Developing effective data-driven models to enhance energy forecasting accuracy would greatly improve the planning of operations for HD vehicles. This paper investigates the use of machine learning models, trained on historical operational data from commercial HD BEVs, to forecast the auxiliary energy consumption. A variety of data-driven approaches have been evaluated and compared based on forecasting accuracy and computational efficiency. We explore the trade-offs between these factors to identify optimal solutions for practical deployment.

Keywords: Electric Vehicles, Heavy Duty electric Vehicles & Buses, AI - Artificial Intelligence for EVs, Auxiliary Components & Sensors.

1 Introduction

The electric truck market is expanding and the size is projected to reach 1,067,985 units by 2030 [1]. Electric vehicles (EVs) quickly gain traction, enabling rapid acceleration and delivering a highly responsive driving experience. Take into account their zero tailpipe emissions, making them an efficient and environmentally friendly alternative to internal combustion engine (ICE) vehicles [2]. However, one of the key challenges facing electric trucks is their limited driving range compared to ICE trucks. For example, many electric trucks have ranges around 200-300 miles per charge, whereas diesel trucks can often travel over 1,000 miles on a single tank [3]. Therefore, heavy-duty (HD) truck drivers may experience ”range anxiety” when transitioning from ICE trucks to electric ones, due to the shorter range of fully electric vehicles compared to their ICE counterparts. The “range anxiety” (i.e., user’s concern about the insufficient all-electric range of an EV to reach a destination or charging point) is considered a major barrier that limits the wide adoption of HD electric trucks [5]. To mitigate this problem, not only does an effective route planning mechanism need to be deployed (making sure the electric truck will reach a destination or a charging station), but also evidence showing energy sufficiency for the given task is needed [6]. In both cases, reliable and accurate prediction of energy consumption is essential. Therefore, a key feature for both sales teams and customers is the ability to forecast energy consumption for HD battery electric vehicles (BEV) to choose the optimal configuration, route planning, and charging activities.

The energy consumption of HD BEVs can generally be divided into two parts: propulsion energy and energy consumed by auxiliary components. The propulsion energy consumption is modeled using a dedicated physics-based simulation framework that has been extensively developed and validated for decades at Volvo. The other part, the auxiliary energy consumption in an HD BEV can normally be more than 10% of the total energy consumption or even more than 20% in some extreme conditions. A baseline approach for forecasting auxiliary energy consumption via basic statistics, e.g., mean and median values, yields an mean squared error (MSE) of around 16 kW, and results in approximately 50% error in forecasting the consumption. A data-driven model, empowered by machine learning algorithms, for auxiliary energy consumption, is expected to significantly improve the accuracy of predicting energy consumption for HD trucks, strengthening and speeding up the electrification in the transportation system. In this study, a comprehensive evaluation is conducted to benchmark the performance of a variety of well-known machine learning methods for forecasting auxiliary energy consumption of HD BEVs. Most energy consumption forecasting studies for vehicles focuses on personal/passenger cars[7] [8], electric buses[9] and commercial vehicles running on fossil fuel. To the authors' best knowledge, a comprehensive evaluation on data-driven approaches for HD electric truck fleets with real-world operation data are not available.

2 Challenge

The challenges encountered in this study on forecasting auxiliary energy consumption are three-fold. First and foremost, it is very costly to build dedicated forecasting methods, using physics-based models and expert knowledge, for each auxiliary system – the traditional approach is not cost-effective, and difficult to scale, as the auxiliary systems include multiple components, e.g., heaters, compressors, and power electronic converters. Moreover, these components are of different consumption characteristics and related to many signals, resembling a complex multivariate system. In addition, commercial vehicles operate in many different situations and perform many different tasks. As transportation tasks, ambient conditions, and drivers vary, so do the specifics of internal processes of the driveline, as well as auxiliary components. The possible circumstances that affect all the essential subsystems are too numerous to account for explicitly. Therefore, an AI-driven system for energy consumption forecasting that is capable of handling heterogeneous population, adapting to different types of transportation tasks, ambient condition etc., is needed.

The main strength of a data-driven approach is its capacity to model even highly complex and partially or fully unknown systems, provided that a sufficient amount of data is available. Chen et al., in their work [4], summarized that the data-driven methods have been the preferred choice for predicting BEV energy consumption, which is most likely because of the complexity in a system with many stochastic parameters, for example, driver behaviour and external conditions such as weather, road, traffic, etc., where a data-driven approach can uncover patterns and the complex relationships without the need to understand the inner system. In addition to providing more accurate and reliable energy consumption forecasts, these models can adapt to new data, continuously improving their predictions over time.

Even though huge amounts of data of HD trucks is available, it is heterogeneity which presents a significant challenge for data-driven models because it involves a vast array of variables that are highly variable and often interdependent. This heterogeneity arises from several factors, including diverse routes and terrains, varying driver behaviors, external conditions, different vehicle configurations, and various mission types. Because of these diverse and fluctuating factors, data-driven models struggle to accurately predict auxiliary energy consumption of electric trucks. The models must account for complex interactions between variables, which is challenging with heterogeneous data.

3 Methodology

3.1 Data Description

The dataset consists of signals transmitted via the CAN bus in the vehicle, including parameters such as speed, acceleration, road inclination and other vehicle operating signals. This multivariate time-series data was collected from a CAN logger installed in HD vehicles. The data spans 100 vehicles operating across more than 20 countries. The majority of the HD trucks in the dataset are equipped with 4x2 and 6x2 axle configurations, with a tractor type design. The battery capacity of each truck is 540 kWh (6 battery packs) by Lithium Nickel-Cobalt-Aluminum Oxide (NCA) technology. The dataset comprises over 200k data instances collected over more than 12 months.

The segmented trips refer to a single trip that is broken into distinct sections, either due to stops, changes in routes, or data collection intervals. For example, if a truck makes several deliveries or pauses during a longer trip in a day, each portion between stops is considered a segmented trip. Furthermore, the segmented trips are filtered to include only the periods when the vehicle is in driving mode, characterized by the key being on and the vehicle not connected to a charger. In addition, a subset is extracted, including 13k segmented trips collected from a homogeneous fleet of 19 vehicles, operating in the same area, and undertaking similar transportation missions.

3.2 Data Processing

The features are aggregated for each segmented trip. Subsequently, the data preprocessing involves cleaning outliers and null values, followed by feature selection using a correlation matrix. Finally, the data is normalized for training the ML models. It is worth mentioning that the auxiliary energy consumption is closely related to the elapsed time. Hence, auxiliary power consumption is used as a relevant metric. The power consumption of individual auxiliary components is summed to obtain the total auxiliary energy consumption.

The distribution of four features is depicted in Figure 1 and Figure 2. The distributions are illustrated using probability density functions, which are normalized histograms. The x-labels show the minimum and maximum range of the features. The ambient temperature follows a Gaussian distribution, whereas the vehicle speed and energy consumption present exponential distributions. It is notable that the weight distribution exhibits a bi-modal pattern, suggesting the presence of both loaded and unloaded trips. The left peak of the distribution corresponds to the truck's body weight, which ranges from 10 to 13 tonnes, depending on the configuration.

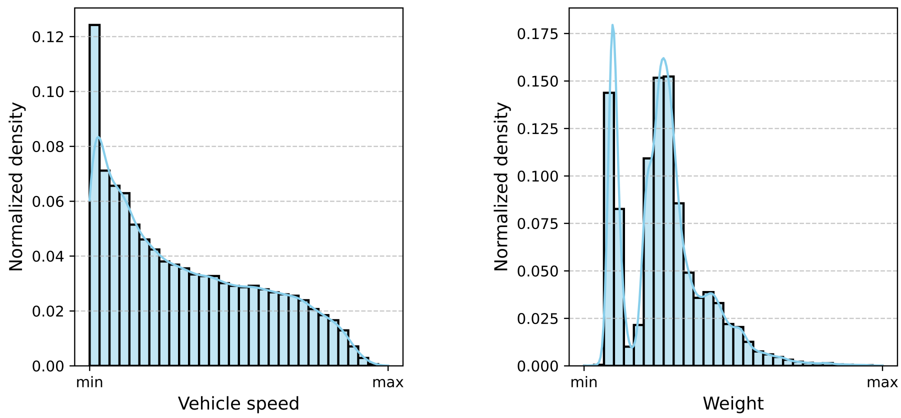


Figure 1: Distribution of vehicle speed and weight in the dataset

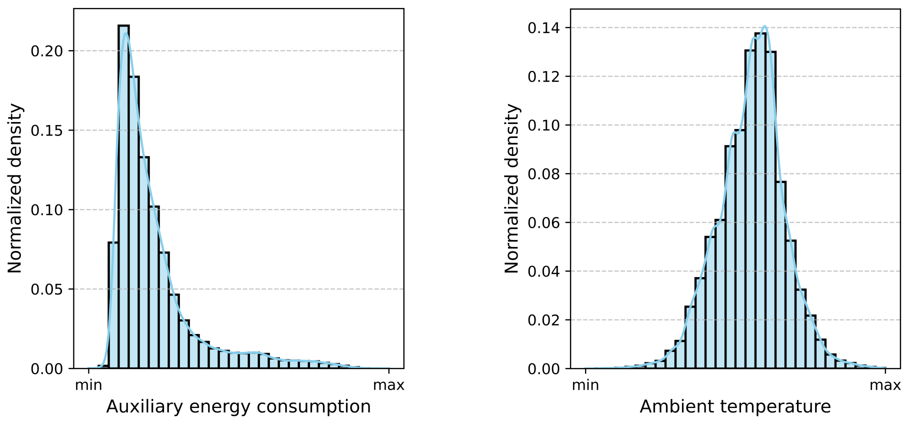


Figure 2: Distribution of total auxiliary energy consumption and ambient temperature in the dataset

3.3 Modelling

3.3.1 Expert Model

Since the auxiliary power consumption, to a very large extent, depends on the ambient temperature [10], an expert model has been developed based on aggregated trip segment data to capture this relationship, as depicted in Figure 3. The figure illustrates the dependency of auxiliary power on ambient temperature displaying an asymmetrical 'U' shape, as noted by Want et al.[10]. The auxiliary power consumption of the HD BEVs tends to be minimal within mid-range of ambient temperature, slightly skewed towards warmer temperatures. On the contrary, auxiliary power consumption increases largely at temperature extremes, particularly under lower ambient temperature conditions. This indicates the critical impact of heating and cooling systems on auxiliary power demands, especially for electric trucks.

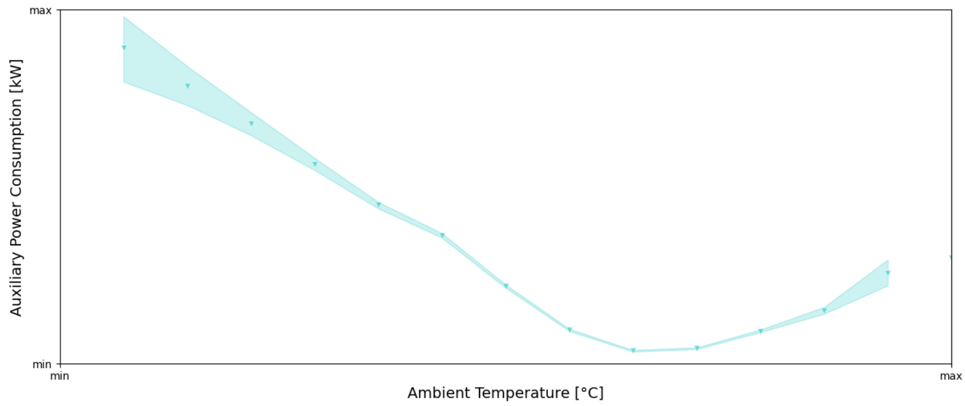


Figure 3: Auxiliary power consumption versus ambient temperature

3.3.2 Machine Learning Models

While auxiliary power consumption clearly depends significantly on ambient temperature, employing ML models can further enhance the analysis by integrating additional features and discovering more accurate correlations between these features and auxiliary power consumption. Therefore, classic ML models implemented in scikit-learn [11] were employed to identify and model these complex relations: i) Classic regression: Linear Regression and Polynomial Regression; ii) Regularized regression: Ridge, BayesianRidge and ElasticNet; iii) Instance-based learning: K-Nearest Neighbors (KNN); iv) Support Vector Regression (SVR); v) Tree-based Regression Algorithms: DecisionTree, RandomForest, Light Gradient Boosting Machine (LightGBM), Extreme Gradient Boosting(XGBoost); vi) Artificial Neural Network(ANN): Multilayer Perceptron(MLP).

A grid search approach was employed to identify the optimal hyperparameters for each ML model, utilizing five-fold cross-validation to ensure robustness and better generalization. Furthermore, feature selection was refined by ranking feature importance scores and retaining only those features with an importance greater than 0.05. Additional, all ML models were trained on each dataset.

A segmented trip can be further divided into smaller parts, referred to as micro-segments. The data of a segmented trip exhibits a sequential structure as it carries temporal information. Consequently, it is natural to think that power consumption in each micro-segment depends not only on the input signals of that micro-segment but also on the information from micro-segments that have come before it. Recurrent neural networks (RNN) are efficient in finding sequential relationships between time-series data instances [12]. Thus, an architecture with a long short-term memory (LSTM) [13] layer, followed by two fully connected layers, was used. In addition, a simpler recurrent architecture was also used for the sequential data. For example, gated recurrent units (GRU) combine the forget and input gates into a single update gate, reducing computational complexity while maintaining performance in sequence modeling [14].

The configuration of the LSTM model contains two hidden layers with size of 64 units on each, followed by two full connected layers also with size of 64 units based on pytorch library [15]. The GRU model follows a similar architecture, i.e. two GRU layers with 64 units each and two full connected layers. In addition, the MSE loss function was utilized, and the models were optimized using the Adam optimizer. During training, the best-performing model was selected and saved based on the lowest validation error criterion. A ReduceLROnPlateau scheduler was employed to dynamically adjust the learning rate according to the validation loss, using a patience parameter of 15 epochs.

3.3.3 Evaluation Metrics

In this study, mean absolute error (MAE), MSE, and a proposed accuracy measure (based on the sum of absolute errors for each forecast divided by the total consumption of each segmented trip) were selected to quantify the performance of forecasting methods. Mathematically MAE is defined as in equation (1), where n represents the total number of segmented trips, y_i is the actual value for the i-th segmented trip, $\hat{y}_i$ is the predicted value for the i-th segmented trip, and $|\cdot|$ denotes the absolute value.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{1}$$

MSE is defined as in equation (2), where n represents the total number of segmented trips, y_i is the actual value for the i-th segmented trip, and $\hat{y}_i$ is the predicted value for the i-th segmented trip.

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{2}$$

The proposed accuracy metric evaluates the forecasting performance by normalized the sum of absolute errors with the total consumption. This measure is similar to mean absolute percentage error (MAPE) and mean absolute range normalized error (MARNE) mentioned in the review [16], but with a different scaling factor. MAPE penalizes forecasting errors at low consumption more than errors at high consumption, which is unsuitable for this application, given that low energy consumption periods are less critical to the operation of BEVs than high consumption periods. Additionally, MARNE normalizes the error by the maximum forecasting value, which may lead to misleading results as well.

Thus, we have chosen to scale the error based on the mean consumption. It is calculated in equation (3), where n represents the total number of segmented trips, y_i is the actual value for the i-th segmented trip, and $\hat{y}_i$ is the predicted value for the i-th segmented trip.

$$\text{Accuray} = \left(1 - \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{\sum_{i=1}^{n} y_i}\right) \times 100 \tag{3}$$

3.4 Experimental Settings

Table 1 summarizes the dataset characteristics and forecasting methodologies in 9 experiment scenarios. The first three scenarios utilize aggregated data from segmented trips, where the models predict the average power consumption for each segmented trip in the test dataset. Scenario 4 replicates the methodologies of the first three scenarios using a subset of data from a fleet of 19 vehicles. Scenarios 5 to 7 operate at a higher temporal resolution using micro-segments, which are multivariate time series data of each trip segment. Scenario 5 (Mean of a sliding window) employs a conventional energy forecasting strategy assuming future consumption is consistent with the past [17]. We use the mean consumption value of the last segment as the forecast for the subsequent segment. Trip segments shorter than the required sliding window length plus the prediction horizon are excluded. The sliding window moves along the time axis with a stride equal to the prediction horizon, continuing until the remaining data at the end of a trip is shorter than the horizon. In such cases, the final forecast values are set to the last value from the most recent prediction. For each sliding window, the model forecasts the next set of values within the defined prediction horizon (scenario 6 and 7), of which the inputs are always based on ground truth. This setup reflects multi-step direct forecasting with clean, non-recursive inputs. The last two scenarios (8 and 9) simulate a more practical forecasting scenario, where only the initial portion of the time series is available. Scenario 8 uses the mean of this initial window as the forecast for the rest of the trip, while Scenario 9 employs recursive forecasting using RNN/LSTM models, in which predictions from each step are fed back as inputs for subsequent forecasts. This approach reflects real-world conditions better, where only partial trip data may be available at a prediction time. A 5-fold cross-validation strategy is employed to ensure robust model evaluation and to provide a reliable estimate of the models' generalization performance. The dataset is split into 80% training and 20% testing.

4 Results and Discussion

4.1 Expert and ML Models

The metrics in the result tables are calculated based on normalized data, and the proposed accuracy measure is determined by applying the inverse transformation of the scalar in relation to the mean value of auxiliary energy consumption. In addition, the time column in the table indicates the duration for training a model.

Table 1: Experimental settings.

Scenario	Dataset	Method
1	Segmented trip aggregation	Global mean
2	Segmented trip aggregation	Expert model
3	Segmented trip aggregation	Classic ML models
4	A subset of segmented trip aggregation	Scenario 1 - 3
5	Time series data	Mean of a sliding window
6	Time series data	RandomForest model prediction of a sliding window
7	Time series data	RNN/LSTM models
8	Initial period of time series data	Mean of the initial window
9	Initial period of time series data	Recursive forecasting of RNN/LSTM models

As expected, the first scenario yields the largest prediction error. Due to the complexity and multivariate nature of the auxiliary system consumption, using the global mean of all segmented trips, as a baseline, yields an accuracy of approximately 55%.

Performance metrics for the Expert and various ML models for Scenarios 2 and 3, as shown in Table 1, are outlined in Table 2. The Expert model, illustrated in Figure 3, achieves slightly improved accuracy (64%) by clustering auxiliary power consumption based on ambient temperature. In addition, classic regression and regularized regression models present performance similar to the expert model, reflecting the nonlinear and complicated nature of auxiliary power consumption. The KNN model performs marginally better than classic regression, while the MLP, SVR and tree-based regression algorithms draw the highest accuracy, reaching approximately 82%.

Additional evaluations conducted on the previously described sub-dataset are presented in Table 3. The results show that accuracy improves when using the sub-dataset, which possesses better homogeneity. Especially, the expert model, classic regression and regularized regression models demonstrate an approximate 8% increase in accuracy. Furthermore, the highest accuracy achieved by the same models when applied to the full dataset with a larger number of vehicles reaches 84%, as shown in 2. Therefore, these models are capable of more effectively capturing the average auxiliary power consumption of each segmented trip.

However, these predictions apply to entire segmented trips where aggregated data simplifies model training, reducing complexity and noise, with training times generally under one hour. Furthermore, training on granular time-series data can result in more accurate predictions, as the model is able to capture the nuances of the data. The performances of these RNN models for Scenarios 5 and 7 are compared in Table 4.

Table 2: Performance metrics for the Scenario 1-3.

Model	MAE	MSE	Accuracy	Time
Global mean	0.6950 ± 0.0020	1.0000 ± 0.0050	0.5579 ± 0.0011	<1 second
Expert model	0.6802 ± 0.0050	0.9670 ± 0.0095	0.6377 ± 0.0065	<1 second
LinearRegression	0.5206 ± 0.0020	0.4742 ± 0.0062	0.6689 ± 0.0013	<10 seconds
Polynomial Regression	0.3392 ± 0.0036	0.2514 ± 0.0057	0.7843 ± 0.0023	<16 minutes
Ridge	0.5206 ± 0.0020	0.4742 ± 0.0062	0.6689 ± 0.0013	<10 seconds
BayesianRegression	0.5206 ± 0.0020	0.4742 ± 0.0062	0.6689 ± 0.0013	<10 seconds
ElasticNet	0.5337 ± 0.0025	0.5068 ± 0.0083	0.6614 ± 0.0025	<3 minutes
KNN	0.3228 ± 0.0025	0.2512 ± 0.0056	0.7946 ± 0.0016	<2 minutes
SVR	0.3310 ± 0.0018	0.2221 ± 0.0032	**0.8128 ± 0.0017**	<40 minutes
DecisionTree	0.3312 ± 0.0025	0.2510 ± 0.0055	0.7894 ± 0.0016	<2 seconds
RandomForest	0.3156 ± 0.0018	0.2188 ± 0.0042	0.7993 ± 0.0010	<3 minutes
LightGBM	0.2890 ± 0.0019	0.1899 ± 0.0041	**0.8162 ± 0.0012**	<14 seconds
XGBoost	0.3515 ± 0.0023	0.2572 ± 0.0045	0.7764 ± 0.0014	<6 seconds
MLP	0.2961 ± 0.0029	0.1997 ± 0.0044	**0.8115 ± 0.0017**	<1 minute

4.2 Forecast Models with a Sliding Window

Two model configurations for time-series forecasting are summarized in Table 4. In the first configuration, a sliding window of 10 samples was used with a forecasting horizon of 10 samples, which generates over 700k distinct sequences with no overlap between trips or vehicles. In the second configuration, the sliding window

Table 3: Performance metrics for the Scenario 4.

Model	MAE	MSE	Accuracy	Time
Global mean	0.7395 ± 0.0110	1.0000 ± 0.0740	0.5911 ± 0.0063	<1 second
Expert model	0.7365 ± 0.0131	0.9925 ± 0.0407	0.7224 ± 0.0133	<1 second
LinearRegression	0.4609 ± 0.0056	0.3717 ± 0.0072	0.7430 ± 0.0031	<4 seconds
Polynomial Regression	0.3157 ± 0.0050	0.2520 ± 0.0899	0.8240 ± 0.0028	<2 minutes
Ridge	0.4609 ± 0.0056	0.3717 ± 0.0072	0.7430 ± 0.0031	<5 seconds
BayesianRegression	0.4610 ± 0.0056	0.3717 ± 0.0072	0.7430 ± 0.0031	<5 seconds
ElasticNet	0.4714 ± 0.0052	0.3892 ± 0.0072	0.7376 ± 0.0030	<3 minutes
KNN	0.3319 ± 0.0057	0.2056 ± 0.0074	0.8150 ± 0.0032	<8 seconds
SVR	0.3155 ± 0.0023	0.1735 ± 0.0034	**0.8407 ± 0.0011**	<1 minute
DecisionTree	0.3348 ± 0.0034	0.2171 ± 0.0074	0.8133 ± 0.0020	<1 second
RandomForest	0.3062 ± 0.0033	0.1743 ± 0.0042	0.8292 ± 0.0018	<30 seconds
LightGBM	0.2864 ± 0.0039	0.1552 ± 0.0044	**0.8404 ± 0.0021**	< 8 seconds
XGBoost	0.3160 ± 0.0041	0.1872 ± 0.0056	0.8238 ± 0.0022	< 3 seconds
MLP	0.2936 ± 0.0020	0.1607 ± 0.0034	**0.8370 ± 0.0023**	< 30 seconds

was extended to 30 samples with a forecasting horizon of 5 samples, resulting in approximately 550k distinct non-overlapping sequences. For each sequence, forecasts were made within the trip by sliding the window, and predictions were averaged to represent the trip-level auxiliary power consumption. These predictions were then compared to the ground truth mean values to compute performance metrics.

In the first configuration, the RandomForest model achieves an accuracy of 87%, outperforming the best model with aggregation data of each segmented trips by 3%. Additionally, the conventional method of forecasting the next segment by applying the mean consumption value from the last segment yields results comparable to those of the RandomForest model.

When applying RNN models such as LSTM or GRU, performance further improves due to the model's ability to capture temporal dependencies in the data. The LSTM model with the second configuration accomplishes the highest accuracy of over 95%, effectively forecasting the next 5 samples from a window of 30 samples. Since the GRU model exhibits approximately 1% lower performance and 2 additional hours of training time compared to the LSTM model, its results are not reported, and it was not selected for subsequent tasks. Despite the high accuracy of RNN models, these models requires the longest training time which exceeds 5 hours using GPUs. For larger datasets, training time could increase significantly. Additionally, it's important to note that the training process requires access to ground truth input features at each step of sliding window, which may not be available in real-world deployment.

Table 4: Performance metrics for Scenarios 5-7. *SW* and *PH* correspond to the length of the sliding window and the size of the prediction horizon

$\text{Model}_{\{SW,PH\}}$	MAE	MSE	Accuracy	Time
$\text{Mean}_{\{10,10\}}$	0.2145 ± 0.0080	0.2469 ± 0.0199	0.8742 ± 0.0047	<1 second
$\text{RandomForest}_{\{10,10\}}$	0.1970 ± 0.0073	0.4438 ± 0.0269	0.8712 ± 0.0046	<10 minutes
$\text{LSTM}_{\{10,10\}}$	0.0985 ± 0.0028	0.0301 ± 0.0031	0.9422 ± 0.0016	>6 hours
$\text{Mean}_{\{30,5\}}$	0.1326 ± 0.0042	0.0641 ± 0.0062	0.9223 ± 0.0024	<1 second
$\text{LSTM}_{\{30,5\}}$	0.0808 ± 0.0011	0.0209 ± 0.0015	**0.9526 ± 0.0006**	>5 hours

4.3 Forecasting with a Initial Window

Considering limited computation resources on the edge devices, consumption forecasts are likely made less frequently and rely primarily on the data from the initial period of a trip. Results from such a deployment scenario are shown in Table 5. The accuracy of forecasting the rest of the trip based on the first window of data varies significantly with window size. A 30-sample window achieves 77% accuracy using the window's mean value, while the LSTM model obtains 88% accuracy by learning the time-series patterns in the initial window and recursively predicting future values.

It is worth mentioning that the LSTM architecture remained consistent with that used in the sliding window forecasting experiments. Model tuning could potentially yield even better performance.

In summary, to effectively forecast future auxiliary power consumption during vehicle operation, clustering vehicles by configuration, mission, and geographical location into homogeneous groups/sub-fleets, and conducting modeling on individual groups is expected to improve performance. Using the global mean consumption yields 59% accuracy. For representing the remaining of the trip with the mean of the first 30-sample window achieves 77%. Furthermore, deploying an LSTM model can raise performance to 88%, in the pursuit of a higher accuracy.

Table 5: Performance metrics for Scenarios 8 and 9. *IW* corresponds to the length of the initial window while *RW* corresponds to the length of the recursive window for LSTM models.

Model$_{\{IW,RW\}}$	MAE	MSE	Accuracy	Time
Mean$_{\{10,-\}}$	0.8986 ± 0.0179	1.7347 ± 0.0615	0.4731 ± 0.0105	<1 second
LSTM$_{\{10,10\}}$	0.2322 ± 0.0075	0.1289 ± 0.0074	0.8638 ± 0.0044	>6 hours
Mean$_{\{30,-\}}$	0.3817 ± 0.0070	0.3334 ± 0.0175	0.7762 ± 0.0041	<1 second
LSTM$_{\{30,5\}}$	0.1989 ± 0.0060	0.0993 ± 0.0062	**0.8834 ± 0.0035**	>5 hours

4.4 Model Trade-offs

Figure 4 illustrates the trade-off between model performance and resource requirements, including training time and model size. In the bottom-left region of the figure, models based on aggregations of segmented trip data are presented. These models generally achieve training times below 100 seconds, although their model sizes vary significantly. In particular, the RandomForest models require the most memory, exceeding 2 MB, whereas most other models remain around 1 MB.

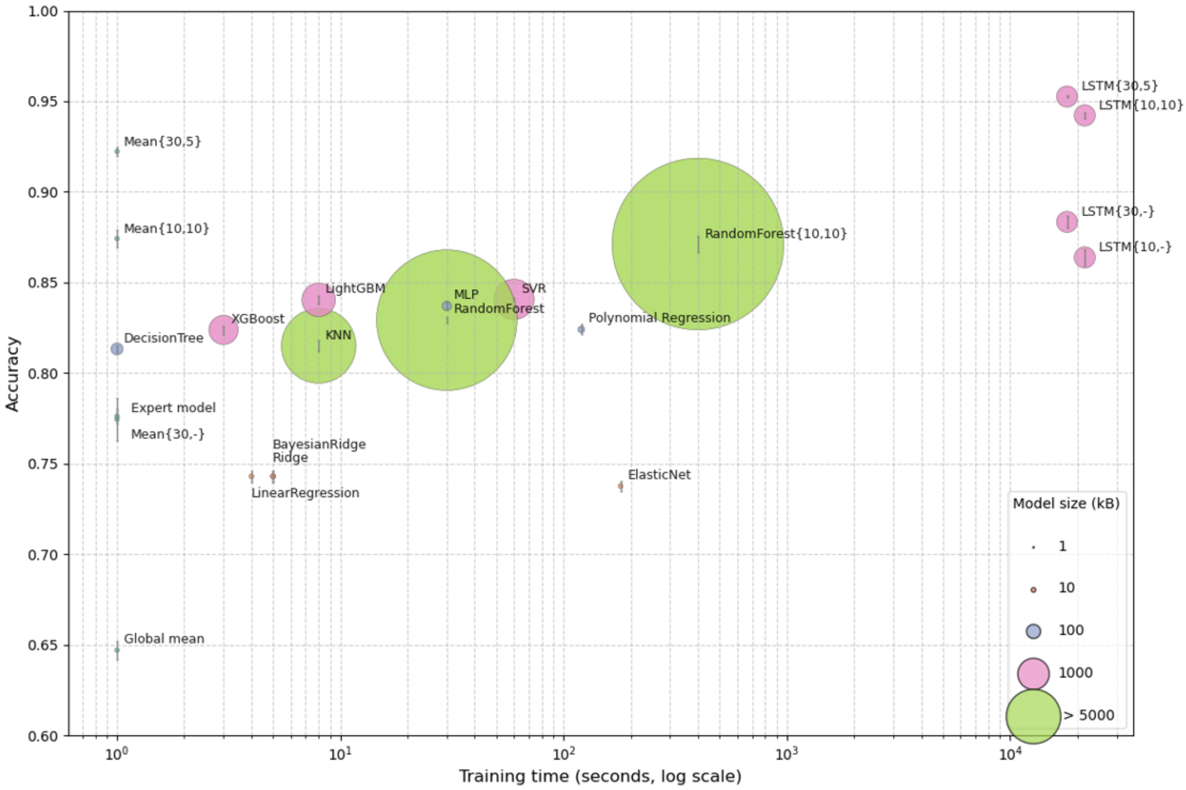


Figure 4: Model trade-offs between performance and resources, training time and memory, required

When high-resolution time-series data is utilized, better forecasting accuracy can be achieved. For example, the top-right region of Figure 4 highlights models such as LSTM$_{\{30,5\}}$, which deliver the highest accuracy but require significantly longer training times. Although these models maintain moderate sizes (around 1 MB), the computational cost during training is substantial. When models with lower training computational cost and small model sizes are preferred, approaches such as using the mean of samples with a sliding window (e.g., the Mean$_{\{30,5\}}$ model) offer an effective alternative. The models achieve relatively high accuracy, between 87% and 92%, while keeping the model compact. However, they are limited to forecasting very short time horizons, such as five samples. When attempting to forecast the entire remaining segment of a trip, the performance falls below that of the classic ML models in Scenario 6 or the expert model, which uses only ambient temperature as the input feature for auxiliary power consumption prediction. On the contrary, LSTM models trained under Scenario 7, which employs sliding windows and fresh ground truth inputs, show strong performance. Taking the first 30 samples as the input, the LSTM model can forecast the remaining duration of the trip with 88% accuracy. Therefore, if sufficient resources are available for training, LSTM models are recommended for edge deployment to achieve high forecasting accuracy.

5 Conclusion

This study presents a comprehensive analysis of 100 HD BEVs operating over more than 12 months in more than 20 countries. It explores data-driven modeling approaches for forecasting auxiliary energy consumption, which can account for up to 20% of total vehicle energy use. Both aggregated segmented trip data and high-resolution time-series data are used for forecasting. To determine the most suitable modeling approach, a extensive evaluation is conducted to benchmark the performance of various well-known ML models.

The simplest forecasting approach, based on the global mean consumption value, yields the lowest accuracy, around 50%. Incorporating an expert model that correlates auxiliary energy consumption with ambient temperature improves accuracy by approximately 10%. Among classical ML models, SVR, MLP, and tree-based regression algorithms obtain the highest accuracy, reaching around 82% with the aggregation data of each segmented trip.

Model performance further improves when training on subsets consisting of homogeneous fleets of vehicles, selected according to configuration, mission profile, and geographical location. This approach generates more consistent datasets and improves forecast accuracy up to 84%. The LSTM model, particularly when using high-temporal-resolution data, achieves the best overall performance, exceeding 95% accuracy, albeit with a higher computational cost. However, even simpler models, such as using the mean of samples within a sliding window (e.g., the $\text{Mean}_{\{10,10\}}$ and $\text{Mean}_{\{30,5\}}$ models), can achieve satisfying accuracy levels between 87% and 92% when taking advantage of high-resolution time-series data. This performance is about 8% higher than the best models based on aggregated segmented trip data.

For edge deployment scenarios, i.e., where only data from the initial period of a trip is available, forecasting the remaining consumption of the trips based on the mean consumption from a 30-sample window achieves an accuracy of 77% accuracy. When leveraging LSTM models with the same setting, accuracy improves to 88%, with the models training under Scenario 7, which requires considerable computational cost. Thus, if sufficient resources are available for training, LSTM models are recommended for edge deployment to achieve high forecasting accuracy.

In conclusion, this work outlines a practical framework for predicting and forecasting auxiliary energy consumption using data-driven models, with a thorough investigation on the trade-off between model accuracy, size, and computational efforts. The findings offer a solid foundation for evaluating and structuring forecasting methods while highlighting the limitations of primitive approaches, e.g., with constant values such as global mean consumption. This balance between accuracy and deployment feasibility is crucial for real-world applications.

Acknowledgments

The work has been carried out with support from the Knowledge Foundation and Vinnova (Sweden's innovation agency) through the Vehicle Strategic Research and Innovation Programme FFI.

References

[1] MarketsandMarkets, "Electric Truck Market by Propulsion (BEV, PHEV, FCEV), Type (Light-Duty Trucks, Medium-Duty Trucks, Heavy-Duty Trucks), Range, Battery Type, Battery Capacity," available at: https://www.marketsandmarkets.com/Market-Reports/electric-truck-market-221011937.html.

[2] Alanazi, Fayez. 2023. "Electric Vehicles: Benefits, Challenges, and Potential Solutions for Widespread Adaptation." Applied Sciences 13, no. 10: 6016. https://doi.org/10.3390/app13106016.

[3] Volvo Trucks, "Volvo's electric truck in first independent efficiency test," available at: https://www.volvotrucks.com/en-en/news/2022/jan/volvo-electric-truck-efficiency-test.html.

[4] Chen, Y., Wu, G., Sun, R., Dubey, A. et al., "A Review and Outlook on Energy Consumption Estimation Models for Electric Vehicles," SAE J. STEEP 2(1):79-96, 2021, https://doi.org/10.4271/13-02-01-0005.

[5] Smuts M, Scholtz B, Wesson J. A critical review of factors influencing the remaining driving range of electric vehicles. IEEE 2017 1st International Conference on Next Generation Computing Applications (NextComp) 2017;196-201.

[6] Li W, Long R, Chen H, Geng J. A review of factors influencing consumer intentions to adopt battery electric vehicles. Renewable and Sustainable Energy Reviews 2017;78:318-28.

[7] Schäfers L, Franke K, Savelberg R, Pischinger S. Auxiliaries' power and energy demand prediction of battery electric vehicles using system identification and deep learning. IET Intelligent Transport Systems 2023;17(2):123–132. https://doi.org/10.1049/itr2.12467.

[8] Achariyaviriya W, Wongsapai W, Janpoom K, Katongtung T, Mona Y, Tippayawong N, Suttakul P. Estimating energy consumption of battery electric vehicles using vehicle sensor data and machine learning approaches. Energies 2023;16(17):6351. https://doi.org/10.3390/en16176351.

[9] Zhang X, Zhang Z, Liu Y, Xu Z, Qu X. A review of machine learning approaches for electric vehicle energy consumption modelling in urban transportation. Renewable Energy 2024;234:121243. https://doi.org/10.1016/j.renene.2024.121243.

[10] Wang J. B., Liu K., Yamamoto T., Morikawa T. Improving estimation accuracy for electric vehicle energy consumption considering the effects of ambient temperature. Energy Procedia 2017;105:2904-9. 8th International Conference on Applied Energy, ICAE2016; 8-11 October 2016; Beijing, China. Available from: https://www.sciencedirect.com/science/article/pii/S1876610217307099

[11] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[12] Jain, L.C., Medsker, L.R.: Recurrent Neural Networks: Design and Applications. CRC Press, Boca Raton, FL, USA (1999)

[13] Hochreiter, S., Schmidhuber, J.: Long Short-Term Memory. Neural Computation 9(8), 1735–1780 (1997). doi: 10.1162/neco.1997.9.8.1735

[14] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 1724–1734, 2014.

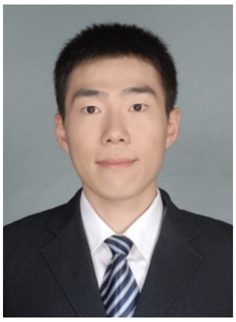
[15] Paszke, Adam and Gross, Sam and Chintala, Soumith and Chanan, Gregory and Yang, Edward and DeVito, Zachary and Lin, Zeming and Desmaison, Alban and Antiga, Luca and Lerer, Adam. Automatic differentiation in PyTorch, 2017.

[16] Wei, Nan, et al. ”Conventional models and artificial intelligence-based models for energy consumption forecasting: A review.” Journal of Petroleum Science and Engineering 181 (2019): 106187.

[17] Zhang, Jin, Zhenpo Wang, Peng Liu, and Zhaosheng Zhang. ”Energy consumption analysis and prediction of electric vehicles based on real-world driving data.” Applied Energy 275 (2020): 115408.

[18] Ahn, Hyunjin, Heran Shen, Xingyu Zhou, Yung-Chi Kung, John Maweu, and Junmin Wang. ”Velocity and energy consumption prediction of medium-duty electric trucks considering road features and traffic conditions.” Journal of Dynamic Systems, Measurement, and Control 146, no. 6 (2024).

Presenter Biography



Zhenkan Wang, PhD, is a Data Scientist specializing in simulation, machine learning, and data analysis at GTT, Volvo Group. Over the past 5 years, he has gained experience in R&D, innovation, advanced engineering, and product development within GTT and PSD.

Beforehand he gathered 10 years of academic experience within R&D of automotive engineering including advanced measurements and simulations focusing on improving energy efficiency for passenger and commercial vehicles. His research has been widely published in high-impact journals, highlighting his expertise in applying advanced techniques to complex physical phenomena, particularly in energy systems and diagnostics.