

Development of an EV HVAC Power Consumption Prediction Model Based on Multivariate Analysis and Machine Learning of Probe Data

Koichi Hasegawa Haruhisa Tsuchikawa Hiroki Matsui Haruki Fukui

*Nissan Motor Co., Ltd., Nissan Technical Center
560-2 Okatsukoku, Atsugi-shi, Kanagawa, 243-0192, Japan
(E-mail: hasegawa_ko@mail.nissan.co.jp)*

KEY WORDS: EV and HV systems, Electrical accessories/air conditioning, Energy management[A3]

To relieve the range anxiety EV users feel on long-distance trips, Nissan equips its navigation system with the Intelligent Route Planner (IRP), which predicts the state of charge (SOC) at the destination by estimating the energy consumed along the route. While traction energy is largely determined by vehicle characteristics and driving conditions, HVAC power varies considerably with ambient temperature, weather, and occupant usage. Prediction adapted to real-world conditions is therefore essential for further improving arrival SOC accuracy.

A factor analysis of the IRP arrival SOC prediction error identified four main contributors: deviation between map-based and actual vehicle speed, the difference between design and on-road rolling resistance, variation in aerodynamic drag caused by wind, and HVAC power prediction error. HVAC power accounts for 14.5% of the error in highway scenes. In congested scenes, traction energy is small because vehicle speed is low, so the relative share of HVAC power reaches 80%.

Probe data recorded by in-market Nissan EVs for each trip and accumulated on a server were used. Trips with the HVAC operating recorded in Japan during 2024 were targeted, and 100,000 trips were extracted. The objective variable is the trip-average power [W] of the high-voltage HVAC components. The data were randomly split into 70,000 training and 30,000 evaluation trips. Ambient temperature is generally known as the dominant feature, since a larger difference between cabin and set temperature increases HVAC power. To clarify other contributions, a provisional model was built using the probe data labels as features, and SHAP (SHapley Additive exPlanations) values were applied to its predictions. The contribution was largest for ambient temperature, followed by trip time, month, departure time, solar radiation, humidity, and average vehicle speed. Each contribution can be explained physically; short trips, for example, have a higher proportion of the initial boarding period with large HVAC load. These seven items were selected as the features of the prediction model.

Six supervised regression methods (LightGBM, neural network, random forest, k-NN, linear regression, and decision tree) were compared by R^2 , training time, and prediction time. LightGBM and the neural network were tied at the top for R^2 , while the training time of LightGBM was about 1/90 of that of the neural network. Assuming periodic retraining as probe data accumulate, LightGBM was adopted.

The model was compared with the conventional method, a two-dimensional HVAC power map based on ambient temperature and vehicle speed (Table 1). RMSE was reduced from 487 W to 378 W (22%) and MAE from 319 W to 254 W (20%), and R^2 improved from 0.24 to 0.55.

The prediction error, defined as predicted minus measured value, was analyzed for low ($<10^\circ\text{C}$), intermediate, and high ($\geq 25^\circ\text{C}$) temperature conditions. The conventional method has its error distribution centered on the underprediction side, which makes IRP predict a higher arrival SOC than actual and should be reduced to avoid battery depletion. The proposed model centers the distribution on zero and nearly eliminates this bias; under the low-temperature heating condition the 1σ point on the underprediction side was reduced from -917 W to -533 W (Fig.1). MBD verification confirmed that the arrival SOC error attributable to HVAC power is halved in both highway and congested scenes. Cumulative feature addition showed R^2 of 0.40 with ambient temperature alone and 0.55 with all seven features; departure time (+0.072), trip time (+0.040), and solar radiation (+0.021) contributed most. The improvement stems not from machine learning itself but from systematically extracting the features affecting the HVAC load and expressing them with a decision tree model. The model will be implemented in IRP to support confident long-distance driving.

Table1 Prediction accuracy:
current method vs. LightGBM

	Current	LightGBM
RMSE [W]	487	378
MAE [W]	319	254
$R^2(\text{evaluation})$	0.24	0.55
$R^2(\text{training})$	—	0.59

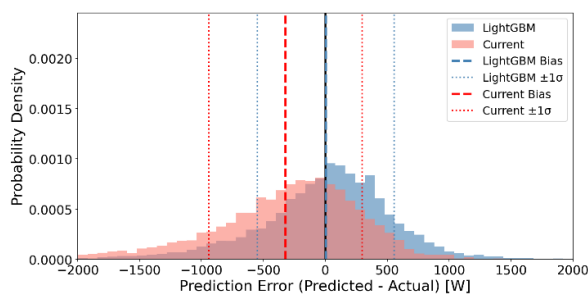


Fig.1 Prediction error distribution:
LightGBM vs. current method ($<10^\circ\text{C}$)