Proceedings of the Sixth International Conference on
Railway Technology: Research, Development and Maintenance
Edited by: J. Pombo
Civil-Comp Conferences, Volume 7, Paper 12.6
Civil-Comp Press, Edinburgh, United Kingdom, 2024
ISSN: 2753-3239, doi: 10.4203/ccc.7.12.6

Research Methods on Reducing Power Consumption Through Big Data Analysis in the Railway Field

**T. Fujimasu, D. Horikoshi, M. Yoshizawa, M. Adachi,
H. Moriyama and A. Hibino**

**Technology Research and Development Department, Central Japan Railway Company
Japan**

Abstract

The aim of this study is to achieve a reduction in power consumption of the Tokaido Shinkansen by deriving the characteristics of operational practices that lead to lower power consumption from a large amount of field data. A method was developed to efficiently calculate and visualize the power consumption using each running data, for understanding the trends in power consumption with respect to travel time. To accurately evaluate operational practices regarding power consumption while eliminating the influence of ambient temperature and occupancy rate, another method was established to correct the effects by those factors on computed power consumption, using an equation that calculates running resistance based on actual running data. Additionally, clustering analyses were performed on notch operations to effectively understand common characteristics from various running patterns for grasping the characteristics that realize a reduction in power consumption. By combining these approaches, it became possible to derive running methods that result in lower power consumption from a large amount of field data.

Keywords: big data, rolling stock, carbon neutral, driving operation, machine learning, data driven.

1 Introduction

In order to achieve carbon neutrality by 2050, efforts are being made in the railway sector to improve energy efficiency. The Tokaido Shinkansen, operated by Central Japan Railway Company, is a high-speed railway that connects Tokyo and Shin-Osaka over a distance of 515km at a maximum speed of 285km/h. It is characterized by its safety, high frequency, and punctuality. Since its operations commenced in 1964, there have been no accidents resulting in fatalities or injuries of passengers on board. We continuously strive for the development and deployment of rolling stock, with a focus on reducing weight, lowering aerodynamic resistance, and improving the efficiency of equipment. As a result of these efforts, we have achieved a reduction of approximately 32% in unit energy consumption (energy required to operate one vehicle for one kilometer) compared to the 1990 fiscal year, as of the end of the fiscal year 2021 [1].

Existing research on energy reduction in the railway sector includes the development of models for simulating power consumption [2], studies on deriving optimal driving methods using genetic algorithms for energy optimization [3], research on predicting driving curves and power consumption to support driving [4], and studies on generating energy-efficient train timetables [5].

While various theoretical methods have been proposed, there are not many research examples that utilize actual large-scale running data to aim for energy reduction. Our objective is to construct a methodology that utilizes the large amount of field data obtained from the Tokaido Shinkansen through analyses in pursuit of operational methods for further energy savings.

2 Visual Analysis of Power Consumption

2.1 Ensuring Consistency of Method to Calculate Power Consumption Using Vehicle Data

The vehicles on the Tokaido Shinkansen accumulate a large amount of running data from various equipment. Among this data, we have decided to calculate power consumption by using the current and voltage values of the traction circuit.

Figure 1. illustrates the time variation of the supplied power from a certain substation (measured at the substation) and the total power consumption of all trains running within the coverage of that substation (calculated from vehicle data). In the time period shown in Figure 1., the actual power supplied by the substation and the total power consumption of the trains generally match in terms of their time variation. This confirms the validity of the method to calculate power consumption using vehicle data . Hereafter, various analyses will be conducted based on the measurement data obtained from the vehicles.

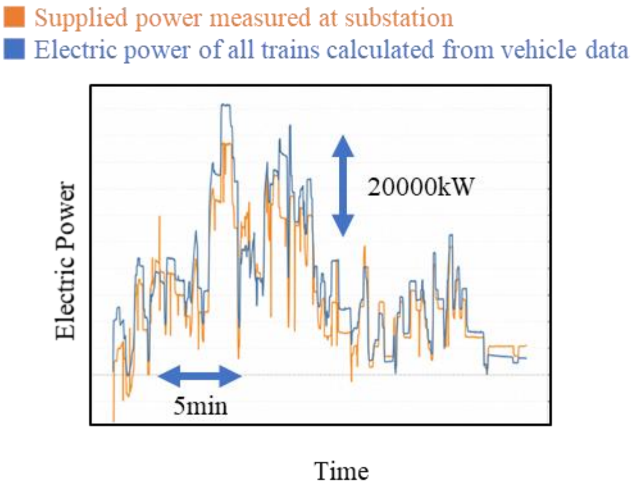


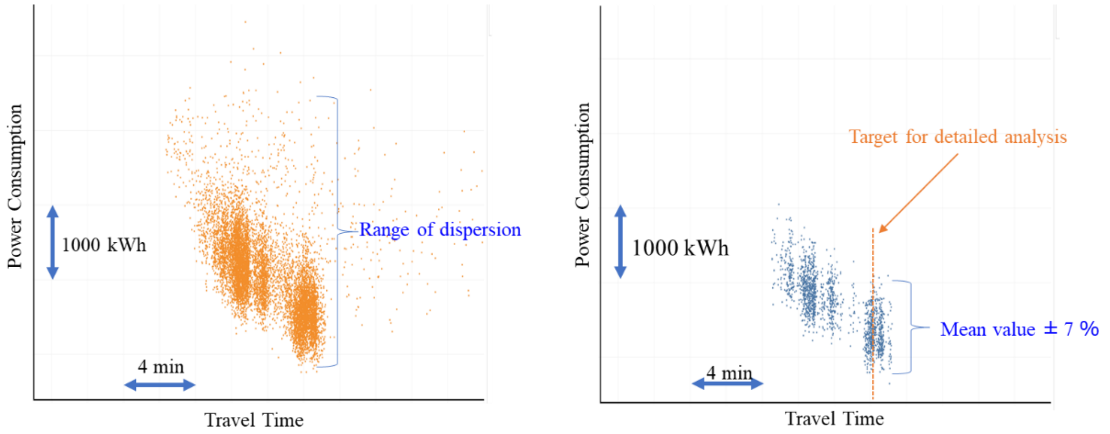
Figure 1: Supplied power from substation and total electric power of trains.

2.2 Trends in Power Consumption between Stations

Improving the operation methods of the Tokaido Shinkansen to reduce power consumption requires an assessment of trends in the power consumption of each train. To ensure consistent running conditions for data evaluation, analyses were conducted on data from trains that ran on the same timetable and had almost identical travel times (within a range of about half a minute). As power consumption generally declines as travel time increases, it is necessary to compare the power consumption of trains that have nearly the same travel time in a specific section. For a preprocessing step, vehicle data such as speed, notch, distance, current, and voltage were combined with timetable information, weather data, and occupancy rates. The combined data was then aggregated by station and evaluated for delay time.

The left chart of Figure 2. shows an example of analysis on 9,027 trains while running between stations A and F using the same timetable. The horizontal axis represents the travel time from station A to station F, and the vertical axis represents power consumption. It was observed that power consumption declines as travel time increases. Additionally, it was found that even for the same travel time, there are differences in power consumption.

The right chart of Figure 2. illustrates an example of analysis on 1,676 trains with delayed ones excluded (behind the schedule by more than 15 seconds). As a result of aligning the running conditions, the range of dispersion compared to the left chart is smaller, however, there are differences in power consumption even for the same travel time. Therefore, detailed analysis of the data indicated by the orange line in right chart of Figure 2. will be discussed in the next section.



Before excluding delayed trains

After excluding delayed trains

Figure 2: Trends in power consumption between stations A and F.

2.3 Case Study of Analysis on Running Performance

Figure 3. depicts an analysis of 209 trains which were referred in the previous section (running on the same timetable with delayed trains excluded). The horizontal axis represents the distance in kilometers between Stations B and F. The vertical axis represents both the speed and power consumption, where the power consumption is set to zero upon departure from Station A.

The darker the green color, the lower the power consumption, while the darker the red color, the higher the power consumption. The figure indicates that power consumption varies despite having the same travel time. This can be attributed to the difference by individual trains in running speeds at each location. Assuming that the variations of power consumption are due to driving operations, a detailed analysis was conducted by taking cases of smallest and largest power consumption among the 209 trains. The results of this analysis are presented in the next section.

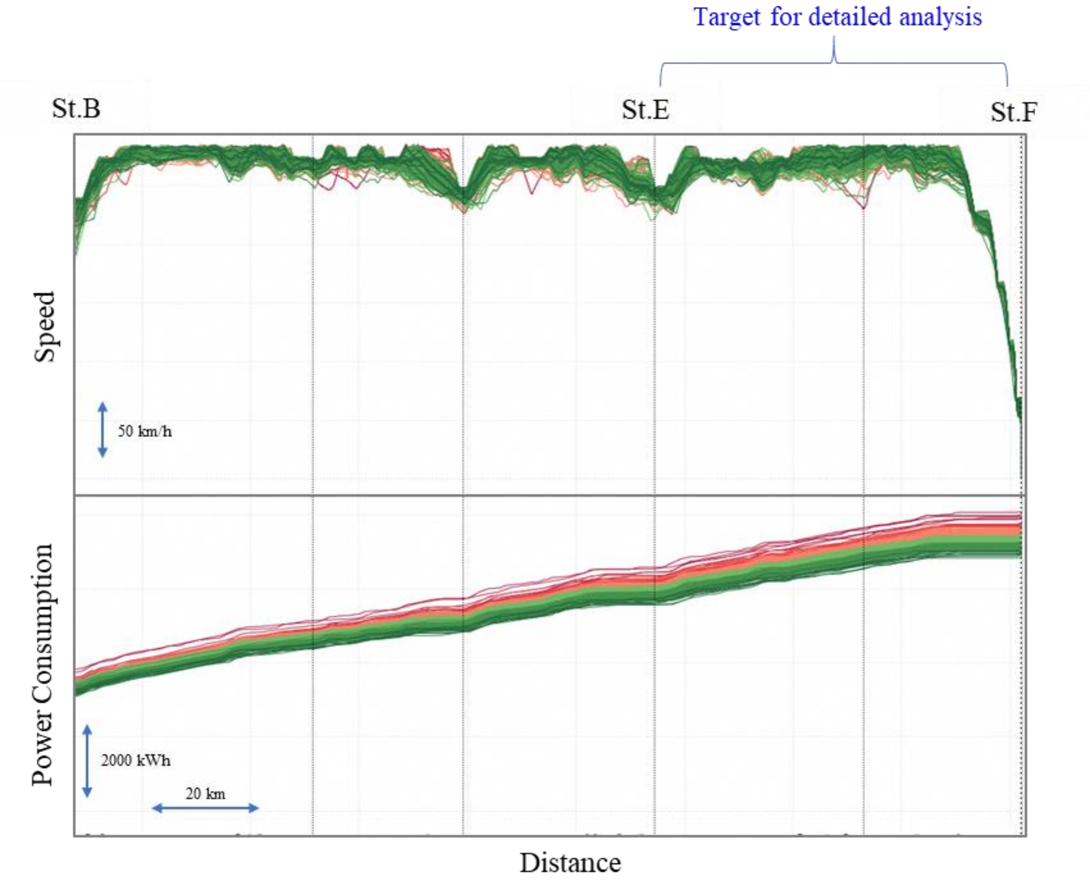


Figure 3: Example of analysis: trains running at the same travel time.

2.4 Analysis of Differences in Power Consumption

Figure 4. illustrates visualization of trains with largest and smallest power consumption between Stations E and F (specified as target for detailed analysis in Figure 3.). The horizontal axis represents the distance in kilometers, while the vertical axis represents the speed, relative elevation, electric power, and input notch. The train with smallest power consumption is shown in green, while one with largest power consumption is shown in red.

The area enclosed by the horizontal axis and the change in electric power represents the power consumption. The correlation between power consumption and the selection of notches is evident, with higher notches resulting in greater power consumption. When comparing the selection of notches, the red line apparently uses notches for longer durations and tends to use higher notch levels more frequently. Analyzing the selection of notches along with the relative elevation reveals that in the case of the red line, after passing station E and running for a while, there is a tendency to use higher notches in the uphill gradient section where the relative elevation increases.

On the other hand, in the case of the green line, a relatively constant notch level is observed even in the uphill gradient section. In the downhill gradient section where the elevation decreases towards station F, the green line performs coasting (notch off), while the red line continues to use notches and then decelerates using brakes after increasing the speed. This serves as a comparative analysis of extreme examples, however, these differences in notch operations are believed to directly affect the power consumption, resulting in significant variations in total power consumption. Therefore, reducing power consumption and minimizing the deviation among trains can be achieved by appropriately selecting notches and effectively utilizing coasting in each section.

Hence, it is important to conduct a more detailed analysis of the relationship between power consumption and notch operations, as well as quantitatively understand the impact of external factors other than notch operations (such as ambient temperature, occupancy rate, etc.) on power consumption. In the next section 2.5, we will discuss the analysis of notch operations, and in Chapter 3, we will examine the analysis of external factors such as ambient temperature and occupancy rate.

Figure 4: Visualization of trains with largest and smallest power consumption.

2.5 Impact of Notch Duration on Power Consumption

Based on the comparison in Figure 4., effective use of coasting can reduce power consumption. Therefore, a quantitative analysis was conducted on the impact of notch durations, particularly the length of coasting time, on power consumption. For 43 trains, which runs the section between stations A and F at approximately the same travel time, the durations of acceleration, coasting, and braking were calculated, and their relationship with power consumption was shown in Figure 5.

Trains with longer coasting times exhibit both higher and lower power consumption. In other words, the length of coasting time is not necessarily a significant indicator for reducing power consumption. This is a natural result, as the overall power consumption in the section is influenced by the amount of reduction achieved in cumulative power consumption during notch input time. This means that evaluating based solely on durations of notches is difficult, and it is necessary to analyze the notch operations according to the characteristic of sections.

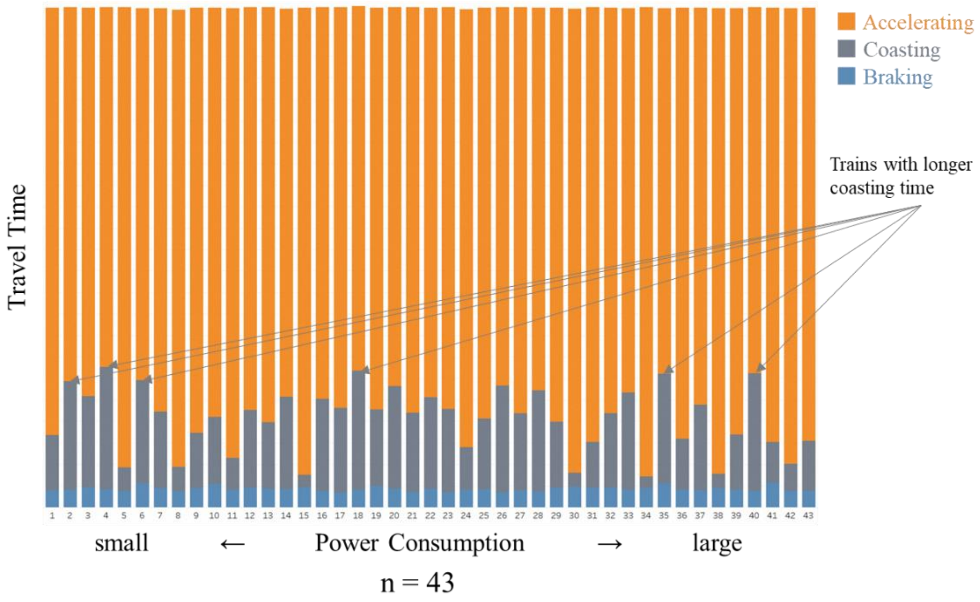


Figure 5: Relationship between durations of acceleration, coasting, braking, and power consumption.

3 Correction of power consumption based on ambient temperature and occupancy rate

Factors that contribute to variations in power consumption, apart from driving operations, include external factors like changes in ambient temperature that impact air density, differences in occupancy rate that alter train weight, and the presence of wind. These external factors can cause differences in power consumption. Additionally, variations in efficiency of equipment for each train composition may also contribute to disparities in power consumption. Given the significant impact of ambient temperature and occupancy rate on train resistance, a method to correct power consumption based on these two factors is being considered.

3.1 Relationship between ambient temperature, occupancy rate, and power consumption based on actual data

Figures 6. and 7. show the relationship between power consumption and ambient temperature, as well as power consumption and occupancy rate, for a total of 940 trains running on the same schedule without delays between stations B and C. From these figures, it can be observed that power consumption declines with increasing ambient temperature and decreasing occupancy rate. This is consistent with the theoretical trend that as the ambient temperature increases, air density decreases, leading to a reduction in train resistance. Similarly, a drop in occupancy rate corresponds to a decrease in train weight including passengers, which also lowers train resistance.

However, accurately evaluating the influence of ambient temperature and occupancy rate on power consumption using the figures is challenging. This difficulty arises from the variations in driving operations and the inclusion of diverse ambient temperatures (Figure 6.) and occupancy rates (Figure 7.). Therefore, in section 3.2, a method to correct power consumption based on ambient temperature and occupancy rate will be discussed.

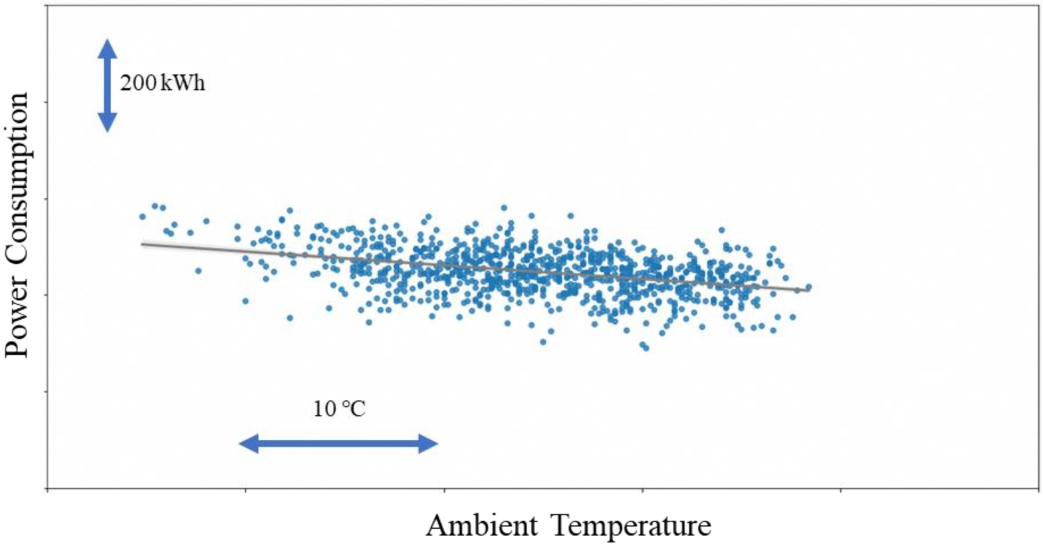


Figure 6: Relationship between ambient temperature and power consumption.

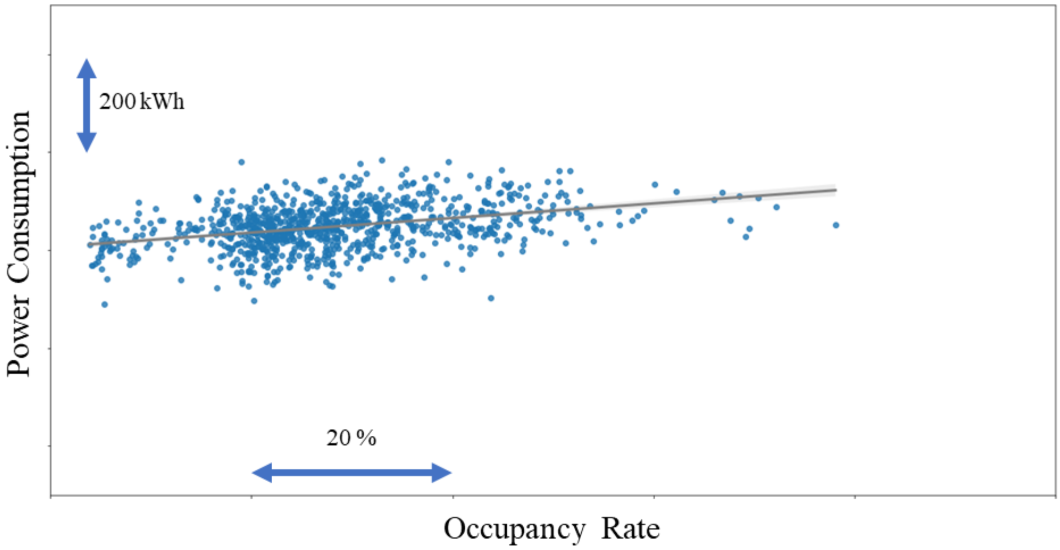


Figure 7: Relationship between occupancy rate and power consumption.

3.2 Development of Method to Correct Power Consumption based on Ambient Temperature and Occupancy Rate

The power consumed by a train is the energy required to move it from a location to another. When a train is in motion, there is a resistance (force) acting in the opposite direction of travel. The tractive force generated by motors is used to overcome this train resistance, resulting in the acceleration of the train. Therefore, if the train resistance is smaller, the tractive force required to accelerate the train will be smaller, resulting in lower power consumption by the motors. Train resistance includes running resistance, gradient resistance, curve resistance, starting resistance, and acceleration resistance. However, in the Tokaido Shinkansen, the duration of high-speed running is long and the radius of curves in the high-speed section is more than 2500 m, the focus was on running resistance and gradient resistance when considering a method to correct the power consumption. It was assumed that power consumption varies based on the difference between running resistance and gradient resistance. The difference in running resistance and gradient resistance was calculated every second for each train, assuming an ambient temperature of 25°C and an occupancy rate of 50 %. Taking into account efficiency of equipment, the amount of correction as power consumption was calculated. Running resistance R_r [kgf] and gradient resistance R_g [kgf] are expressed by Equations (1) and (2) respectively, and power consumption P [kW] is calculated by Equation (3).

$$R_r = (a + b \times V) \times W + c \times \rho \times V^2 \tag{1}$$

$$R_g = h \times W \tag{2}$$

$$P = \frac{1}{367.098} \times \frac{1}{\eta} \times (R_r + R_g) \times V \tag{3}$$

Here, a, *b*, and c are coefficients determined experimentally based on the characteristics of the train composition. *V* represents the speed in kilometers per hour [km/h], *W* is the total weight of the train including passengers [ton], ρ is the air density in kilograms per cubic meter [kg/m^3], h is the gradient [‰], and η is the efficiency of equipment.

When the ambient temperature and occupancy rate are x [°C] and y [%], respectively, the difference between the running resistance and gradient resistance compared to when they were 25 °C and 50 % can be calculated as ΔR_r [kgf] and ΔR_g [kgf] using Equations (4) and (5).

$$\Delta R_r = (a + b \times V) \times \left(W_{50} - W_y\right) + c \times (\rho_{25} - \rho_x) \times V^2 \tag{4}$$

$$\Delta R_g = h \times (W_{50} - W_y) \tag{5}$$

Here, W_{50} [ton] represents the total weight of the train including passengers at an occupancy rate of 50 %, W_y [ton] represents the total weight of the train including passengers at an occupancy rate of y [%], ρ_{25} [kg/m^3] represents the air density at an ambient temperature of 25 °C, and ρ_x [kg/m^3] represents the air density at an ambient temperature of x [°C] . From Equations (4) and (5), the amount of correction as power consumption ΔP [kW] can be calculated using Equation (6).

$$\Delta \mathrm{P} = \frac{1}{367.098} \times \frac{1}{\eta} \times (\Delta R_r + \Delta R_g) \times V \tag{6}$$

To validate the aforementioned method, a sample of 10 trains was selected, considering varying ambient temperatures and occupancy rates, all traveling at the same speed within a specific section. Table 1 shows correction of power consumption for ambient temperature and occupancy rate. From these results, the following observations can be made:

As to Train No. 1, the power consumption remains unchanged following the correction. This is because the ambient temperature and occupancy rate are close to the reference values of 25 °C and 50 % respectively, resulting in minimal amount of correction. Additionally, the effects of ambient temperature and occupancy rate corrections cancel each other out.

Looking at the overall data from the 10 trains, various combinations of ambient temperature and occupancy rate are observed, and there is a tendency for a large variation (standard deviation) in power consumption before correction.

The validity of the corrections is now being discussed. The power consumption of Train No.1 (220.5 kWh) is set as the reference value for an ambient temperature of 25 °C and an occupancy rate of 50 %. We examine the power consumption before correction for Trains No. 3, 6, and 10. The power consumption before correction for

Trains No. 3, 6, and 10 had differences of more than 5 kWh (2.27%) compared to the reference value of 220.5 kWh. When these differences are corrected based on ambient temperature and occupancy rate, they become 1.6 kWh, 0.4 kWh, and 0.9 kWh, respectively. These differences correspond to absolute values of 0.72% to 0.18% compared to the reference value, indicating that the corrections are valid.

No.	Ambient Temperature [°C]	Occupancy Rate [%]	Power Consumption (before correction) [kWh]	Power Consumption (after correction) [kWh]	after correction /before correction
1	24.3	46.4	220.5	220.5	1.00
2	32.5	45.3	216.4	220.7	1.02
3	31.8	43.2	214.8	218.9	1.02
4	17.1	34.8	224.1	221.9	0.99
5	23.2	55.1	222.0	220.6	0.99
6	16.8	54.3	225.5	220.9	0.98
7	32.3	59.0	217.3	219.9	1.01
8	24.5	60.8	226.1	224.7	0.99
9	25.0	66.5	225.9	224.1	0.99
10	19.6	67.9	226.1	221.4	0.98
Average			221.9	221.4	
Standard Deviation			4.38	1.80	

Table 1: Correction of power consumption for ambient temperature and occupancy rate

Figure 8. depicts the frequency distribution of the power consumption before and after correction between stations A and F. After the correction, the standard deviation of the power consumption decreased by 25.8%, showing that the variance of power consumption per train decreased due to the correction. This method is effective in reducing the impact of external factors such as ambient temperature and occupancy rate. As a result, a method was established to understand the difference in power consumption by significantly minimizing the impact of external factors, such as ambient temperature and occupancy rate. The analyses will proceed using the power consumption after correction.

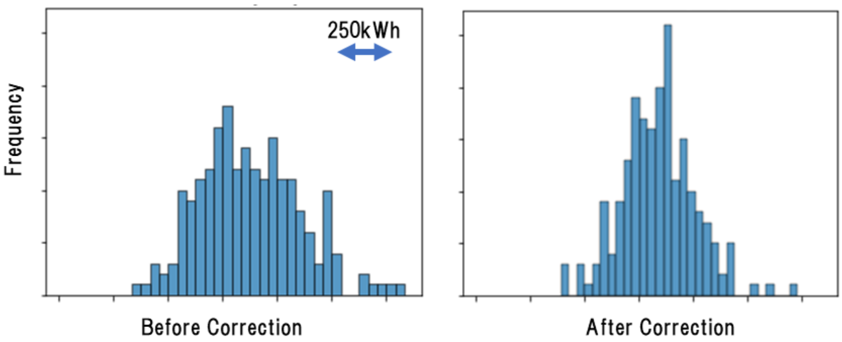


Figure 8: Frequency distribution of the power consumption.

4 Understanding the Characteristics that Contribute to Reducing Power Consumption

In actual operation, there are always variations in driving conditions, and no two driving data are exactly the same. It is difficult to grasp the factors that cause changes in power consumption by comparing multiple driving data because of the large amount of information involved. Therefore, we attempted to efficiently compare the data by clustering each driving data based on the differences in notch operations per kilometer and dividing them into groups with smaller and larger power consumption. The k-means clustering method was adopted for this purpose.

We created a dataset consisting of notch values at each location for 344 trains between stations B and D and performed clustering on this dataset. Number of clusters was set as 100 using the elbow method, which identifies the point where the decrease in residual sum of squares (RSS) becomes gradual.

As a result of clustering, each cluster was assigned 1 to 21 trains. The average power consumption for trains belonging to each cluster was calculated, and they were arranged from left to right in ascending order of average power consumption. The vertical axis represents the number of trains belonging to each cluster (Figure 9.). This graph illustrates the relationship between the characteristics of notch operations and power consumption.

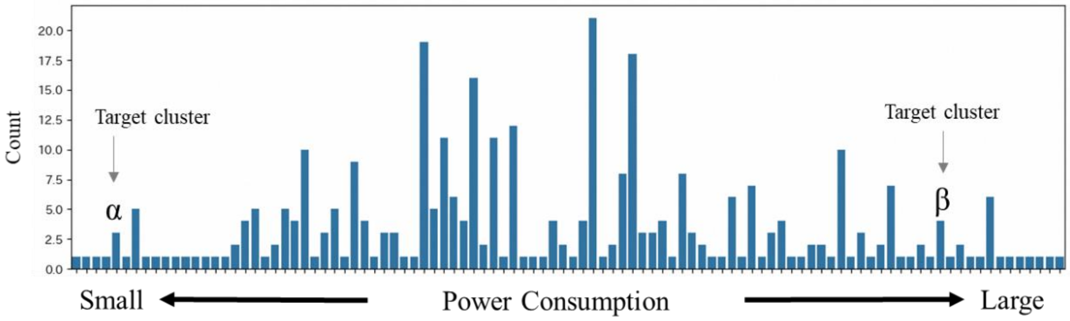
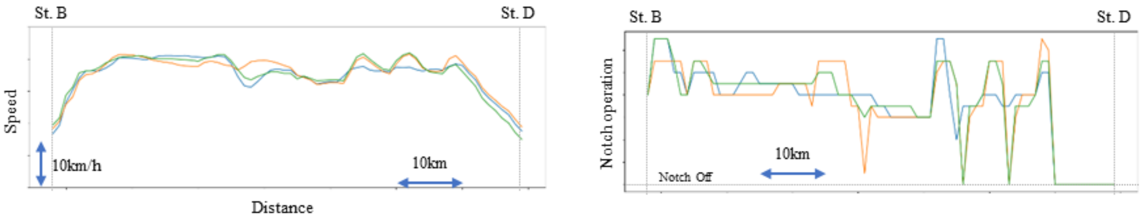
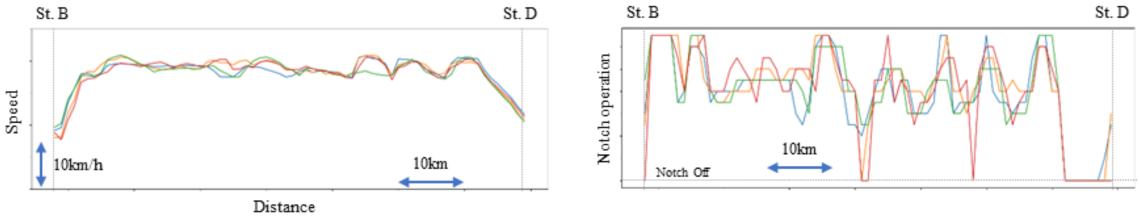


Figure 9: Frequency distribution of clusters.

To understand the trends in notch operations concerning cases with varying power consumption, this can be accomplished by appropriately selecting relevant clusters from the figure above for comparing the driving curve and notch operations. As an example, Figure 10. shows the results of plotting the speed and notch operation on the vertical axis and distance in kilometers on the horizontal axis for two clusters (α and β) shown in Figure 9. In the cluster with smaller power consumption, α (n=3), there is less variation in notch and a greater frequency of using low notch levels. On the other hand, in the cluster with larger power consumption, β (n=4), there is more variation in notch and a greater frequency of using high notch levels. This aligns with the correlation between power consumption and selection of notches discussed in section 2.4. By conducting clustering on multiple driving data, it has become possible to extract the characteristics of various running patterns.



Cluster α (n=3) with smaller power consumption.



Cluster β (n=4) with larger power consumption.

Figure 10: Example of plotting: driving curve and notch operation.

5 Conclusions and Contributions

In this study, we aimed to reduce power consumption of trains in the Tokaido Shinkansen by analyzing a large amount of field data.

We examined a method for calculating power consumption from the vehicle data. This allowed us to visualize and analyze trends in the power consumption of actual running trains. Another method was developed to address the influence of ambient temperature and occupancy rate on power consumption. By applying corrections to the data under various conditions of temperature and occupancy rate, we became able to standardize the conditions and make meaningful comparisons to identify differences in power consumption.

Furthermore, we employed clustering analysis to extract distinctive features from the extensive dataset. This allowed us to identify common characteristics among the diverse running patterns observed. We intend to derive optimal running methods that further reduce power consumption based on the insights gained from the massive field data. These methods will be validated using actual trains to assess their effectiveness.

References

[1] Central Japan Railway Company, 2023, CENTRAL JAPAN RAILWAY COMPANY FACT SHEETS 2023, page10, https://company.jr-central.co.jp/ir/factsheets/_pdf/factsheets2023.pdf

[2] Chuang, H.J., Ho, C.Y., Chen, C.S., Chen, C.S., Lin, C.H., Hsieh, C.H., 2009. Design of Optimal Coasting Speed for MRT Systems Using ANN Models. IEEE Trans. Ind. Appl. 45, 2090–2097. https://doi.org/10.1109/TIA.2009.2031898

[3] Brenna, M., Foiadelli, F., Longo, M., 2016. Application of genetic algorithms for driverless subway train energy optimization. Int. J. Veh. Technol. 2016, 1–14. https://doi.org/10.1155/2016/8073523

[4] K. Ogawa, T. Yokouchi, Y. Takeuchi, T. Saito, 2024. Verifying the Energy-saving Effect of a Driver Advisory System using Speed Estimation for Freight Railways. IEEJ Transactions on Industry Applications, 144(3), 87-93

[5] M. Miyatake, A Simple Mathematical Model for Energy-saving Train Scheduling. IEEJ Transactions on Industry Applications, Vol.131 No.6, 860-861