Print ISSN 0033-9008
Online ISSN 1880-1765

QUARTERLY REPORT of RTRI

Nov. 2026 Vol. 67 No. 4

CONTENTS

PAPERS

A Method for Visualizing Passenger Flow Data from a Passenger Distribution Estimation System in Railway Stations [I][O]

Configuration Proposal for an Onboard Insulation Diagnostic Device for Superconducting Maglev Propulsion Coils Using Partial Discharge Detection [I][R]

Proposal of New Maintenance Standard Values of Track Irregularity for High-speed Type Turnout [I][T]

Simulation and Parametric Study of Ground Behavior Induced by Chemical Grouting Considering Permeation and Fracture Grouting Mechanisms [I]

Fatigue Life Evaluation of Aged Rails Considering Actual Situation on Practical Track [I]

Formalization of Field Engineers' Tacit Knowledge Using a Statistical Occurrence Prediction Model for Lock Adjustment of Point Machines [I]

Survey of Load Conditions and Evaluation of Bending Load Capacity of Aged Shinkansen Prestressed Concrete Sleepers [I]

Estimation of Newly Fallen Snow Density Using Disdrometer Data [N]

Detectability of Residual Stress Anomalies in Various Types of Railway Wheels Subjected to Thermal Overload Based on Back Gauge Change [O][R]

Evaluation of Yield Management with a Randomized Controlled Trial (A/B Test) [O]

Characterization of Instrumented Wheelset Considering Contact Condition between Wheel and Rail [R][T]

Effects of Operating Conditions on Traction Motor Bearing Life [R]

SUMMARIES

Summaries of Papers in RTRI REPORT (in Japanese)

[H] Human factors
[I] Infrastructure
[N] Natural hazards
[O] Operations
[R] Rolling stock
[T] Technical system integration and interaction

RAILWAY TECHNICAL RESEARCH INSTITUTE

QUARTERLY REPORT of RTRI

Print ISSN 0033-9008
Online ISSN 1880-1765

November 2026 Vol. 67 No. 4

CONTENTS

PAPERS

247 A Method for Visualizing Passenger Flow Data from a Passenger Distribution Estimation System in Railway Stations [I][O]
G. TSUSHIMA, M. ISHIZUKI, M. SHIBATA

253 Configuration Proposal for an Onboard Insulation Diagnostic Device for Superconducting Maglev Propulsion Coils Using Partial Discharge Detection [I][R]
S. OTA, R. IKEDA, M. NAKASHIMA

261 Proposal of New Maintenance Standard Values of Track Irregularity for High-speed Type Turnout [I][T]
K. SHIOTA, A. MIYASHITA

270 Simulation and Parametric Study of Ground Behavior Induced by Chemical Grouting Considering Permeation and Fracture Grouting Mechanisms [I]
Y. YAMASHITA, T. NAKAYAMA, T. ISHII

276 Fatigue Life Evaluation of Aged Rails Considering Actual Situation on Practical Track [I]
M. HOSODA, N. TAKAHASHI, T. DESHIMARU

284 Formalization of Field Engineers' Tacit Knowledge Using a Statistical Occurrence Prediction Model for Lock Adjustment of Point Machines [I]
W. INABA, K. MATSUOKA, S. TAMEHIRO, M. FUKUDA, K. KAITO

291 Survey of Load Conditions and Evaluation of Bending Load Capacity of Aged Shinkansen Prestressed Concrete Sleepers [I]
T. WATANABE, S. MINOURA, K. GOTO

298 Estimation of Newly Fallen Snow Density Using Disdrometer Data [N]
K. TAKAMI

303 Detectability of Residual Stress Anomalies in Various Types of Railway Wheels Subjected to Thermal Overload Based on Back Gauge Change [O][R]
K. HANDA

309 Evaluation of Yield Management with a Randomized Controlled Trial (A/B Test) [O]
R. MATSUMOTO

317 Characterization of Instrumented Wheelset Considering Contact Condition between Wheel and Rail [R][T]
T. TANAKA, S. KUNIYUKI

324 Effects of Operating Conditions on Traction Motor Bearing Life [R]
D. SUZUKI, K. TAKAHASHI

SUMMARIES

330 Summaries of Papers in RTRI REPORT (in Japanese)

[H] Human factors
[I] Infrastructure
[N] Natural hazards
[O] Operations
[R] Rolling stock
[T] Technical system integration and interaction

A Method for Visualizing Passenger Flow Data from a Passenger Distribution Estimation System in Railway Stations

Ginga TSUSHIMA **Mitsutaka ISHIZUKI** **Munenori SHIBATA**
Architectural Engineering Laboratory, Structures Technology Division

Monitoring passenger flow in real time within railway stations enables congestion points to be identified effectively, contributing to enhanced station safety management. We verified the estimation accuracy of a passenger distribution estimation system developed previously using real video footage recorded within railway stations. The results confirmed that the system can capture origin-destination (OD) traffic volume trends with a high degree of accuracy. Additionally, we visualized real-time passenger distributions to estimate the passenger flow and verified that the developed system can generally reproduce the actual dynamics of passenger flow in real time.

***Key words*:** *passenger flow, image analysis, security camera, cross-sectional traffic volume*

1. Introduction

For railway operators, monitoring real-time passenger flow conditions within railway stations is important because it serves as fundamental information for identifying congestion points, ensuring station safety, and implementing safe evacuation strategies and evacuation guidance in case of emergencies. However, current methods for real-time passenger flow monitoring are limited to locations where security cameras are installed, and no method presently exists that can determine flow conditions across an entire railway station in real time. Furthermore, real-time passenger flow information is also important for railway passengers themselves as a means of avoiding congestion and understanding emergency situations, the only available approach remains camera image distribution services [1] that anonymize security camera footage.

As a general method for understanding passenger flow that is not restricted to real-time applications, there are techniques that conduct OD traffic surveys and then analyze or predict passenger flow conditions. However, existing OD traffic surveys require considerable cost because they depend on manual counting and questionnaire surveys, and they cannot support long-term or constant measurement. This creates issues when applying them to railway station planning, which is strongly affected by surrounding development, or to construction planning in stations where movement routes frequently change. In recent years, methods [2] for estimating crowd size by using signals from communication devices carried by passengers, such as GPS and Wi-Fi, have also been developed, but these methods remain limited to understanding relative differences in crowd size according to place and time.

Accordingly, we have previously proposed [3] a method that automatically estimates OD traffic volume by utilizing image analysis and mathematical models as an alternative to conventional manpower-dependent OD traffic estimation. We also used the obtained OD traffic data as a basis to develop a passenger distribution estimation system capable of grasping passenger flow conditions in real time through passenger flow simulation [4].

This paper reports the verification results regarding the estimation accuracy of OD traffic volume produced by the developed system, while also presenting proposed visualization methods targeting both operators and passengers for the obtained OD traffic data and real-time passenger flow conditions. As for the organization of this paper, Section 2 outlines the developed passenger distribution estimation system. Section 3 presents verification results concerning the OD traffic estimation accuracy achieved by the system. Section 4 explains proposed methods for visualizing the obtained OD traffic volume as well as proposed visualization methods for real-time passenger flow conditions.

2. Passenger distribution estimation system

Figure 1 illustrates the configuration of the passenger distribution estimation system. The system estimates OD traffic volume while performing simulations by applying OD traffic estimation results (usage ratios of each route) that were analyzed in advance (e.g., during the same time period on the previous day) to cross-sectional traffic volume measured in real time from security camera footage. Through this process, passenger flow conditions within the railway station are monitored in real time. Figure 1 shows steps performed by the system; "Step 1: cross-sectional traffic measurement" and "Step 2: OD traffic volume estimation" in the preliminary analysis, and "Step 3: real-time cross-sectional traffic measurement" and "Step 4: passenger flow simulation" in the real-time analysis. Detailed explanations of each step are given in previous studies [3][4].

Step 1: "Cross-sectional traffic measurement"

Cross-sectional traffic volume is measured using AI from the acquired video footage. The open-source YOLOX [5] is used to conduct people detection (head detection), and ByteTrack [6], which is an open-source multiple object tracking (MOT) method, is used to track the same individuals across consecutive frames.

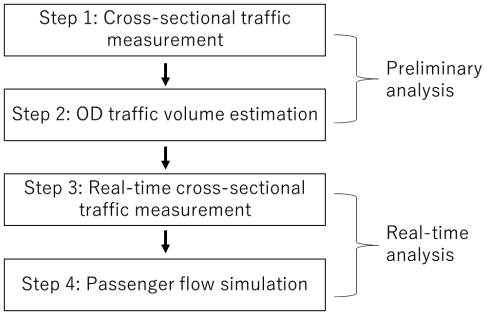


Fig. 1 Configuration of passenger distribution estimation system

Through this process, the number of people passing each section is automatically measured (Figure 2).

Step 2: “OD traffic volume estimation”

The connected movement routes between security camera locations within railway stations are constructed as a network in order to estimate OD traffic volume, namely the number of people moving between cameras, using a mathematical optimization model.

Step 3: “Real-time cross-sectional traffic measurement”

The same processing used in cross-sectional traffic volume measurement (Step 1) is carried out in real time, and the number of people passing each second is measured.

Step 4: “Passenger flow simulation”

Agents are generated and moved in the simulation based on measured cross-sectional traffic volume and previously estimated OD traffic volume, with pedestrian speeds adjusted according to congestion conditions, thereby visualizing passenger flow.

3. OD traffic estimation in railway stations

3.1 Data acquisition method

System verification was conducted at the two railway stations A and B shown in Figures 3 and 4 by using footage from video cameras installed so that their viewing angles closely matched those of the existing security cameras. In general, security camera footage recorded at railway stations often has a resolution of around 640×480 px at approximately 10 fps. However, this system assumes direct analysis of footage before it passes through the recorder, where image quality of at least HD resolution 1280×720 px can generally be expected. Therefore, the system was verified using high-resolution video footage (1440×1080 px at 30 fps).

Railway station A is a station where transfers between platforms rarely occur, and the video cameras targeted for measurement were located at 7 points labeled a-g. Railway station B is a station where transfers occur between all platforms, and 9 locations labeled a-i were selected as targets. In Railway station A, a video camera was also installed at point h, which is a merging location, in order to confirm flow conditions. Data for both railway stations were acquired during weekday morning rush hours from 7:00 to 9:00.

3.2 Measurement and validation of cross-sectional traffic volume

For each video obtained at the respective railway stations, AI-based automatic cross-sectional traffic volume measurement was performed for one peak-hour period, and measurement accuracy was verified. Figures 5 and 6 plot the automatically measured counts against counts manually obtained from the recorded footage, both aggregated in five-minute intervals. These figures show that the error remained within approximately 10%. It should be noted that when cross-sectional traffic volume is large, manual counting itself may also involve errors, but in this paper, the error is defined as the difference between automatically measured counts and manually counted results treated as ground truth. In part of Railway station B, there were sections where the AI-based counts became smaller and fell outside the 10% error range, indicated by the red- framed area in Figure 6. This is thought to be because the video camera recorded an escalator located beneath an artificial ground structure, resulting in dark footage where the black escalator steps blended with human silhouettes, thereby making AI-based measurement difficult. Thus, analysis of video footage from a total of 16 cameras confirmed that



Fig. 2 Cross-sectional traffic measurement by AI

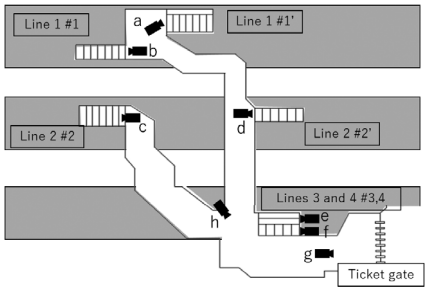


Fig. 3 Floor plan and camera locations (Railway station A)

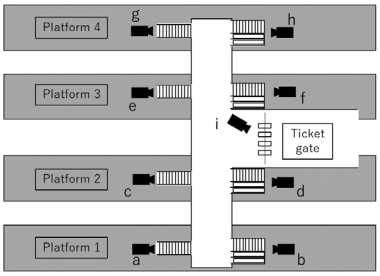


Fig. 4 Floor plan and camera locations (Railway station B)

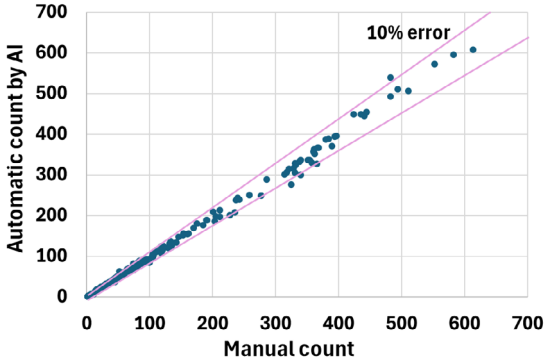


Fig. 5 Manual vs. automatic count (Railway station A, 8:00-9:00, 5-min)

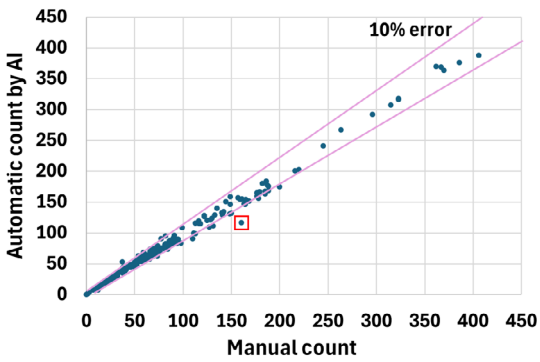


Fig. 6 Manual vs. automatic count (Railway station B, 7:30-8:30, 5-min)

unless conditions are extremely poor, AI-based cross-sectional traffic volume measurement can be achieved with an error of approximately 10% when the camera angle relative to the ceiling is generally between 30° and 50°.

3.3 Estimation and validation of OD traffic volume

In accordance with the method presented in previous research [3], OD traffic volume was estimated through mathematical optimization using the measured cross-sectional traffic volume at each location, and its accuracy was verified. Figures 7 and 8 show the network diagrams of the video cameras used in the mathematical optimization process. The network structure places generation and attraction nodes at locations where video cameras are installed, connects cameras linked by movement routes as links, and sets branch nodes at locations where movement routes merge or diverge.

In cross-sectional traffic volume aggregation, a time lag caused by travel time occurs between the moment a passenger is observed at the origin and the moment the same passenger is observed at the destination. Therefore, depending on the aggregation timing, a large discrepancy may arise between the total trip generation volume and the total trip attraction volume. This discrepancy needs to be minimized when estimating OD traffic volume, and this was achieved by shifting the aggregation timing of the trip attraction volume relative to the aggregation timing of the trip generation volume. Figure 9 shows the correlation coefficient between trip generation volume and trip attraction volume, with the aggregation timing of the trip attraction volume shifted in 5-second increments. The correlation coefficient reached its maximum when a time delay of 35 seconds was applied. Under this condition, the temporal variations in the total trip generation and attraction volumes were generally consistent with each other (Figure 10). Accordingly, the aggregation timing for trip attraction volume was delayed by 35 seconds in Railway Station A and by 40 seconds in Railway Station B for OD traffic volume estimation.

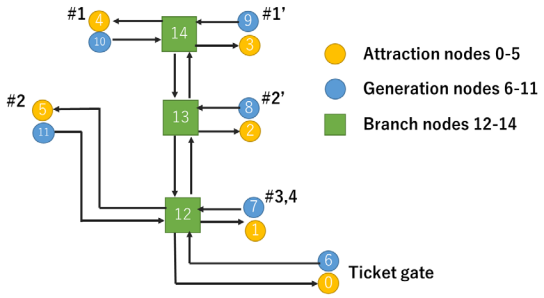


Fig. 7 Camera network diagram (Railway station A)

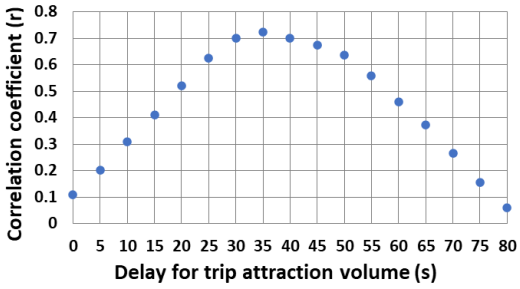


Fig. 9 Correlation between trip generation and attraction volumes (Railway station A, 5-s)

In order to create ground truth OD traffic volume data for verifying estimation accuracy, for this study, we conducted a sampled survey extracting passenger OD data in Railway station A which has few transfer passengers. We used a sampling rate of 10% which is comparable to ordinary OD traffic surveys. The results were then expanded to one hour of observed traffic volume and used as the ground truth values. For Railway station B, due to the large number of transfer passengers, the survey period was restricted to the peak five-minute interval from 7:51 to 7:56 to achieve a higher sampling rate. Passenger OD extraction surveys were conducted with a sampling rate of 72.5%, and the results expanded to the observed traffic volume for the same five-minute period were used as ground truth values.

Tables 1 and 2 present the estimated OD traffic volume results for each railway station, including the estimated number of people and the destination ratios for each origin location. In Railway station A, there were very few transfer passengers moving from one platform to another, whereas in Railway station B, approximately half of the passengers generated on the platforms were transfer passengers. This result is also consistent with observations made on site.

Next, Figures 11 and 12 compare the ground truth OD traffic volume values produced through the sampling survey with the estimation results generated by the estimation model. For both railway stations, the correlation coefficient with the correct values exceeded 0.97, confirming that the proposed estimation model can capture OD traffic volume trends with a high level of accuracy.

4. Visualization of passenger flow

4.1 Visualization of OD traffic volume

Aggregated OD traffic volumes are commonly presented in tabular form, as shown in Tables 1 and 2. They can also be visual-

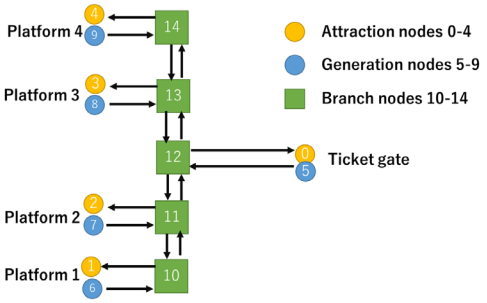


Fig. 8 Camera network diagram (Railway station B)

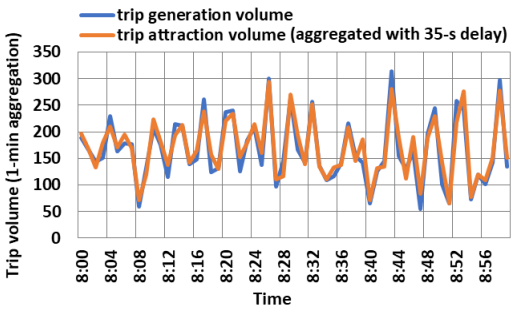


Fig.10 Temporal change in trip generation and attraction volumes (Railway station A)

Table 1 OD traffic volume estimation results (Railway station A, 8:00-9:00)

Origin \ Destination	#1	#1'	#2	#2'	#3,4	Ticket gate
#1			6 (1%)	1 (0%)	6 (1%)	623 (98%)
#1'			6 (1%)	1 (0%)	6 (1%)	623 (97%)
#2	2 (0%)	7 (0%)			5 (0%)	1939 (99%)
#2'	2 (0%)	7 (1%)			5 (1%)	802 (98%)
#3,4	2 (0%)	6 (0%)	5 (0%)	1 (0%)		1990 (99%)
Ticket gate	470 (11%)	1345 (33%)	1057 (26%)	213 (5%)	1033 (25%)	

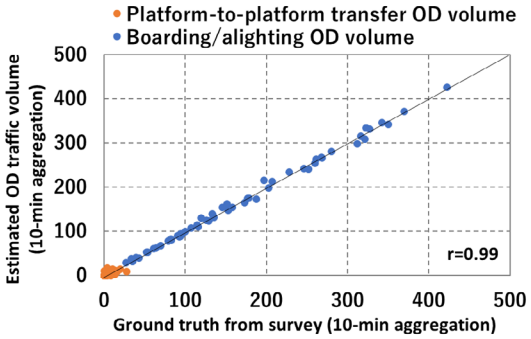


Fig. 11 Ground truth vs. estimated values (Railway station A, 8:00-9:00)

Table 2 OD traffic volume estimation results (Railway station B, 7:51-7:56)

Origin \ Destination	Platform 1	Platform 2	Platform 3	Platform 4	Ticket gate
Platform 1		11 (11%)	8 (8%)	22 (22%)	61 (60%)
Platform 2	23 (28%)		5 (6%)	14 (17%)	39 (48%)
Platform 3	76 (28%)	23 (8%)		47 (17%)	130 (47%)
Platform 4	117 (31%)	36 (9%)	27 (7%)		199 (53%)
Ticket gate	72 (46%)	22 (14%)	17 (11%)	45 (29%)	

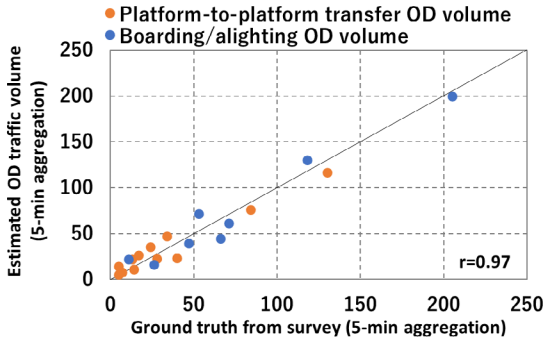


Fig. 12 Ground truth vs. estimated values (Railway station B, 7:51-7:56)

ized using a circular chord diagram (Figure 13), where OD traffic volume is expressed by the width of the bands connecting origins and destinations, namely the width occupied by each color along the circumference. However, in railway stations, large numbers of people move in conjunction with train arrivals and departures, so these passenger waves caused by train operations must also be captured appropriately together with OD traffic volume. For this reason, conventional presentation methods such as Table 1 and Figure 13, which use aggregation units of 30 minutes or one hour, have the limitation that traffic waves and localized congestion points are difficult to grasp.

Therefore, we developed a method for visualizing OD traffic volume (hereafter, "diagram") based on train operation diagrams: individual passenger movements are represented with time on the horizontal axis and distance on the vertical axis. In this method, passenger flow simulation is used to calculate travel time from the origin to the destination, after which the passenger generation time and attraction time are computed and connected with line segments. In the present study, because the movement routes are simple, generation and attraction are connected directly with line segments and represented using average walking speed. However, when analyzing longer routes or railway stations with higher congestion levels, it is possible to divide the line segments by section and express changes in walking speed by altering the slope of the line segments, namely walking speed, at crossing points and other locations. Figure 14 presents an image of visualization using the diagram method. With this method, congestion conditions in arbitrary time periods and passageways can be understood from the number of lines drawn, whereas the degree of passenger crossing can be grasped from the number of line intersections.

Figure 15 shows an example of a 30-minute diagram for Railway station A during the period from 8:00 to 8:30. The number of lines makes congestion conditions graspable. For example, from 8:10 to 8:15, there are relatively few boarding passengers, indicated by the larger white areas in the background; whereas from 8:15 to 8:20, there are many boarding passengers, represented by orange lines, and congestion is especially visible near the ticket gate. Next, Figure 16 shows a representation in which the thickness of the lines representing passengers is made inversely proportional to passageway width so that congestion level, namely density relative to each section, can be grasped rather than simply the absolute number of passengers. Since congestion level can be expressed as the proportion of passageway space occupied by passengers, when the passageway widths of Sections A and B are L and 2L, respectively, a single passenger is represented with a line thickness of 1/L in Section A and 1/2L in Section B. Compared with Figure 15, the density of lines in Figure 16 appears more uniform, which indicates that although the absolute number of passengers is greater near the ticket gate, the passenger density relative to passageway width does not differ greatly throughout the station as a whole.

4.2 Visualization of real-time passenger flow

We used video cameras a-g in Railway station A as targets to conduct real-time analysis by the passenger distribution estimation system, namely processes (Figure 1, Steps 3 and 4), in order to grasp real-time passenger flow conditions, and the visualization method is presented here. In this study, analysis was performed while replaying recorded footage, thereby pseudo-reproducing real-time analysis conditions.

In the real-time analysis, the number of passengers passing each camera in real time was allocated to each route based on previously estimated OD traffic volume results, after which passenger movement conditions were predicted through passenger flow simulation so that congestion conditions throughout the railway station can be monitored. As in the preliminary analysis, cross-sectional traffic volume was measured by applying YOLOX and ByteTrack, and passenger flow within the railway station was grasped by gener-

ating and moving agents within the simulation according to the cross-sectional traffic volume entering each second.

In the visualization method proposed in this study, the railway station floor plan was divided into grid-like meshes, particles were generated according to passenger density within each mesh, and the number of passengers together with movement direction was represented by changing the particle colors according to OD category. The visualization used previously estimated OD traffic volume results, so real-time visualization can be achieved without time lag. An example of the visualization is shown in Figure 17. Around 8:10:15 in Figure 17, boarding passengers occupied most of the railway station, whereas around 8:17:30, many alighting passengers were arriving from Lines 2 and 3/4, and it can be seen that boarding and alighting passengers intermingled, causing congestion throughout the station. These conditions also correspond to footage from video camera h shown in Figure 3 at the same time.

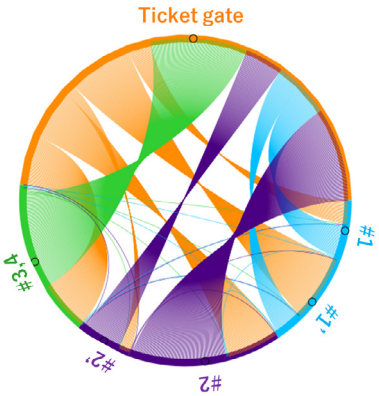


Fig. 13 Circular chord diagram (Railway station A)

Next, YOLOX analysis was applied to footage from video camera h, and the temporal change in the number of passengers detected within the camera area was compared with the number of people at the same location calculated by passenger flow simulation (Figure 18). When passenger numbers were small, the results generally matched, whereas as passenger numbers increased, the discrepancy became larger. The reason for this is thought to be that when passenger density increases, occlusion becomes significant, and AI measurement by video camera h fails to capture people located farther away, reducing the number of detected individuals.

5. Conclusion

When the developed passenger distribution estimation system was applied to footage actually recorded within railway stations and the estimation accuracy of OD traffic volume was verified, the correlation coefficient with conventional methods exceeded 0.97, confirming

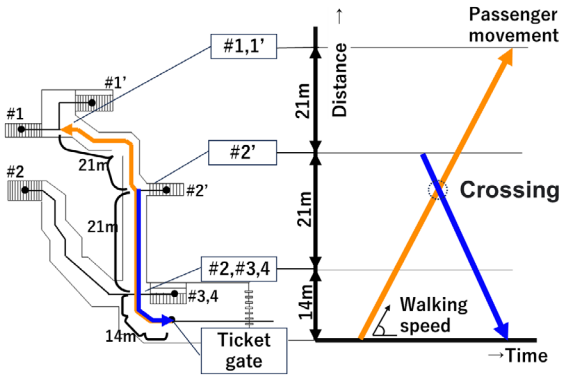


Fig. 14 Visualization method for OD traffic volume

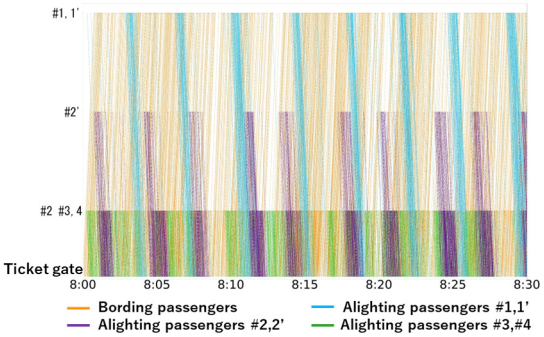


Fig. 15 Visualization of OD traffic volume (Railway station A: path width not considered)

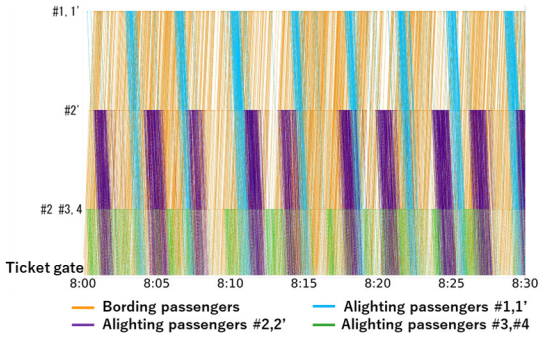


Fig. 16 Visualization of OD traffic volume (Railway station A: path width considered)

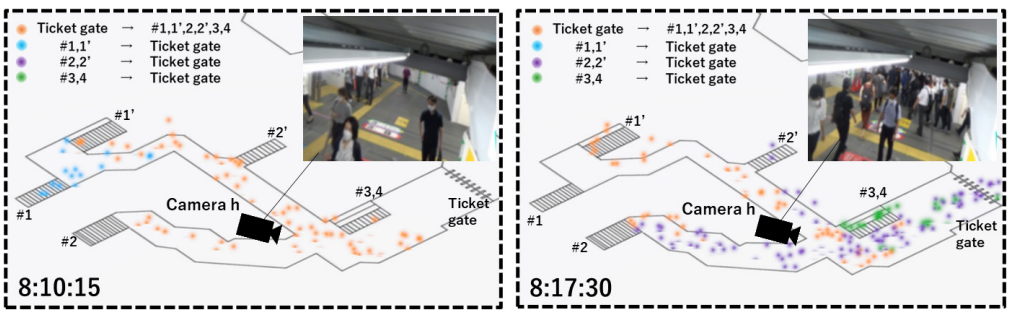


Fig. 17 Visualization of real-time passenger flow (Railway station A)

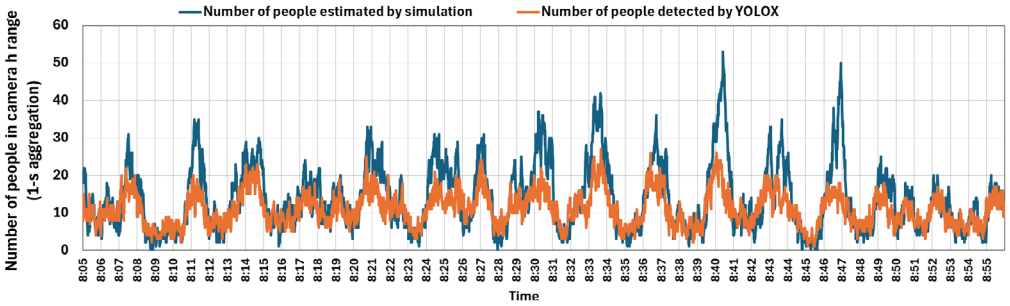


Fig. 18 Simulated vs. YOLOX-detected number of people (Railway station A, camera h)

that OD traffic volume trends can be captured with a high level of accuracy. In addition, a new diagram-based visualization method was proposed for the obtained OD traffic volume, demonstrating that temporal changes in OD traffic volume and cross-sectional traffic volume at each location can be represented. Furthermore, real-time passenger distribution was estimated and visualized, confirming that actual real-time flow conditions could generally be reproduced.

Tasks for future work include examining methods for grasping and predicting passenger flow in areas where security cameras are not installed and for linking with other systems that require dynamic changes according to passenger flow conditions in order to build railway station environments that are safer and more passenger-friendly.

Acknowledgment

We would like to express our thanks to staff at the East Japan Railway Company and West Japan Railway Company who assisted this study in carrying out passenger flow surveys, including the installation of video cameras within railway stations.

References

[1] TOKYU CORPORATION, News Release, 2016, https://www.tokyu.co.jp/company/news/list/Pid%3D2470.html (accessed: 2024-05-09) (in Japanese).
[2] Sachiko O., Toshiyuki N., Tomotaka U., and Shigeharu T., "A study on survey method by wi-fi packet sensor for flow monitoring system," Journal of Civil Engineering Planning and Management, Vol. 36 (No. 5), pp. I_799–I_807, 2019 (in Japanese).
[3] Munenori S., Mitsutaka I., Ginga T., "Estimation method of pedestrian od flow in railway station with applying movie processing and mathematical optimization modeling," Journal of Architecture and Planning (Transactions of AIJ), 88(803), pp. 56-67, 2023 (in Japanese).
[4] Hiromichi O., Toshiyuki A., "The station planning with the simulation system of passenger flow," Proceedings of the J-RAIL Symposium JRail'95, pp. 469-472, 1995 (in Japanese).
[5] Zheng G., Songtao L., Feng W., Zeming L., Jian S., "YOLOX: Exceeding YOLO Series in 2021," arXiv: 2107.08430, 2021.
[6] Yifu Z., Peize S., Yi J., Dongdong Y., Fucheng W., Zehuan Y., Ping L., Wenyu L., Xinggang W., "ByteTrack: Multi-Object Tracking by Associating Every Detection Box", arXiv 2110.06864, 2021.

Authors



Ginga TSUSHIMA
Senior Researcher, Architectural Engineering Laboratory, Structures Technology Division
Research Areas: Passenger Flow, Architectural Planning



Mitsutaka ISHIZUKI
Senior Researcher, Architectural Engineering Laboratory, Structures Technology Division
Research Areas: Passenger Flow, Architectural Planning, Evacuation Planning



Munenori SHIBATA, Ph.D.
Senior Researcher, Architectural Engineering Laboratory, Structures Technology Division
Research Areas: Passenger Flow, Architectural Planning, Transportation Planning

Configuration Proposal for an Onboard Insulation Diagnostic Device for Superconducting Maglev Propulsion Coils Using Partial Discharge Detection

Satoru OTA **Ryohei IKEDA**
Magnetic Levitation Laboratory, Advanced Superconductivity Technology Division

Minoru NAKASHIMA
Magnetic Levitation Laboratory, Advanced Superconductivity Technology Division (Former)

Ground coils are essential components of superconducting maglev systems. Propulsion coils, a category of ground coils, require rigorous insulation diagnostics to withstand mechanical, electrical, and environmental stresses, strong dynamic electromagnetic interactions, and high voltages. This paper reviews the requirements for insulation diagnostics of propulsion coils, emphasizing previously developed efficient methodologies. One such methodology is an advanced diagnostic system that analyzes electromagnetic waves from partial discharges to evaluate coil deterioration and identify discharge locations. Following the discussion, we propose the configuration of an onboard insulation diagnostic device and verify the required specifications and functionality of its detection, measurement, recording, and analysis functions.

***Key words*:** *superconducting maglev, ground coil, partial discharge, electromagnetic waves, insulation degradation diagnosis*

1. Introduction

Superconducting maglev vehicles generate their propulsion and levitation forces through electromagnetic interaction between the superconducting magnets mounted on the bogies and the ground coils installed along the guideway. Therefore, a large number of ground coils must be installed along the entire line. Since ground coils are subjected to various loads depending on their function and location, various factors must be accounted for, including mechanical vibration, electrical stress, and environmental degradation. Propulsion coils, in particular, provide longitudinal force to the vehicles and are subjected to the resulting reaction force. Additionally, they experience high voltages induced by the fluctuating magnetic field of the superconducting magnets, as well as the large current flowing through the coils themselves. Since many propulsion coils are installed, maintaining sufficient insulation performance against high voltage is critical; therefore, efficient inspection methods are required.

Previously, we investigated the effectiveness of onboard insulation diagnosis for evaluating propulsion coil insulation performance and demonstrated partial discharge detection using a single antenna [1]. Additionally, we investigated position resolution during operation at 500 km/h. However, methods for measurement and recording after detection, and methods for diagnosing ground coil degradation and identifying discharge locations through data analysis, remain insufficiently explored. This paper outlines the required functions of an insulation diagnostic system using onboard partial discharge detection (hereinafter referred to as the "onboard system") and reports the results of performance verification.

It should be noted that this study verified each function of the onboard system individually. The results do not constitute an integrated operational verification of the entire system, nor do they include insulation diagnostic tests conducted during operation using an actual superconducting magnetic levitation vehicle.

2. Requirements for insulation diagnosis of propulsion coils

The following sections describe an onboard system designed to efficiently perform insulation diagnosis of propulsion coils. First, we explain insulation defects and partial discharge, which are the primary factors necessitating insulation diagnosis. Next, the configuration of the proposed onboard system for efficient partial discharge detection is described.

2.1 Insulation diagnosis of propulsion coils using partial discharge

Propulsion coils typically consist of metallic conductors and coil resin, which is used for electrical insulation and provides insulation from ground potential. However, insulation defects, such as voids, can form within the resin. As shown in Fig. 1, the defective region has a lower dielectric constant and dielectric breakdown strength than the surrounding resin, resulting in electric field concentration. Consequently, partial discharge, a localized dielectric breakdown phenomenon, occurs.

As shown in Fig. 2, the repeated occurrence of partial discharge causes progressive deterioration of the insulation, which eventually leads to complete dielectric breakdown once electrical insulation from ground potential can no longer be maintained. If partial discharge is detected early using the onboard system, proactive actions, such as coil replacement, can be implemented before dielectric breakdown occurs.

Partial discharge is a microscopic dielectric breakdown occurring inside the insulation of electrical equipment. Partial discharge can be detected through various physical and chemical phenomena, as summarized in Table 1 [2]. One or more of these phenomena can be selected for use in the onboard system.

When applying the phenomena and detection methods in Table 1 to ground coil insulation diagnostics based on insulation material degradation, this method is considered unsuitable because propulsion coils use resin insulation. Additionally, light emission and heat generation detection methods are unsuitable because propulsion coils are enclosed by levitation and guidance coils. Regard-

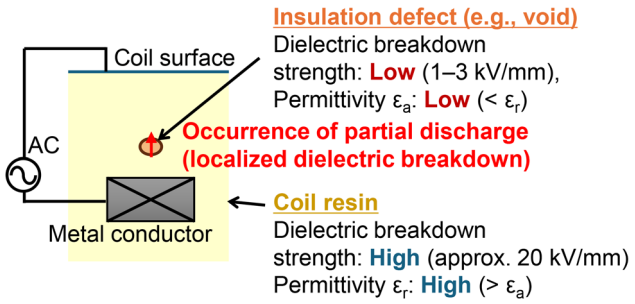


Fig. 1 Occurrence of partial discharge within coil resin

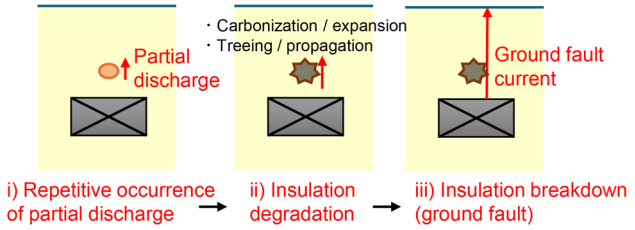


Fig. 2 Repetitive occurrence of partial discharge, progression of damage, and insulation breakdown

ing ultrasonic detection, the material composition of propulsion coils—including polymer materials and fillers, such as silica—is expected to cause significant scattering, making detection difficult.

Conversely, pulse current or voltage detection methods are effective, as are the conventional partial discharge measurement techniques used in manufacturing plants [3]. However, when applied at the installation site of propulsion coils, additional ground-side equipment is required, such as high-voltage test power supplies and multiple current transformers (CTs). This leads to a large-scale system configuration with significant operational constraints.

On the other hand, appropriately implementing methods for detecting electromagnetic waves or magnetic fields makes onboard insulation diagnosis feasible, enabling inspection from vehicles traveling at 500 km/h. Furthermore, electromagnetic wave-based measurements have been verified to be feasible even when the propulsion coils are enclosed by the levitation and guidance coils [1]. Therefore, this method is advantageous for achieving compact system configuration and inspection efficiency.

Based on these advantages, this study employs an inspection method capable of detecting electromagnetic waves and magnetic fields.

2.2 Positioning of electromagnetic-wave-based insulation diagnosis for propulsion coils

Various test methods are used to diagnose the insulation of propulsion coils. Conventional methods include measurement of circuit parameters and insulation resistance; other methods include withstand voltage, tan δ , and partial discharge testing. Each method has its own characteristics, applicable conditions, and limitations.

Circuit parameter and insulation resistance measurements are used to evaluate the overall condition of equipment, but they offer limited capacity for detailed diagnosis. In contrast, while withstand voltage testing (such as commercial-frequency and lightning impulse tests), and tan δ testing, are effective for evaluating insulation performance, it is difficult to use these methods to diagnose individual coils during testing on a main line. Additionally, conducting these tests in the field with circuits composed of many inter-connected coils requires the supply of high voltage, which in turn calls for large-scale test transformers, resulting in significant equipment-related constraints.

Conventional partial discharge testing, typically performed during manufacturing plant inspections, also faces difficulties when applied to individual coil diagnosis on a main line. In contrast, the inspection method based on electromagnetic wave and magnetic field detection, as described in Section 2.1, enables efficient diagnosis. However, since this method uses the operating voltage of the propulsion coils as the test voltage, values exceeding this level cannot be applied, unlike in conventional partial discharge tests. Consequently, this limits the types and magnitudes of partial discharges that can be evaluated.

The following sections discuss the insulation diagnosis of propulsion coils using this electromagnetic wave-based measurement method, focusing on the system configuration and the requirements for each function.

2.3 Requirements for the onboard system

This paper investigated the conceptual configuration of an onboard system for diagnosing the insulation of propulsion coils on a high-speed moving vehicle. Figure 3 shows the system configuration and a conceptual diagram of the inspection process. When a

Table 1 Physical and chemical phenomena of partial discharges, their detection methods, and applicability to propulsion coil insulation diagnostics

Phenomenon	Detection Method	Overview of Detection Method	Applicability
Insulation degradation	Gas sensors, chemical analysis	Detection of chemical changes in insulating materials (gas or insulating oil) by component analysis and related methods	Not applicable
Light emission / heat generation	Optical sensors, ultraviolet monitors	Detection of light emission and heat generated by partial discharges using sensors or cameras	Not applicable
Ultrasonic waves	Acoustic sensors, AE sensors	Detection of sound or ultrasonic waves generated by partial discharges using sensors	Not applicable
Current / voltage	Coupling capacitors, high-frequency CTs	Detection of pulse currents caused by partial discharges and voltage variations along the propagation path using measuring instruments	Applicable (with constraints)
Electromagnetic waves / magnetic fields	Magnetic sensors, antennas	Detection of electromagnetic waves or magnetic fields generated by partial discharges using sensors or antennas	Applicable (without constraints)

vehicle passes through an inspection section, the propulsion coil circuits are energized, producing the operating voltage of the propulsion coils. In addition, fluctuating magnetic fields generated by the movement of the vehicle equipped with superconducting magnets induce additional voltages in the propulsion coils. Together, these two voltages result in the high voltage applied to the propulsion coils. If any propulsion coil within the inspection section has degraded insulation, the above-described partial discharge is expected to occur. The proposed system is designed to continuously measure the electromagnetic waves generated by these partial discharges.

The onboard system must fulfill the following four functions:

1. A detection function to capture electromagnetic waves as physical phenomena;
2. A measurement function to quantitatively evaluate the detected electromagnetic waves;
3. A recording function to store electromagnetic wave data;
4. An analysis function to determine the insulation condition of propulsion coils from the recorded data.

Table 2 summarizes the assumptions and conditions established for each function.

The objectives for each function were defined and used as a basis to conduct functional verification as follows:

1. Detection function: detects UHF electromagnetic waves radiated from propulsion coils due to partial discharge.
2. Measurement function: identifies system configurations, and enables selection of methods for quantitative evaluation of electromagnetic waves.
3. Recording function: allows continuous operation for more than 70 minutes (to exceed the assumed 67-minute one-way travel time between Shinagawa and Osaka).
4. Analysis function: identifies signs of insulation degradation in the propulsion coils from the electromagnetic wave measurement results. Additionally, the locations of the electromagnetic wave sources must be estimated with a spatial resolution equal to or higher than that of individual propulsion coils.

Based on these objectives, the configuration of the onboard system was determined and then verified.

3. Proposed configuration and functional verification of the onboard insulation diagnostic system

3.1 Detection function

Regarding the detection function, this section first describes the resulting characteristics of partial discharge and the electromagnetic waves. Assuming the microcurrent generated by partial discharge inside a propulsion coil can be represented as a pulse waveform with a duration on the order of nanoseconds, its frequency spectrum is represented by a sinc function, $\mathrm{sinc}(x) = \sin(x)/x$. Consequently, the electromagnetic waves generated by partial discharge exhibit a broadband frequency spectrum. Based on this broadband characteristic, we investigated the specifications of the detection function. In addition to the magnitude of the frequency components, factors such as the size of the detection antenna must be considered, depending on the frequency range.

Figure 4 illustrates the relationship between the magnitude of the frequency components, obtained by assuming various partial discharge current pulse durations, and the dimensions of a half-wavelength dipole antenna used for electromagnetic wave detection.

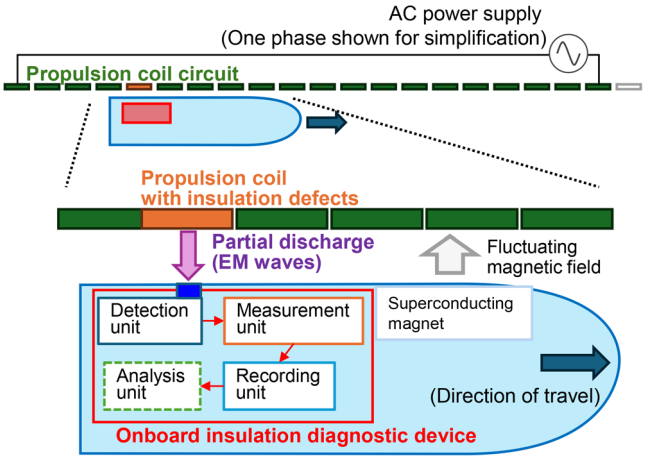


Fig. 3 System configuration of onboard diagnostic device and conceptual diagram of inspection

Table 2 Prerequisites for onboard diagnostic device

Item	Consideration (Function)	Prerequisite
Detection target	Frequency characteristics of the detector (Detection)	Electromagnetic waves (UHF band*) *300 MHz-3 GHz
Coil applied voltage	No special consideration required (Non-contact measurement)	Up to approximately 20 kV
Coil spacing and passing speed	Available measurement time (Measurement), Position estimation resolution (Analysis)	0.9 m, maximum 500 km/h (6.5 ms passage per coil)
Measurement duration	Recording media capacity (Recording)	Up to 70 min

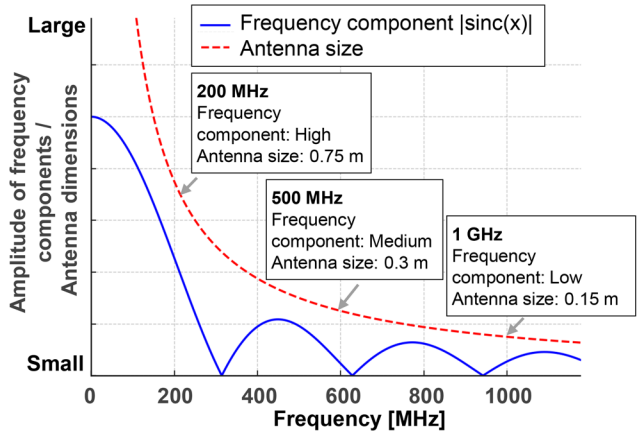


Fig. 4 Frequency component distribution of partial discharges and antenna dimensions

To improve the signal-to-noise (S/N) ratio in insulation diagnosis, larger frequency components of the partial-discharge electro-

magnetic waves are preferable. However, when considering installation as an onboard system, the antenna size is subject to physical constraints. Therefore, a frequency range that enables electromagnetic-wave detection while allowing a relatively compact antenna is desirable. As shown in the example measurement results of partial-discharge electromagnetic waves in Fig. 5, detection was insufficient when using an antenna frequency of 1 GHz, whereas detection was feasible at 500 MHz. Based on these results, an antenna frequency of 500 MHz was selected.

3.2 Measurement function

Several candidate configurations for the measurement function were selected, focusing on commercially available instruments. The three types of instruments listed in Table 3 were evaluated through performance verification tests as candidate devices to acquire and record the electromagnetic waves detected by the antenna described in the previous section.

Table 3 Types and overview of measurement functions

Measuring Instrument	Overview
Digital oscilloscope	Suitable for acquiring signal waveforms in the time domain and directly observing waveform shape, amplitude, and frequency by capturing voltage fluctuations in real time
Spectrum analyzer	Useful for signal analysis in the frequency domain, including frequency spectrum analysis of electromagnetic waves and analysis of noise and interference
Vector signal analyzer	In addition to the above, capable of handling phase information and performing more advanced and highly customizable signal analysis

This study focuses on the detection of partial-discharge electromagnetic waves in the UHF band (300 MHz-3 GHz). However, when configuring the measurement function, the required performance specifications must be carefully defined to ensure accurate identification of partial discharge in the subsequent analysis function. In particular, the time and frequency resolutions are important.

(1) Requirements in the time domain

The temporal resolution of the measurement function must be fine enough to capture the nanosecond duration of the current pulses generated by partial discharge. Therefore, instruments with high sampling rates, such as 2 to 5 gigasamples per second (GS/s), were selected for the measurement function to ensure accurate temporal resolution. This capability enables highly accurate capture of the timing of partial discharge events (Fig. 6).

(2) Requirements in the frequency domain

During the implementation phase, settings of the measurement parameters were confirmed to significantly affect detection performance for partial discharge electromagnetic waves. In particular, the settings for the reference level, center frequency, and measurement bandwidth were critical.

Reference Level

The reference level affects the dynamic range and, therefore, requires appropriate adjustment. Setting the reference level too high may prevent the detection of weak electromagnetic waves generated by partial discharge. Conversely, setting the reference level too low may cause over-range conditions, interfering with stable measurement. For example, as shown in Fig. 7(a), an excessively high reference level results in insufficient dynamic range, burying the partial discharge signal in background noise.

Center Frequency

The frequency components of electromagnetic waves generated by partial discharge exhibit a broadband distribution that follows a sinc function. Considering the environmental noise and characteristics of the antenna reception in the test environment, the 500 MHz band was initially selected as suitable for measurement. However, the test results revealed the presence of other radio-frequency signals in the 500 MHz band, as shown in Fig. 7(b). Subsequently, the center frequency was adjusted to 465 MHz, and measurements were conducted under this condition, as shown in Fig. 8. Additionally, certain frequency bands contained interfering wireless signals that degraded the post-processing and analysis of the acquired data. Furthermore, as shown in Fig. 7(c), desired signals could not be obtained in frequency bands where partial discharge was weak. These results indicate that frequency selection must consider both external noise conditions and partial discharge characteristics.

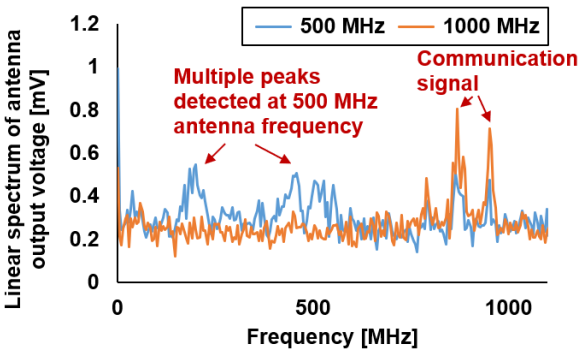


Fig. 5 Electromagnetic wave measurement results for selecting antenna frequency

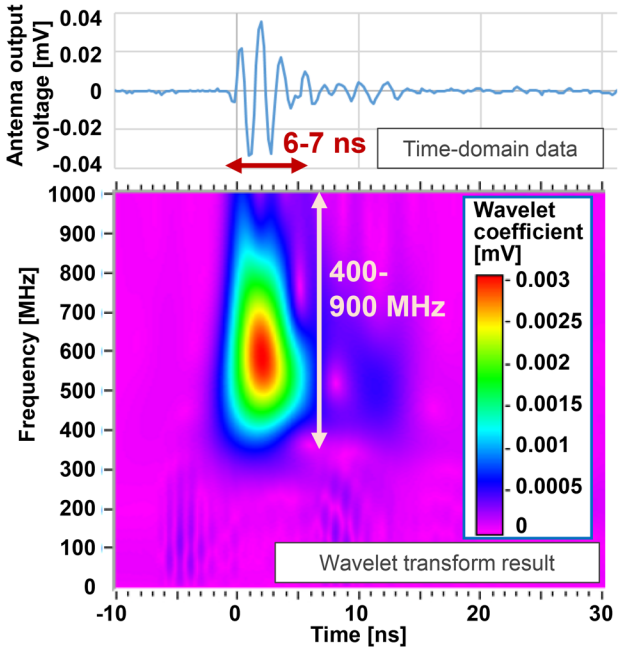
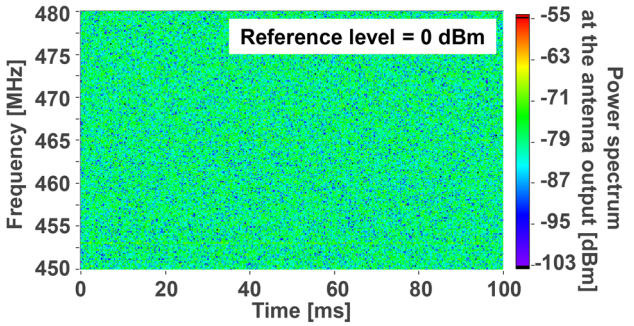
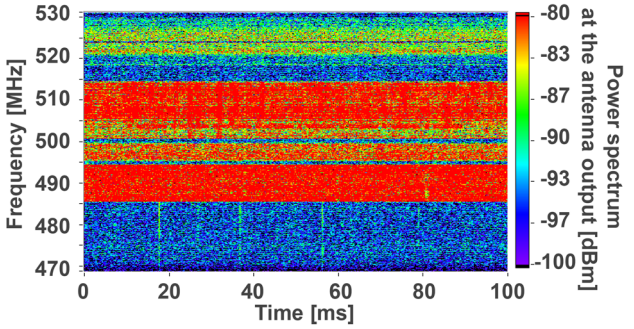


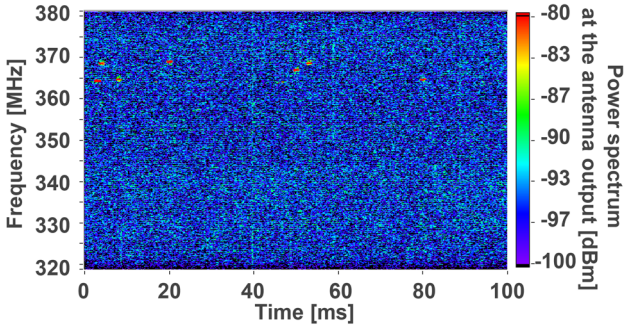
Fig. 6 Example of captured electromagnetic waveforms



(a) Inappropriate reference level setting (0 dBm)



(b) Other radio signals near the center frequency



(c) Weak PD waves near the center frequency

Fig. 7 Example of acquired electromagnetic waveforms

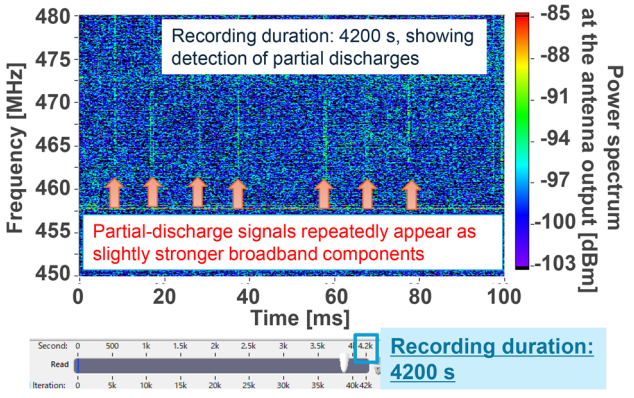


Fig. 8 Electromagnetic wave detection results of partial discharges under recording function settings

Measurement Bandwidth

Improper configuration of the measurement bandwidth may adversely affect the recording function. An excessively wide bandwidth setting was found to cause memory overflow, thereby reducing the probability of successful measurement. To balance the trade-off between measurement performance and memory usage during recording, the measurement bandwidth was set to 30 MHz for the recording-function tests described in Section 3.3 (Fig. 8). The next section demonstrates that this setting enables stable measurement operation.

Appropriate configuration of these parameters ensured the stable detection of electromagnetic waves generated by partial discharge.

However, in addition to stable detection, the appropriate design of the recording function is essential for acquiring long-duration data under actual operating conditions. The following section describes the configuration and performance verification of the recording function.

3.3 Recording function

The recording function must support long-duration operations, with the capacity to store data for a period corresponding to a single run between Shinagawa and Osaka. This requirement was established because implementing system redundancy is difficult due to physical constraints and cost, requiring measurements to be performed using a single onboard system.

As listed in Table 4, an investigation of configurations satisfying the recording requirements revealed that digital oscilloscopes are unsuitable for long-duration recordings due to limitations inherent in their measurement principles. Although spectrum analyzers are capable of long-duration recording, they measure only a single frequency component and are therefore unsuitable for capturing the broadband frequency components associated with partial discharge. In contrast, vector signal analyzers were found to be capable of recording broadband signals over extended durations. Furthermore, we confirmed that the recording function requirements could be met by appropriately customizing the hardware and software configurations.

Accordingly, to verify the recording function, the configuration shown in Table 5 was set up, and a long-duration recording test was conducted using a simulated partial discharge source [4]. Figure 8 shows the test results.

As shown in Fig. 8, the target recording duration of 70 minutes and stable detection of partial discharge electromagnetic waves were successfully achieved by appropriately configuring the hardware and recording software.

Table 4 Measurement duration of instruments for selecting recording function

Type	Example Model	Measurement Duration
Digital oscilloscope	Tektronix DPO70000X Series	0.5 s
Spectrum analyzer	Tektronix RSA5000B Series	5 s-6 h[*1]
Vector signal analyzer	NI PXIe-5663	More than 70 min[*2]

[*1] Extremely narrow frequency span
[*2] Customization required

Table 5 Configuration for verification of recording function

Category	Component / Parameter	Configuration / Setting
Hardware	Measuring instrument / detector	NI PXIe-5663 / Anritsu MP534B
	Recording media	Samsung Portable SSD T7 (2 TB)
Software settings	Measurement frequency span	30 MHz
	Center frequency	465 MHz
	Reference level	−70 dBm
	Recording duration	4200 s (70 min)

3.4 Analysis function

Using the measurement data obtained from the onboard system configuration described above, we performed an insulation degradation diagnosis of the propulsion coils. This section describes methods for detecting signs of insulation degradation from electromagnetic wave data and estimating the locations of the sources.

Although the analysis function does not need to be integrated into the onboard system, performing real-time analysis within the system enables a rapid response to abnormalities. Therefore, this study incorporated the analysis function as part of the onboard system to improve overall operational performance.

3.4.1 Insulation degradation diagnosis

To investigate the behavior of molded specimens that are structurally similar to propulsion coils and emit partial discharges prior to dielectric breakdown, diagnostic tests were conducted using the configuration shown in Fig. 9 [5]. Electromagnetic waves were collected three times per minute. The frequency components of the waveform $f(t)$ at time t were obtained via a discrete Fourier transform and expressed as $F(t)$. Then, changes in the index G_i(t), defined as the sum of the frequency components within each frequency band, were evaluated.

The three measurements obtained each minute were analyzed independently in this study, without averaging or maximum-value processing. The frequency components of each individual measurement were evaluated separately.

$$G_i(t)=\sum_{j=1}^{n_i} Fij \quad (1)$$

F_{0j}: Frequency components from 0 MHz to less than 50 MHz

F_{ij}: Frequency components from $100i$ − 50 MHz to less than $100i$ + 50 MHz ($1 \le i \le 10$)

n_i: Number of frequency components within the same frequency band

In Equation (1), the evaluation is performed using the sum of frequency components within each frequency band rather than tracking a specific frequency component individually, because the electromagnetic waves generated by partial discharges exhibit broadband frequency characteristics.

Figure 10 shows the temporal variation of G_i until the simulated discharge source experiences dielectric breakdown. After approximately 1.7 hours, significant increases were observed in the G_4 indicator (centered around 400 MHz) and the G_5 indicator (centered around 500 MHz). Dielectric breakdown occurred at 2.7 hours.

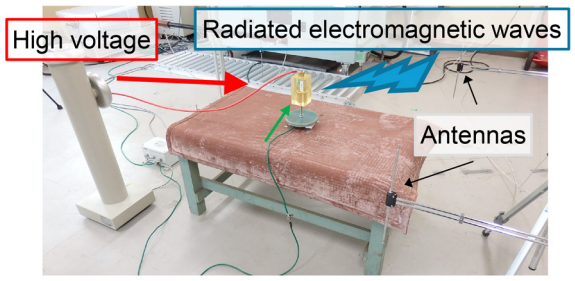


Fig. 9 Experimental setup for insulation degradation diagnosis using a simulated discharge source

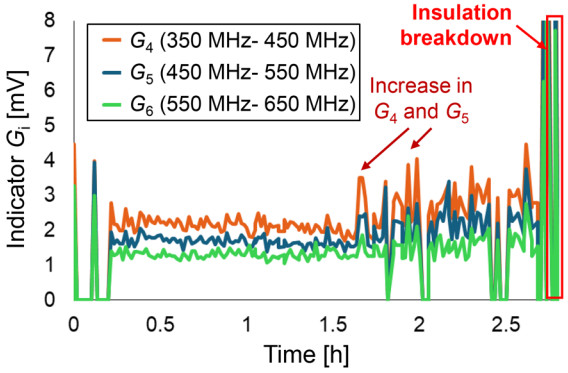


Fig. 10 Insulation degradation diagnosis results

These results suggest that monitoring G_4 and G_5 indicators may enable the detection of insulation degradation signs leading to dielectric breakdown.

It should be noted that this test was conducted under accelerated degradation conditions. Unlike actual operating environments, voltage application was continuous. For example, if the voltage application duration during operation is approximately 20 seconds per run, one hour of continuous testing corresponds to approximately 180 runs. However, since actual operating conditions differ from continuous voltage energization, it is feasible to identify long-term degradation trends by accumulating measurement results over multiple runs and monitoring variations in indicators such as G_4 and G_5.

Operational approaches are considered effective, such as recommending early replacement when indicator variations exceed predefined thresholds or determining appropriate maintenance timing through long-term monitoring of indicator trends.

3.4.2 Electromagnetic-wave source localization

In actual inspection environments, propulsion coils are arranged as shown in Fig. 3. Therefore, when electromagnetic waves attributed to partial discharge are detected, it is necessary to identify the propulsion coil acting as the radiation source. To localize the source, the time difference t_{ij} that maximizes the cross-correlation function was obtained from the received signals f_i and f_j of two antennas using Equation (2). These time differences were then converted into distance differences. Furthermore, multiple antenna configurations were used to improve the accuracy of source localization [6].

$$t_{ij}=\arg\max_t\left|\int_{t_s}^{t_e} f_i(T)\,f_j(T+t)\,dT\right| \quad (2)$$

The source localization curve **S**, derived from the antenna positions $\mathbf{P}_i$ and $\mathbf{P}_j$ and the time difference t_{ij} obtained from Equation (2),

is expressed by Equation (3). The curve **S** forms a two-dimensional hyperbola or a three-dimensional hyperboloid.

$$|\mathbf{P}_i-\mathbf{S}|-|\mathbf{P}_j-\mathbf{S}|=ct_{ij} \quad (3)$$

c: the speed of light

Using Equation (4), the likelihood indicator E_k for the k-th antenna pair was calculated at a spatial point $\mathbf{S}_e$, based on multiple time differences obtained from various antenna pairs. These values were then integrated using Equation (5) to calculate the overall likelihood indicator, E, which represents the combined evaluation result from multiple antenna combinations.

$$E_k(\mathbf{S}_e)=1-w_d\left|\|\mathbf{P}_i-\mathbf{S}_e|-|\mathbf{P}_j-\mathbf{S}_e|-ct_{ij}\right| \quad (4)$$

k: Index of antenna combinations

w_d: Penalty factor

$$E(\mathbf{S}_e)=\prod_{k=1}^{n} E_k(\mathbf{S}_e) \quad (5)$$

n: Total number of antenna combinations used for evaluation

A discharge source localization test was conducted using one propulsion coil and four antennas to determine the arrival time differences of electromagnetic waves. The configuration of the test is shown in Fig. 11. Ideally, in an onboard installation, the four antennas would be arranged along the X-axis, as viewed relative to Antenna 1. However, due to constraints on equipment placement, the antennas could not always be arranged in this manner. The localization results are shown in Fig. 12. The white region corresponding to high indicator values extends along the right side of the propulsion coil. This demonstrates that discharge source localization with a spatial resolution of approximately half the length of a propulsion coil is achievable.

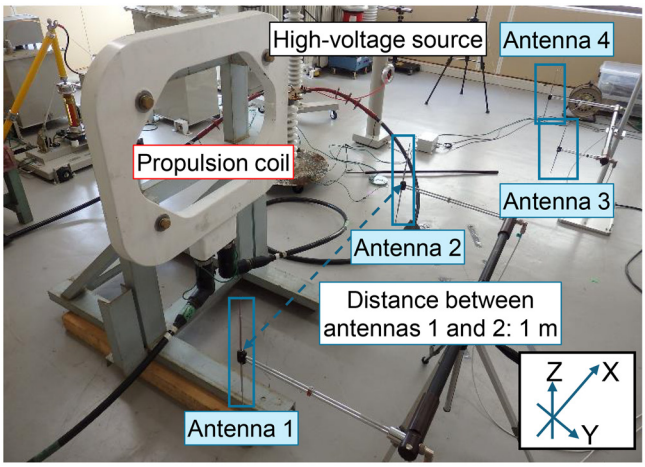


Fig. 11 Experimental setup for discharge source localization using a propulsion coil

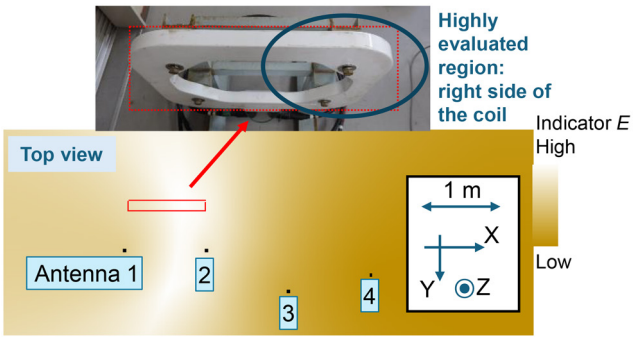


Fig. 12 Results of discharge source localization test (top view near propulsion coil)

4. Functional verification results and remaining issues of the onboard system

Each component of the developed onboard system has been verified at the component level, and configurations that satisfy the required specifications have been proposed. The following sections outline three issues requiring further investigation.

4.1 Issue 1: Phase and amplitude of applied voltage

As shown in Fig. 8, the electromagnetic waves generated by partial discharge in the propulsion coils occur periodically. Therefore, the potential distribution of the propulsion coils must be considered, since they function as components of a synchronous motor. Consequently, appropriate antenna placement must also be determined according to the phase of the periodically occurring partial discharge.

Additionally, the voltage applied to each propulsion coil varies depending on its position within a series connected propulsion coil section. Therefore, as described in Section 2.2, the applied voltage can be lower than the levels used in conventional withstand voltage tests. Consequently, an appropriate testing framework that considers both voltage phase and voltage stress conditions must be established.

4.2 Issue 2: Detailed evaluation of influence of superconducting magnets and other equipment

Although the onboard system was discussed in the context of vehicle installation, many devices required for vehicle operation are also installed onboard. Among these devices, superconducting magnets are unique to superconducting maglev systems, making it essential to evaluate the magnetic-field effects of superconducting magnets on other devices.

Figure 13 shows an example in which radiated electromagnetic waves were measured in the vicinity of an actual superconducting magnet to evaluate magnetic field effects [7]. In this example, the influence of the strong magnetic field from the magnet and the surrounding operational equipment on measurement performance was found to be negligible. Nevertheless, further detailed investigation is still required.

4.3 Issue 3: Integrated system verification and operational issues

The requirements and system configurations presented in this paper are primarily based on the verification of individual functions. Comprehensive operational verification of the entire onboard system remains to be fully executed.

Future work includes determining appropriate data formats and conducting integrated operational verification to confirm that the

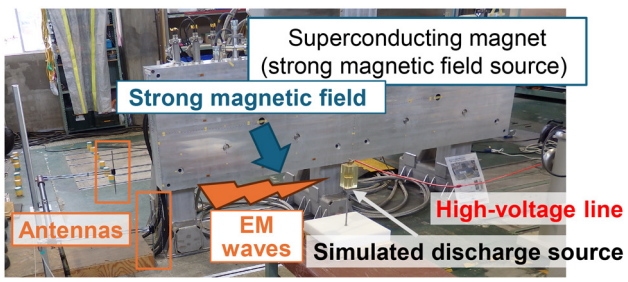


Fig. 13 Example of an evaluation test on influence of other onboard equipment

detection, measurement, recording, and analysis functions operate seamlessly while satisfying the proposed system requirements.

Additionally, further investigation is required to establish testing procedures for practical system implementation, operational methods for onboard use, and concrete approaches for discharge source localization. Other important issues include maintaining durability and measurement accuracy under actual operating conditions, such as vehicle vibration and environmental noise.

5. Conclusion

Propulsion coils used in superconducting maglev systems are installed in large numbers and require high insulation reliability. This study identified the functional requirements of an onboard insulation diagnostic system and conducted various tests to verify its performance.

The required specifications and functionalities for the detection, measurement, recording, and analysis functions were successfully validated. However, several challenges remain to be addressed, including the evaluation of unverified items and the integrated operational verification of the overall system.

Future work will involve refining the system configuration and analysis methods further, as well as continuing to explore solutions to address the challenges identified in this study.

References

[1] Ota, S., Yoda, H., and Ikeda, R., "Development of Efficient Non-contact Insulation Diagnosis for the Propulsion Coil of the Superconducting Maglev," *RTRI Report*, Vol. 29, No. 11, pp. 23-28, 2015 (in Japanese).

[2] Noda, K. and Nagai, Y., "Partial Discharge Measurement Equipment," *Meiden Review*, Vol. 337, No. 4, p. 13, 2012 (in Japanese).

[3] *International Electrotechnical Commission*, *IEC 60270: High-voltage Test Techniques – Partial Discharge Measurements*, 2000.

[4] Ota, S., "Long-Term Electromagnetic Wave Recording in the Onboard Inspection System for Ground Coil Insulation Diagnosis," presented at the *2024 IEEJ National Convention*, Paper 5-193, p. 326, 2024 (in Japanese).

[5] Ota, S. and Ikeda, R., "Insulation Deterioration Diagnosis Using Time and Frequency Domain Characteristics of Electromagnetic Wave Radiated from the Partial Discharge Generation Molded Specimen," presented at the *2022 IEEJ National Convention*, pp. 180-182, 2022 (in Japanese).

[6] Ota, S., "Estimation of Partial Discharge Occurrence Position of Ground Coil by Detection of Radiated Electromagnetic Waves," presented at the *Joint Technical Meeting on Transportation and Electric Railway / Linear Drives*, TER-23-002/LD-23-002, pp. 7-12, 2023 (in Japanese).

[7] Ota, S. and Nakajima, M., "Evaluation of the Impact of High Magnetic Fields from Superconducting Coil Magnets for Maglev Trains on Electromagnetic Wave Sensing," presented at the *2023 IEEJ Industry Applications Society Conference*, pp. III-241-242, 2023 (in Japanese).

Authors



Satoru OTA
Senior Researcher, Magnetic Levitation Laboratory, Advanced Superconductivity Technology Division
Research Areas: Railway Electrical Systems, Insulation Diagnostics



Ryohei IKEDA
Senior Researcher, Magnetic Levitation Laboratory, Advanced Superconductivity Technology Division
Research Areas: Insulation Materials, Material Evaluation



Minoru NAKASHIMA
Senior Researcher, Magnetic Levitation Laboratory, Advanced Superconductivity Technology Division (Former)
Research Areas: Maglev Vehicle Dynamics, Superconducting Maglev Systems

Proposal of New Maintenance Standard Values of Track Irregularity for High-speed Type Turnout

Katsutoshi SHIOTA **Ayano MIYASHITA**
Track Structures and Components Laboratory, Track Technology Division

Maintenance standard values for track irregularity at high-speed type turnouts, where trains pass through the main line of the turnout at a speed of 120 km/h on a conventional line, are managed with stricter values than those for general turnouts. These values were empirically determined from past running tests. In this study, turnout running simulation models were developed to quantitatively evaluate the effects of track irregularity on turnout component damage and vehicle dynamics for establishing mechanically based maintenance values. Using these models, various analyses of track irregularity were conducted, and new maintenance standard values were proposed.

***Key words*:** *high-speed type turnout, track irregularity, FEM, MBD, turnout component damage, vehicle dynamics*

1. Introduction

In the Former Japan National Railways (hereinafter referred to as "Former JNR"), the allowable passing speed on the main line side of turnouts on conventional lines (narrow gauge) was raised to a maximum of 100 km/h in 1968. Subsequently, with the increasing demand for higher speeds, running tests at 120 km/h were conducted in 1981. As a result, the occurrence of excessive lateral force (defined as 60 kN or more in this study) was confirmed in some turnouts. The factors contributing to this were identified as rail joint misalignment, a large track gauge at the crossing, and a small check gauge. Additionally, track irregularity (longitudinal level, alignment, and cross level) within the turnouts was considered one of the factors increasing lateral force at these locations. Consequently, it was suggested that various countermeasures would be required at these locations for high-speed passage through turnouts [1]. In response to this, Former JNR implemented stricter maintenance standard values for rail joint misalignment, crossing track gauge, and check gauge compared to general turnouts. Furthermore, for track irregularity (longitudinal level, alignment, and cross level), the strictest upper bound values for the highest-grade line classifications in general sections were applied. After implementing these countermeasures, running tests were conducted again. Since no excessive lateral force was detected, it was decided to manage turnouts where trains pass through the main line side at a maximum speed of 120 km/h (hereinafter referred to as "high-speed type turnouts") in accordance with Table 1.

Table 1 Control items for high-speed type turnouts and general turnouts

Control Item	High-Speed Type Turnout	General Turnout*
Longitudinal level irregularity (10m chord)	13 mm	23 mm
Alignment irregularity (10m chord)	13 mm	23 mm
Cross level irregularity (10m chord)	11 mm	Track twist 23 mm
Crossing track gauge deviation	+3, −3 mm	+5, −3 mm
Check gauge	1,023-1,030 mm	1,022-1,030 mm
Stock rail fit	3 mm	4 mm
Joint stagger	1 mm	2 mm
Guard flangeway width	56 mm	None

*Maintenance standard values for the highest-grade line

2. Turnout running simulation models

In this study, the evaluation of turnout component damage and vehicle dynamics was investigated by combining two analysis methods: dynamic FEM and MBD. Generally, while it is possible to simultaneously evaluate component damage and vehicle dynamics using dynamic FEM, the computational load is extremely heavy, requiring a large amount of time for analysis. On the other hand, MBD specializes in evaluating the mechanical behavior of vehicles; although it cannot calculate the response of components constituting the turnout, its computational load is small, enabling high-accuracy evaluation in a short time. Therefore, in this study, leveraging the characteristics of both methods, a simulation model using dynamic FEM was applied to evaluate turnout component damage, and a model using MBD was applied to evaluate vehicle dynamics, properly utilizing the evaluation models according to the objective.

Each constructed model is shown in Fig. 1, and the main analysis specifications are listed in Table 2. In both models, the vehicle is modeled with rigid bodies consisting of one carbody, two bogies, and four wheelsets, with each mass point connected by spring and damper elements.

In the dynamic FEM model, the track is modeled with beam elements and spring elements for the rails, sleepers, and substructure, while the turnout components to be evaluated are modeled using solid elements, enabling the calculation of stresses and other metrics. Furthermore, to precisely simulate the dynamic behavior of the rails and the switching/locking system during vehicle passage, loading tests and switching tests of the switch panel components were conducted, reflecting the non-linear response of each component into the model [2].In the MBD model, the track is modeled as masses for the rails and sleepers that follow directly underneath the

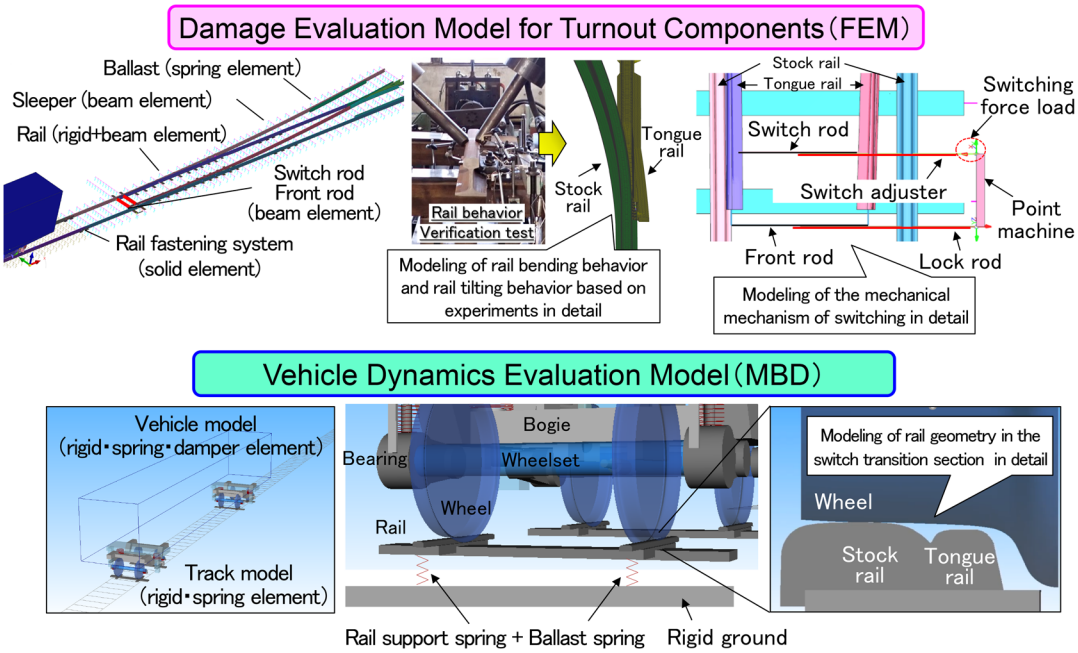


Fig. 1 Overview of turnout running simulation models

Table 2 Main analysis specifications

Item	Specification
Turnout number	No. 10 simple turnout (Elastic point)
Rail type	JIS 50N
Track gauge	1,067 mm
Slack	5 mm
Young's modulus (Rail)	2.1×10^5 N/mm^2
Poisson's ratio (Rail)	0.3
Vehicle type	Limited express vehicle (Bolsterless bogie)
Distance between bogie	14.15 m
Wheelbase	2.1 m
Friction coefficient	0.3
Bogie mass	2.83×10^4 kg/bogie
Wheelset mass	1.52×10^3 kg/wheelset

wheelsets, with these elements connected by springs and dampers. In addition, the shape of each rail at the transition section between the stock rail and the tongue rail was precisely modeled to enable high-accuracy evaluation of vehicle dynamics at the transition. The validity of each constructed model has been verified through comparison with vehicle running test results [2, 3].

3. Evaluation of effects of track irregularity on turnout component damage and vehicle dynamics

3.1 Evaluation items

To evaluate the effects of track irregularity in high-speed type turnouts on turnout component damage and vehicle dynamics, turnout running simulations were conducted using the constructed models.

The evaluation items and their limit values are listed in Table 3. These contents were established with reference to the items stipulated in the Manual of Conventional Railway Operation Speed Improvement Test [4] (hereinafter referred to as the "Speed Improvement Manual").

Among the turnout components, for the track components, the generated stress in the tongue rail elastic part, the generated stress in the rail fastening system, and the generated stress in the rail fastening bolts at the hinged heel switch were evaluated. Regarding the generated stress in the rail fastening system within the turnout, the joint of the switch panel was targeted, considering that the impact load during joint passage would be the most severe.

For the signaling components, the lock rod displacement of the switching/locking system, the generated stress in the switch adjuster, and the generated stress in the front rod were selected as control items. Note that the Speed Improvement Manual indicates that the lateral force acting on the backside of the wheel of guardrail should also be checked in addition to the above control items. However, the lateral force acting on backside of wheel of guardrail was excluded from this study because it is less susceptible to the effects of track irregularity due to the individual countermeasures taken for the crossing track gauge and back gauge (Table 1), and its structural design is already compatible with high-speed operation [5].

Next, regarding vehicle dynamics, running safety and vehicle vibration were selected as control items. For running safety, the duration of the derailment coefficient continuously exceeding 0.8 [6] (hereinafter referred to as the "derailment coefficient exceeds time") was checked, and for vehicle vibration, longitudinal level and later-

Table 3 Evaluation item and limit value

Evaluation target	Evaluation item		Limit value
Turnout components	Track components	Elastic part of switch rail stress	137 N/mm^2
		Rail fastening system stress	115 N/mm^2
		Heel part of hinged heel switch tightening bolt stress	240 N/mm^2
	Signaling components	Lock rod displacement	1.5 mm
		Switch adjuster stress	240 N/mm^2
		Front rod stress	240 N/mm^2
Vehicle dynamics	Running safety	Margin to derailment coefficient exceeds time	15 msec
	Vehicle ride motion	Longitudinal level acceleration	0.5 G
		Lateral acceleration	0.4 G

Table 4 Analysis conditions

Item	Analysis condition
Track irregularity	Longitudinal level irregularity, alignment irregularity, cross level irregularity
Track irregularity waveform	One Cosine wave (Fig. 2)
Track irregularity wavelength	10 m
Track irregularity amplitude	5 mm to 25 mm
Vehicle running speed	120 km/h
Vehicle running direction	Facing movement

al vibrations were reviewed.

Furthermore, among the switch, closure, and crossing panels, the evaluation position within the turnout was focused on the switch panel, where the impact on the components was considered to be the most severe. Here, it is known that at the crossing, which was the cause of excessive lateral force in past running tests, the track gauge and check gauge are managed more strictly compared to general turnouts; since this serves as a countermeasure against excessive lateral force, the effect of track irregularity is small [5]. Regarding closure panel, preliminary investigations have confirmed that the effects of track irregularity on component damage and vehicle dynamics are minimal [7].

3.2 Analysis conditions

The analysis conditions are listed in Table 4. The target track irregularities were longitudinal level, alignment, and cross level irregularities, which are strictly managed under the current maintenance standard values for high-speed type turnouts. To quantitatively evaluate the effect of each track irregularity, situations where they exist independently were set. The waveform of the track irregularity to be set (Fig. 2) was a single cosine wave with a wavelength of 10 m, using the amplitude as a parameter. Although the maintenance standard values for high-speed type turnouts are 13 mm for longitudinal level and alignment irregularities and 11 mm for cross level irregularity, this study conducted case studies setting values larger than these maintenance standard values to confirm the limits of the track irregularity effects. Increasing wheel load and lateral force conditions were selected for direction of track irregularity. Longitudinal level irregularity was set as a downward-facing convex; alignment irregularity was set as a convex facing the rail side where the wheels transition; and the cross level irregularity was set so that the tongue rail side, where the wheels transition, became a downward facing convex. Furthermore, the setting position for each track irregularity was determined through preliminary analysis to be the location where the wheel load and lateral force reached their maximum. The vehicle running speed was set to 120 km/h, and the running direction was set to facing travel, which is a condition that increases the fluctuation of wheel load and lateral force. It should be noted that since the stress evaluation of the tongue rail elastic part involves an extremely heavy computational load in dynamic analysis, it was evaluated through static analysis by applying the wheel load and lateral force calculated by MBD as loads.

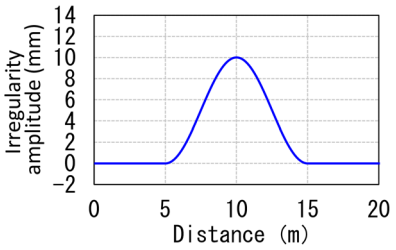
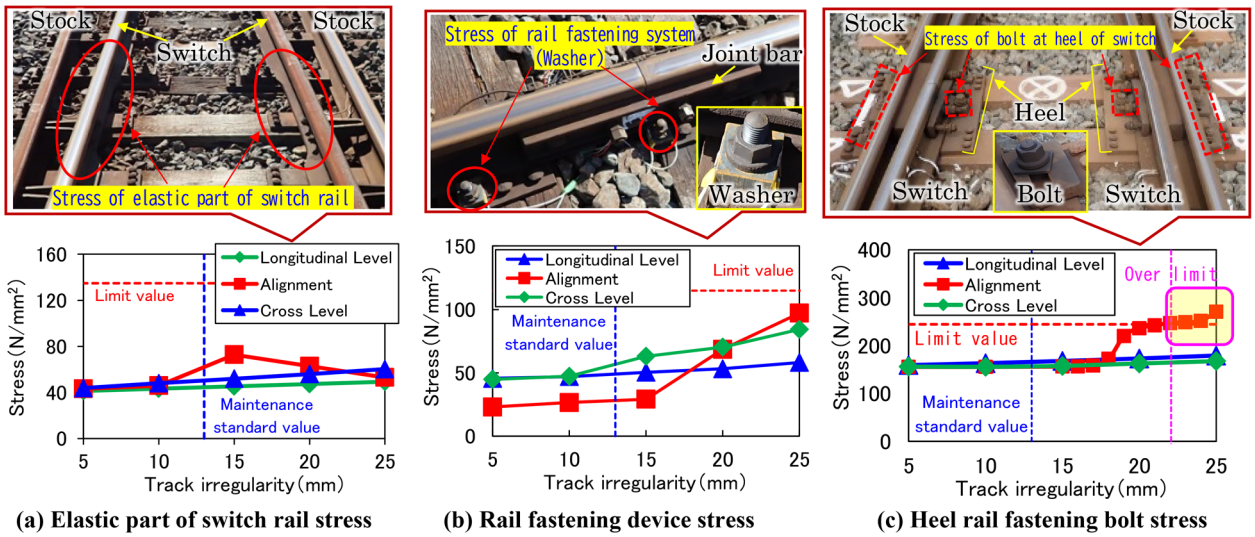


Fig. 2 Example waveform of track irregularity to be applied

3.3 Evaluation of the effects of track irregularity on turnout component damage

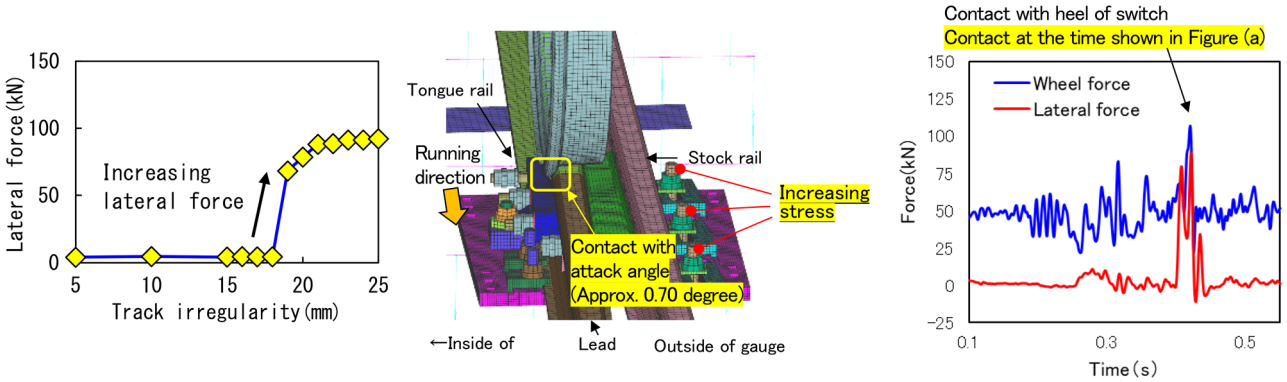
3.3.1 Evaluation results for track components

Figure 3 shows the analysis results regarding the effects of track irregularity on track component damage. Within the range below the current maintenance standard values, the generated stresses for all control items were sufficiently below their respective limit values. Next, regarding cases where the track irregularity exceeded the maintenance standard values, the generated stresses for all cases remained below the limit values with respect to longitudinal level



(a) Elastic part of switch rail stress (b) Rail fastening device stress (c) Heel rail fastening bolt stress

Fig. 3 Analysis results of effect of track irregularity on damage to track components



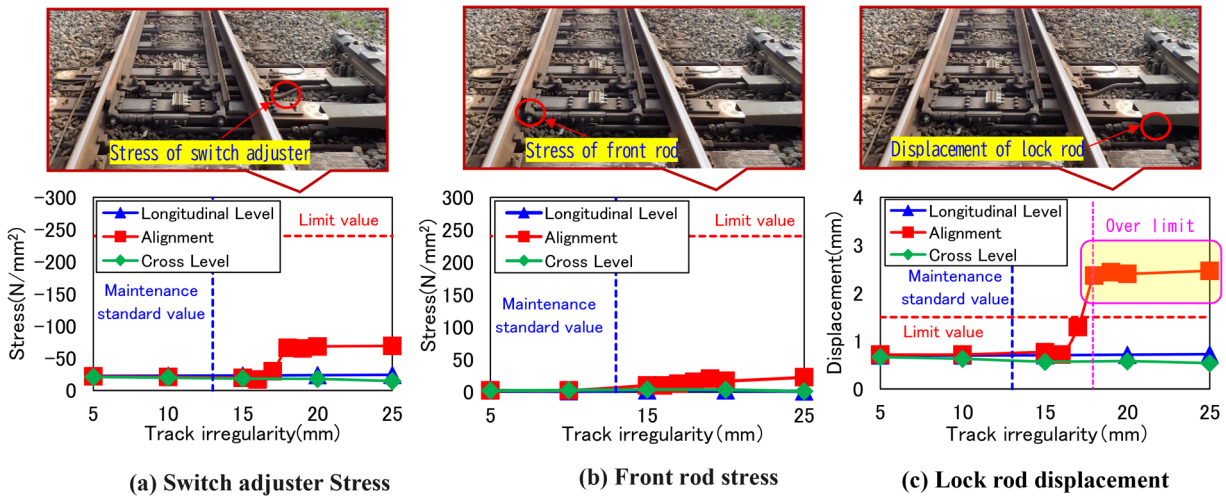
(a) Relationship between alignment irregularity and lateral force (b) Wheel–heel contact condition (c) Generation of wheel load and lateral force

Fig. 4 Mechanism of increased heel rail fastening bolt stress

and cross level irregularities. However, the rail fastening bolt stress at the heeled point section exceeded the limit value when the alignment irregularity was 21 mm, which is 8 mm larger than the maintenance standard value. Here, Fig. 4(a) shows the relationship between alignment irregularity and lateral force regarding the rail fastening bolt stress. It is confirmed that while the lateral force remains small up to an alignment irregularity of 18 mm, it increases significantly from 19 mm onward. This tendency matches the increasing trend of the rail fastening bolt stress shown in Fig. 3(c). This increase in lateral force occurs because when a vehicle passes at high speed through the location where there is an alignment irregularity before the heeled point section, it has the effect of displacing the vehicle in the lateral direction. When the irregularity exceeds 19 mm, the wheel flange comes into contact with the heeled point section with an angle of attack of approximately 0.7 degrees (Figs. 4(b) and 4(c)). Under the influence of this increased lateral force, it is considered that a large load was generated on the heeled point bolts that secure the heeled point section.

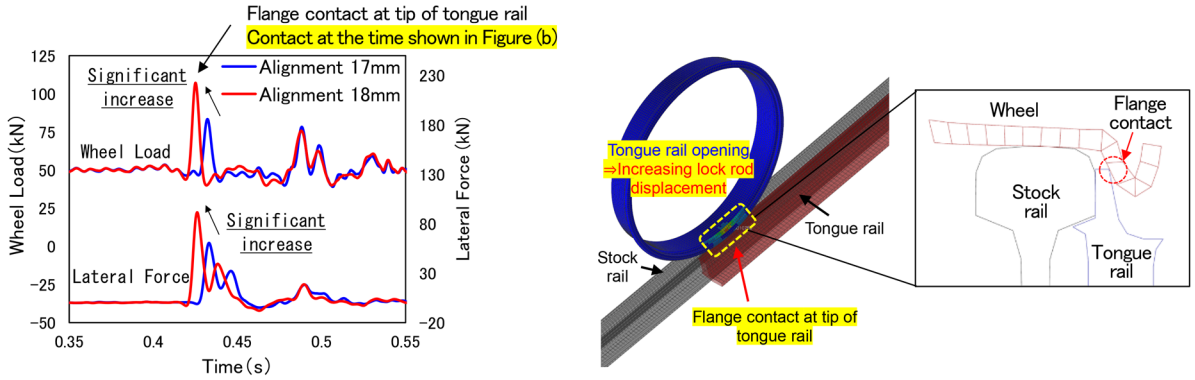
3.3.2 Evaluation results for signaling components

Figure 5 shows the analysis results regarding the effects of track irregularity on signaling component damage. Within the range below the current maintenance standard values, the generated stresses and displacements for all control items were sufficiently below their respective limit values. Next, in cases where the track irregularity exceeded the maintenance standard values, the values remained below the limit values in all cases with respect to longitudinal level and cross level irregularities. However, in the case of alignment irregularity the lock rod displacement exceeded the limit value when the irregularity was 18 mm, which is 5 mm larger than the maintenance standard value. Here, Fig. 6(a) shows a comparison of the wheel load and lateral force waveforms between the case with an alignment irregularity of 17 mm (before the lock rod displacement exceeds the limit value) and the case with 18 mm (after exceeding the limit value). The waveforms confirm that the wheel load and lateral force increase significantly when the alignment irregularity reaches 18 mm. This is because when a vehicle passes at high speed through the location where there is alignment irregularity at the front end of the turnout, this irregularity displaces the vehicle in the lateral direction. When the alignment irregularity exceeds 18 mm, the lower part of the wheel flange comes into contact with the toe of the tongue rail, as shown in Fig. 6(b). It is considered that a large lateral force was generated by this contact between the lower



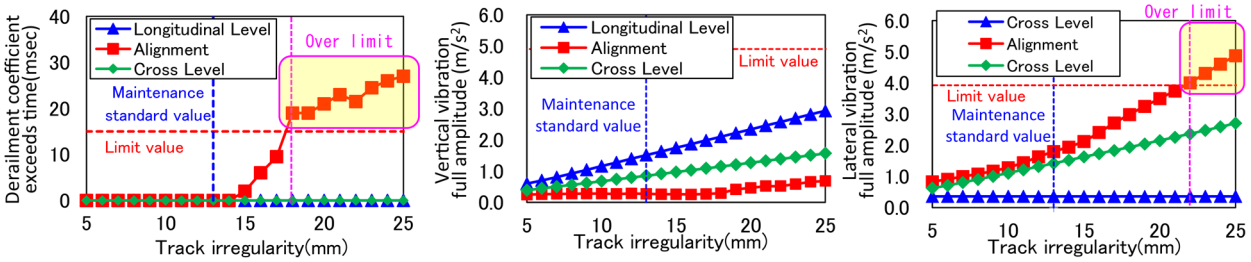
(a) Switch adjuster Stress (b) Front rod stress (c) Lock rod displacement

Fig. 5 Analysis results of the effects of track irregularity on damage to signaling components



(a) Generation of wheel load and lateral force due to alignment irregularity (b) Contact condition between the wheel flange and the tongue rail tip

Fig. 6 Mechanism of lock rod displacement increase caused by alignment irregularity



(a) Derailment coefficient exceeds time (running Safety) (b) Vehicle longitudinal level vibration (c) Vehicle lateral vibration

Fig. 7 Evaluation results of effects of track irregularity on vehicle dynamic behavior

part of the wheel flange and the toe of the tongue rail, and the resulting reaction caused the toe of the tongue rail to open, leading to a significant increase in the lock rod displacement.

3.4 Evaluation of the effects of track irregularity on vehicle dynamics

Figure 7 shows the analysis results regarding the effects of track irregularity on vehicle dynamics. Within the range below the current maintenance standard values, the values for all control items were sufficiently below their respective limit values. In cases where the track irregularity exceeded maintenance standard values, the values for longitudinal level and cross level irregularities nevertheless remained below the limit values in all cases. However, when alignment irregularity reached 18 mm, which is 5 mm over the maintenance standard value, it exceeded the limit value for derailment coefficient exceeds time. In addition, the lateral vibration exceeded its limit value when the alignment irregularity reached 22

mm, which is 9 mm over the maintenance standard value. The reason for exceeding the limit value for derailment coefficient exceeds time is considered to be that, similar to the mechanism of the lock rod displacement exceeding the limit value described in Section 3.3.2, the vehicle was shaken laterally due to the effect of the alignment irregularity, and the wheel flange came into contact with the toe of the tongue rail, generating an excessive lateral force. In cases of lateral vibration exceeding the limit value, a tendency for lateral vibration to increase in proportion to the increase in alignment irregularity has been observed. Therefore, it is considered that the lateral acceleration of the vehicle rose as the alignment irregularity increased during high-speed passage of the vehicle.

3.5 Investigation of combined track irregularities

Based on investigations reported in the previous section, it was confirmed that under conditions where longitudinal level, alignment, and cross level irregularities exist independently, the effects of each track irregularity on turnout component damage and vehicle dynamics are small as long as they remain below the current maintenance standard values. In the case of longitudinal level and cross level irregularities, the results remained below each limit value even when values larger than the current maintenance standard values were allowed. Based on the above results, because of the significant effect of alignment, it is desirable to manage this irregularity using the current maintenance standard value. However, there is a possibility that the maintenance standard values for longitudinal level and cross level irregularities can be relaxed because their effects are small. Nevertheless, investigations conducted so far have examined situations where each track irregularity exists independently. Therefore, even if the alignment irregularity is within the current maintenance standard value, if it occurs in combination with other track irregularities, its effect on component stress, displacement, and vehicle dynamics might become significant. Accordingly, this section conducted a case study assuming a situation where track irregularities occur in combination. Generally, under conditions where alignment and cross level irregularities coexist in opposite phases, wheel load falls and lateral force increases simultaneously, affecting running safety. Caution is thus required in their management [8]. Furthermore, the interaction between alignment and cross level irregularities has also been investigated in a previous study, which confirmed that when alignment and cross level irregularities occur in opposite phases, the lateral force affecting component damage reaches its maximum [9]. Therefore, this study also set a track irregularity where alignment and cross level irregularities occur in opposite phases, and evaluated the effects of this on component stress, displacement, and vehicle dynamics. When longitudinal level and alignment irregularities are combined, longitudinal level irregularity does not cause vehicle rolling, and its effect on running safety is considered small. In addition, because rolling does not occur, lateral irregularity of the vehicle does not occur either and no increase in lateral force is generated. Thus, its effect on component damage is also considered small. Based on this, longitudinal level irregularity was excluded from the investigation.

The track irregularity wavelength was set at 10 m. The alignment irregularity was set to 13 mm, which is the current maintenance standard value where the response values of each evaluation item were within the limit values in the investigations up to the previous section, and the cross level irregularity was set in the range of 5 mm to 25 mm. The vehicle running speed was set to 120 km/h, and the direction and setting position of the track irregularity were the same as in the investigation reported in the previous section.

Figure 8 shows the analysis results for the effect on track component damage, and Fig. 9 shows the analysis results for the effect on signaling component damage. Under all conditions, the generated stresses and displacements remained below the limit values. However, stress in the tongue rail elastic part, the rail fastening system, and the heeled point bolt, tended to increase as the cross

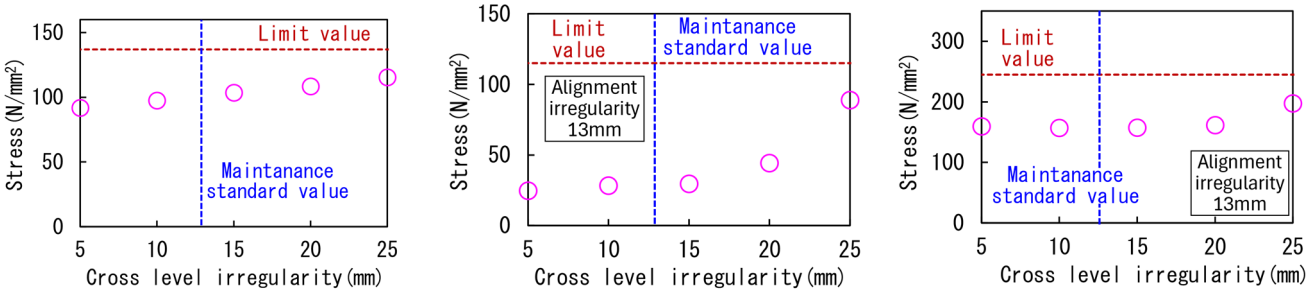


Fig. 8 Evaluation results of effects of combined track irregularities on damage to track components

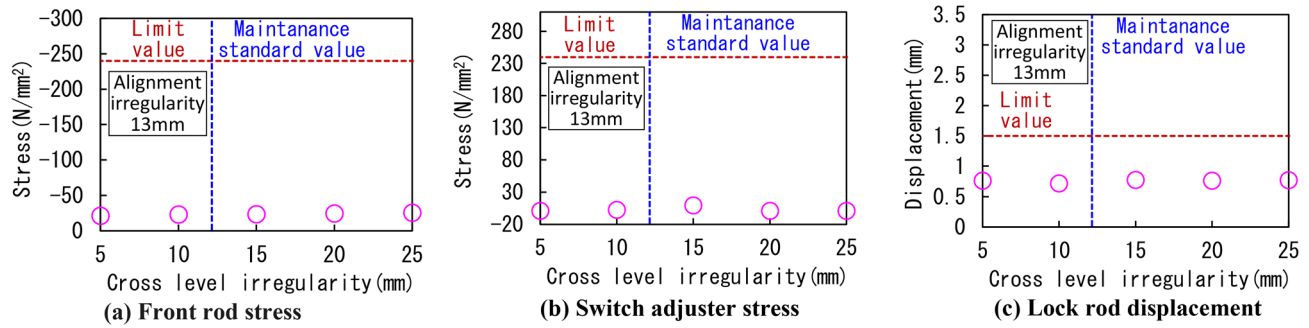


Fig. 9 Evaluation results of effects of combined track irregularities on damage to signaling components

level irregularity increased. In particular, stress increased significantly in the rail fastening system and the heeled point bolt (Figs. 8(b) and 8(c)) when the cross level irregularity exceeded 20 mm. Although both target the joint, it is considered that when the irregularity exceeds 20 mm, the vehicle rolling and the angle of attack increase, and the wheel flange comes into contact with the joint section, thereby increasing the lateral force. For the signaling components, no increase in generated stress or displacement due to the increase in cross level irregularity was observed. As mentioned above in Section 3.3.2, the increase in stress and displacement of signaling components is caused by the vehicle being displaced in the lateral direction due to the effect of alignment irregularity, and the lower part of the wheel flange coming into contact with the toe of the tongue rail, generating excessive lateral force. On the other hand, in this analysis, since the alignment irregularity is fixed at 13 mm, which is the current maintenance standard value, the effect of the alignment irregularity is small, and no contact between the wheel flange and the toe of the tongue rail was observed even when the cross level irregularity increased. For this reason, it is considered that the effect on the stress and displacement of signaling components was small.

Figure 10 shows the analysis results for the vehicle dynamics evaluation. The derailment coefficient exceeds time for running safety (Fig. 10(a)) was 0 msec under all conditions. The longitudinal level vibration of the vehicle (Fig. 10(b)), tended to increase with the increase in cross level irregularity. However, even when the cross level irregularity was 25 mm, it was 0.16 G, which is 32% of the limit value of 0.5 G, and was sufficiently small. Vehicle lateral vibration (Fig. 10(c)) was 0.41 G when the cross level irregularity reached 24 mm, resulting in a value exceeding the limit value. This is considered to be because the vehicle passed through the location of the cross level irregularity while being shaken laterally due to the high-speed passage through the location with alignment irregularity, and the further accumulation of vehicle rolling increased the lateral vibration. However, the maintenance standard value for cross level irregularity in general turnouts is 23 mm, and it was found that the values remain within the limit values up to that level.

3.6 Discussion

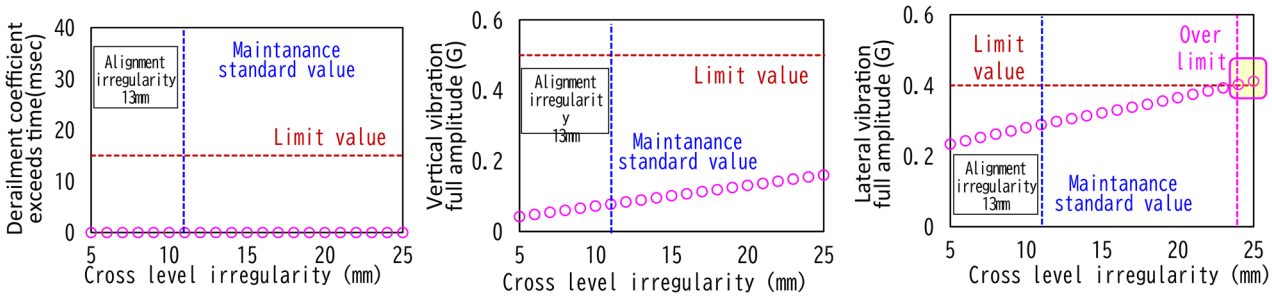
A summary of the evaluation results in this investigation is shown in Table 5. Under conditions below the current maintenance standard values (longitudinal level irregularity: 13 mm, alignment irregularity: 13 mm, cross level irregularity: 11 mm), the values for all control items were below their respective limit values, revealing that the effects on turnout component damage and vehicle dynamics are small. Furthermore, regarding longitudinal level and cross level irregularities, even when the irregularity was 25 mm, which is significantly larger than the maintenance standard values, the values remained below the limit values for each control item, confirming that the effects are small. On the other hand, regarding alignment irregularity, when the irregularity was 18 mm, which is 5 mm larger than the current maintenance standard value (Table 1), the lock rod displacement of the signaling component and the derailment coefficient exceeds time related to running safety exceeded their limit values. Furthermore, when the irregularity was 21 mm, which is 8 mm larger than the current maintenance standard value (Table 1), the rail fastening bolt stress of the track component at the heeled point section exceeded the limit value. Moreover, when the irregularity was 22 mm, which is 9 mm larger than the current maintenance standard value (Table 1), the lateral vibration of the vehicle exceeded the limit value. Based on these results, since alignment irregularity affects component damage and vehicle dynamics when it becomes larger than the maintenance standard value, it is considered desirable to manage alignment irregularity within the current maintenance standard value.

Furthermore, assuming a situation where track irregularities occur in combination, an investigation was conducted by combining an alignment irregularity fixed at the current maintenance standard value of 13 mm with a cross level irregularity ranging from 5 mm to 25 mm. As a result, it was confirmed that the lateral vibration of the vehicle exceeded the limit value when the cross level irregularity reached 24 mm. However, the values remained below the limit value up to 23 mm, which is the maintenance standard value for cross level irregularity in general turnouts. In addition, for other items regarding component damage and vehicle dynamics, it was confirmed that the values remained below their respective limit values under all conditions.

Based on the above results, under the conditions investigated in this study, it is considered that the maintenance standard values for longitudinal level and cross level irregularities in high-speed type turnouts can be allowed up to 23 mm, which is the same as the maintenance standard value for general turnouts that are not high-speed type turnouts.

4. Conclusions

In this study, turnout running simulation models were constructed with the objective of evaluating the effects of track irregularity in high-speed type turnouts on turnout component damage and vehicle dynamics. Furthermore, using the constructed models, turnout running simulations were conducted using track irregularity as a parameter, and the effects of track irregularity in high-speed



(a) Derailment coefficient exceeds time (running safety)

(b) Vehicle longitudinal level vibration

(c) Vehicle lateral level vibration

Fig. 10 Evaluation results of effects of combined track irregularities on vehicle dynamics

Table 5 Summary of evaluation results

Evaluation target	Evaluation item	Evaluation index	Evaluation results			
			Longitudinal level irregularity*	Alignment irregularity*	Cross level irregularity*	Alignment & cross level irregularities**
Turnout components	Track components	Tongue rail elastic section stress	No exceedance	No exceedance	No Exceedance	No exceedance
		Rail fastening device stress				
		Heel section bolt stress		Exceeded at amplitude of 21 mm		
	Signaling components	Lock rod displacement	No Exceedance	Exceeded at amplitude of 18 mm	No Exceedance	No exceedance
		Switch adjuster stress		No exceedance		
		Front rod stress				
Vehicle dynamics	Running safety	Duration exceeding the derailment coefficient threshold	No Exceedance	Exceeded at amplitude of 18 mm	No Exceedance	No exceedance
	Vehicle ride motion	Longitudinal level vibration		No exceedance		
		Lateral vibration		Exceeded at amplitude of 22 mm		Exceeded at a cross level irregularity of 24 mm

* Evaluated for track irregularity amplitudes ranging from 5 mm to 25 mm.
** Alignment irregularity fixed at 13 mm; cross level irregularity varied from 5 to 25 mm for evaluation.

type turnouts on turnout components and vehicle dynamics were quantitatively evaluated. The main results are as follows.

(1) It was confirmed that within the range below the current maintenance standard values for high-speed type turnouts (longitudinal level irregularity: 13 mm, alignment irregularity: 13 mm, cross level irregularity: 11 mm), the values for all control items regarding turnout components and vehicle dynamics were below their respective limit values.

(2) To confirm the limits of allowable track irregularity, a case study was conducted setting values larger than the maintenance standard values. As a result, it was confirmed that the values for longitudinal level and cross level irregularities remained below the limit values for each control item even under the condition of 25 mm. However, for alignment, it was confirmed that when this irregularity reached 18 mm or more, the values exceeded the limit values for component stresses, displacements, and vehicle dynamics.

(3) Assuming a track irregularity where alignment and cross level irregularities occur in combination in opposite phases, running analyses were conducted using the cross level irregularity as a parameter while fixing the alignment irregularity at the maintenance standard value of 13 mm. As a result, it was confirmed that the values remained below the limit values in all cases for the turnout component control items, derailment coefficient exceeds time and longitudinal level vehicle vibration. However, it was confirmed that the lateral vibration of the vehicle exceeded the limit value when the cross level irregularity reached 24 mm or more.

(4) Based on these results, under the conditions investigated in this study, it was confirmed that while it is desirable to manage alignment irregularity within the current maintenance standard value, there is no problem in allowing the maintenance standard values for longitudinal level and cross level irregularities up to 23 mm, which is the same as the maintenance standard value for general turnouts.

References

[1] Japanese Railway Engineer's Association, *Research on new track irregularity management methods*, pp. 89–93, 1982 (in Japanese).

[2] Shiota, K., and Sonoda, Y. "Evaluation of the effect of track Irregularity on damage to track components in high speed type turnouts", J-RAIL2024, SS2–8–4, 2024 (in Japanese).

[3] Shiota, K., Miyashita, A., and Sonoda, Y., "Study on the effect of track irregularity on vehicle dynamics at high speed passage of turnouts, Journal of railway engineering", Vol. 28, No. 1, pp. 108–115, 2024 (in Japanese).

[4] Railway Technical Research Institute, *An outline of a manual of conventional railway operation speed improvement test*, pp. 23–32, pp. 53–56, p. 61, 1993 (in Japanese).

[5] Board of track maintenance engineering, *HosenKougaku (Jou)*, pp. 383–386, 2016 (in Japanese).

[6] Ishida, H., Matsuo, M. and Fujioka, T., "Safety Assessment Method of Railway Vehicle under Oscillatory Wheel Load Fluctuation", Transactions of the JSME, Vol. 71, No. 702, pp. 454–461, 2005 (in Japanese).

[7] Shiota, K., Nakahashi, J., "Basic study on the effects of track irregularity in turnouts on running safety and track components

strength", Journal of railway engineering, Vol. 27, No. 1, pp. 147–154, 2023 (in Japanese).
[8] Board of track maintenance engineering, *HosenKougaku (Jou)*, pp. 206–208,2016 (in Japanese).
[9] Shiota, K., Shiomi, S., and Sonoda, Y., "Investigation of the effect of track irregularity on the damage to turnout components at high-speed type turnout", Transactions of the JSME, Vol. 94, No. 951, 2025 (in Japanese).

Authors



Katsutoshi SHIOTA, Ph.D.
Senior Researcher, Track Structures and Components Laboratory, Track Technology Division
Research Areas: Turnout technology, Running simulation



Ayano MIYASHITA
Researcher, Track Structures and Components Laboratory, Track Technology Division
Research Areas: Turnout technology, Running simulation

Simulation and Parametric Study of Ground Behavior Induced by Chemical Grouting Considering Permeation and Fracture Grouting Mechanisms

Yudai YAMASHITA **Takashi NAKAYAMA** **Takahiro ISHII**
Tunnel Engineering Laboratory, Structures Technology Division

Increasing use of underground spaces is driving demand for chemical grouting around existing structures. This increases the importance of controlling ground displacement during grouting operations. To address this, a numerical analysis method capable of estimating ground behavior prior to construction is required. This study extends a numerical prediction model previously limited to permeation grouting, to include fracture grouting. The proposed method is validated through comparison with the results of water injection tests. The final section presents the results of a parametric study using the proposed model.

***Key words*:** *chemical grouting, permeation grouting, fracture grouting (or hydrofracture grouting depending on context), parametric study, finite element analysis (FEA)*

1. Introduction

Chemical grouting is a ground improvement method in which grout materials with controllable setting times (grout) are injected into soil voids to increase its strength and seepage resistance (Fig. 1). Chemical grouting is frequently used in applications such as excavation support works, shield tunneling, and undertrack crossing construction, where it serves to suppress ground deformation and mitigate the displacement of adjacent structures. Owing to the compact size of the equipment and its high operational efficiency, this method has been widely applied, particularly at sites with limited working space and time constraints [1], [2].

However, because injection pressure is exerted on the ground during grouting, ground and structural deformation may occur depending on site conditions. For example, in a case where grouting was carried out to suppress deformation during the passage of a shield tunnel beneath a railway viaduct, monitoring results showed that the track heaved by approximately 2-3 mm per day [3]. In another case involving grouting beneath a station building, an abutment uplift of about 70 mm was reported [4].

Accordingly, guidelines such as the Manual for Construction Measures near Urban Railway Structures [3] classify potential impacts and prescribe corresponding countermeasures based on the extent of the injection zone and its distance from existing structures. For instance, when grouting is performed adjacent to a structure, harmful effects may occur under normal construction conditions if the clearance between the grout zone and the structure is less than 1.5 times the injection depth. Prior countermeasures and impact assessment are considered essential within this range, regardless of overburden thickness or ground conditions.

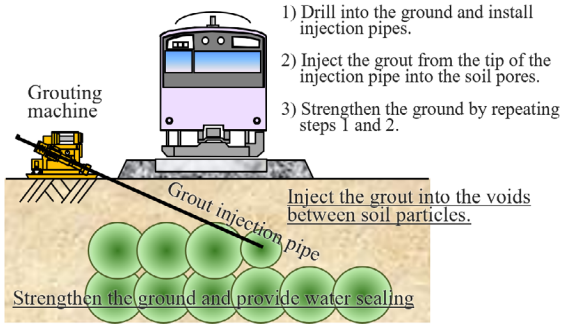


Fig. 1 Overview of chemical injection process

Impact prediction can be carried out either through in-situ injection tests or numerical analysis; this study focuses on prediction methods based on numerical analysis.

Research on in-situ injection tests includes, for example, studies by Mori et al [5]. and Fukui et al [6]., who, based on laboratory experiments and field data, proposed relationships between injection pressure and the initiation and propagation of fractures in fracture grouting, where detrimental ground deformation is likely to occur.

In terms of numerical analysis, Akagi et al [7]. created a mathematical model of fractures generated during fracture grouting using joint elements and proposed a ground deformation analysis method that applies pressure to these fractures. The authors have previously proposed a three-dimensional ground deformation analysis method for permeation grouting [8], as well as its two-dimensional formulation [9].

Although numerical studies have been conducted depending on the grouting type, a generalized and practical prediction method for evaluating ground behavior induced by chemical grouting has not yet been established.

This study proposes a method to extend a simulation approach originally developed for permeation grouting to fracture grouting by applying a simplified correction. First, an overview and the results of a water injection test conducted on a test embankment to validate the proposed method are presented. Next, the proposed method is outlined, followed by a comparison of ground surface displacements obtained from experiments and simulations. Finally, the results of a parametric study on surface heave are presented for different types of commonly used chemical grouting methods.

2. Water Injection Experiments Simulating Permeation and Fracture Grouting

2.1 Overview of Injection Tests

In this study, water injection experiments were conducted on a test embankment. Table 1 summarizes the geotechnical properties of the embankment. In these experiments, the soil properties were selected so that both permeation grouting and fracture grouting could be reproduced under relatively low injection pressures.

Figures 2 and 3 illustrate the experimental setup and injection

Table 1 Geotechnical properties of the test embankment

Particle diameter D_{20} (mm)	0.142
Soil particle density ρ_s (kN/m^3)	26.6
Wet density ρ_t (kN/m^3)	16.1
Dry density ρ_d (kN/m^3)	14.3
Water content ω (%)	13.0
Deformation modulus E_{50} (CD) (kN/m^2)	12,000

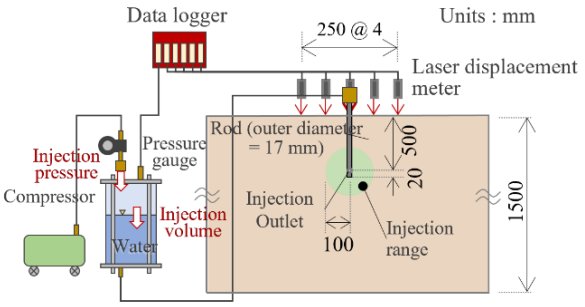


Fig. 2 Schematic diagram of water injection test

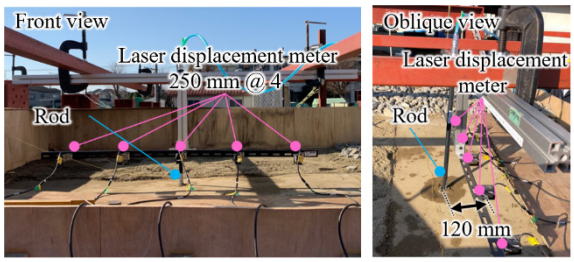


Fig. 3 Water injection test in progress

procedure. For the injection method, a rod with a predrilled injection port was inserted into the ground to a depth of 500 mm. Water was then injected at a constant pressure by applying compressed air pressure to a cylinder filled with water.

The injection volume was defined as the amount of material required to fill a spherical zone with a radius of 100 mm centered at the injection point, expressed in terms of porosity. Ground surface displacement was measured using a laser displacement meter at five measurement points arranged in a straight line at 250 mm intervals on both sides of the rod, with a lateral offset of 120 mm from the rod.

2.2 Results of Water Injection Tests

(1) Relationship between Injection Rate and Injection Pressure

Figure 4 shows the relationship between injection rates and injection pressure. The figure also includes the relationship calculated using Maag's equation [10]. Maag's equation is a theoretical expression for permeation grouting, based on the assumption that excess pore water pressure propagates radially from the injection point and equilibrates with hydrostatic pressure at an infinite distance.

The coefficient of permeability of the test embankment was estimated to be 3.9×10^{-5} m/s based on the relationship between D20 particle size and permeability. The experimental results are in reasonable agreement with the values calculated using Maag's equation at injection rates of approximately 4-6 L/min, indicating that permeation grouting consistent with the assumptions of Maag's theory was reproduced in this range.

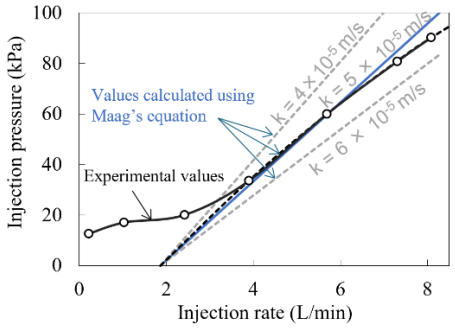


Fig. 4 Relationship between injection rate and injection pressure

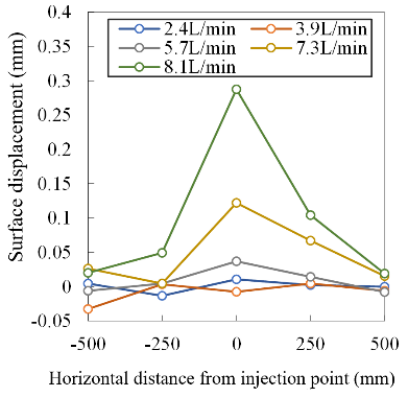


Fig. 5 Relationship between horizontal distance from injection point and surface displacements

In contrast, outside the range of 4-6 L/min, the slopes of the experimental and theoretical relationships differ. At injection rates of approximately 0-4 L/min, this discrepancy is attributed to the initial conditioning process associated with the unsaturated state of the test embankment. At injection rates exceeding approximately 6 L/min, the slope decreases and the curve becomes more gradual. This suggests that, in this range, the injection mode begins to transition from permeation grouting to fracture-permeation grouting.

(2) Distribution of Surface Displacement and Evolution of Maximum Surface Displacement

Figure 5 shows the distribution of ground surface displacement measured by each laser displacement meter. The largest displacement was observed directly above the injection point (horizontal distance of 0 mm), and the displacement decreased with increasing distance from the injection point.

Figure 6 shows the relationship between injection rates and maximum surface displacement. The relationship between injection rates and injection pressure (from Figure 4) is also shown for reference. When the injection rate exceeds 6 L/min, surface heave increases with increasing injection rates.

Previous studies have reported that surface displacement tends to increase once fracture grouting begins. Since the reduction in the pressure gradient and the onset of heave occur almost simultaneously, it is inferred that the injection state transitions to the early stage of fracture-permeation grouting at around 6 L/min.

3. Extension of Ground Behavior Simulation Method to Fracture Grouting

3.1 Overview of Conventional Simulation Method for Permeation Grouting

In permeation grouting, the injected material and pore water infiltrate the surrounding ground, generating hydraulic pressure (seepage pressure) that acts on the soil skeleton. In the method proposed by the authors, this coupled problem of seepage and ground deformation is formulated within a total stress framework. The seepage pressure is evaluated using the excess pore water pressure derived from Maag's equation (Fig. 7).

Maag's equation is expressed as Eq. (1). The injection pressure is given by $p(r_a, t)$, which is obtained by substituting the injection point radius r_a for r.

$$\begin{cases} p(r,t)=\dfrac{\gamma\cdot q}{4\pi k_w}\cdot\left\{\left(\dfrac{\mu_g}{\mu_w}\right)\dfrac{1}{r}+\left(1-\dfrac{\mu_g}{\mu_w}\right)\dfrac{1}{R(t)}\right\}+\gamma\cdot h_w & (r\le R(t)) \\ p(r,t)=\dfrac{\gamma\cdot q}{4\pi k_w}\cdot\dfrac{1}{r}+\gamma\cdot h_w & (r>R(t)) \end{cases} \quad (1)$$

Here, $p(r, t)$ denotes the pressure at a radial distance r (kPa); γ is the unit weight of water (kN/m^3); h_w is the groundwater level near the injection point (m); q is the injection rate (m^3/s); k_w is the coefficient of permeability of water (m/s); μ_w is the viscosity of water (Pa·s); μ_g is the viscosity of the grout (Pa·s); and $R(t)$ is the radius of influence at time t (m).

The radius of influence $R(t)$ at time t is given by Eq. (2), based on the study by Karol et al [11].

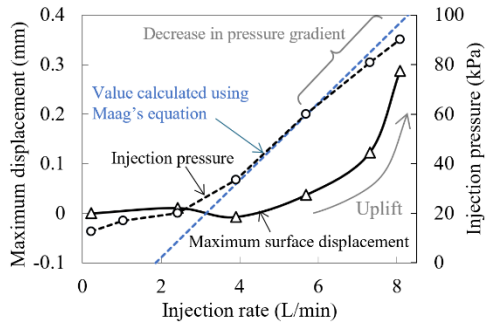


Fig.6 Relationship between injection rate and maximum surface displacement

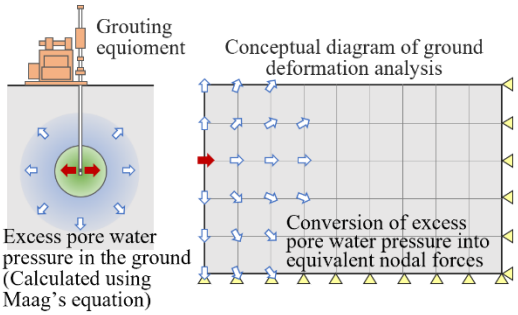


Fig. 7 Overview of excess pore water pressure and ground deformation analysis

$$R(t)=\sqrt[3]{r_a^{\,3}\frac{300\cdot q\cdot t}{4\pi\cdot n}} \quad (2)$$

Here, n denotes the porosity (%), and r_a is the radius of the injection point (m).

Permeation in chemical grouting is inherently a three-dimensional phenomenon, and the excess pore water pressure distribution obtained from Maag's equation is spherical. To account for this in two-dimensional analyses, the authors have previously proposed a correction method based on rotational transformation [9].

Using this approach, ground surface displacement under permeation grouting conditions can be reproduced successfully. However, a challenge remains in that the method cannot adequately capture the increase in surface displacement that occurs once fracture grouting begins.

3.2 Proposed Extension for Fracture Grouting

In chemical grouting, the injection mode depends on the injection rate: at low injection rates, permeation grouting predominates; however, as the injection rate increases, the ground begins to fracture due to the injection pressure, leading to a transition to fracture-permeation grouting, in which both permeation and fracture mechanisms coexist. The initial stage of fracture-permeation grouting is illustrated in Fig. 8.

During fracture-permeation grouting, not only the region around the injection point but also the fracture surfaces act as seepage paths. Accurately evaluating the increase in seepage area and the anisotropy induced by fracturing is essentially equivalent to determining the number and orientation of fracture veins, which is highly challenging.

To address this issue, the present study simplifies the fractured zone as a circular region and applies the injection pressure at its boundary, thereby providing a conservative evaluation.

The fracture initiation pressure P_f, for sandy soils is given by Eq. (3), based on the study by Mori et al [12].

$$P_f=m\cdot\sigma_3'+\sigma_t+R \quad (3)$$

Here, m is a constant; σ_3' is the minimum effective principal stress at the injection point (kN/m^2); σ_t is the tensile strength of the ground (kN/m^2); and R represents the resistance to fracture propagation.

When the excess pore water pressure $p(r, t)$, obtained from Maag's equation in Eq. (1), exceeds the fracture initiation pressure P_f given by Eq. (3), fracturing of the ground is assumed to initiate. The region inside the boundary defined by $P_f = p(r, t)$ (i.e., closer to the injection point) is then assumed to transition to fracture-permeation grouting, as illustrated in Fig. 9.

In Maag's formulation, the seepage surface is assumed to be

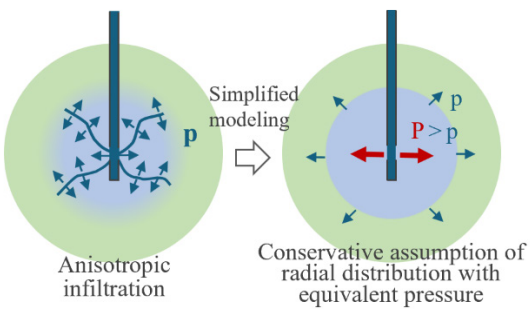


Fig. 8 Conceptual image of initial-stage fracture-permeation grouting in the ground analysis

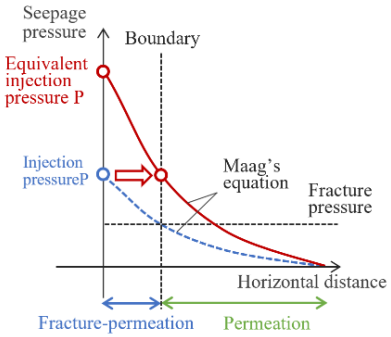


Fig. 9 Correction of seepage pressure for transition to fracture grouting

located at the radius of the injection point, and the seepage pressure at this radius is taken as the injection pressure. However, when permeation grouting and fracture-permeation grouting coexist, the seepage surface is considered to shift to the boundary between these regions. Accordingly, the excess pore water pressure distribution obtained from Maag's equation is modified by shifting it so that a seepage pressure equivalent to the injection pressure acts at this boundary.

The constant m in Eq. (3) has been reported to range from 1 to 2 for typical sandy soils. On the other hand, results from constant-rate injection tests on sandy soils have shown that, when fracturing occurs, the effective injection pressure $P_f'(= P_f - u_0$, where u_0 is the initial pore water pressure) can reach approximately 4-5 times the effective confining pressure $\sigma_c(=\sigma_c - u_0)$ [13].

Accordingly, simulations of the water injection experiments were performed for two cases, $m = 2$ and $m = 5$.

4. Validation of the Proposed Method through Simulation of Experiments

4.1 Verification of the Relationship between Horizontal Distance from the Injection Point and Surface Displacement

Figure 10 shows the relationship between horizontal distance from the injection point and ground surface displacement obtained from the simulations. Consistent with the experimental results, the simulations show that the maximum surface displacement occurs at a horizontal distance of 0 mm from the injection point, and that the displacement decreases with increasing distance from the injection point.

The results obtained using the conventional method (Fig. 10(a)) show that, under permeation grouting conditions with injection rates up to 5.7 L/min, the simulated surface displacements are larger than the experimental values, thus providing a conservative estimate. However, in the range where permeation grouting and fracture-permeation grouting coexist (injection rates of 7.3 L/min and above), the experimental values exceed the simulation results at locations near the injection point. This observation is consistent with the known limitation of the conventional method, namely its inability to adequately capture the increase in surface displacement that occurs once fracture grouting begins.

4.2 Verification of the Relationship between Injection Rate and Maximum Surface Displacement

Figure 11 shows the relationship between injection rates and maximum surface displacement. The proposed method yields conservative results compared with the experimental data, while successfully capturing the trend in maximum surface displacement.

Comparing the cases of $m = 2$, and $m = 5$, in the proposed meth-

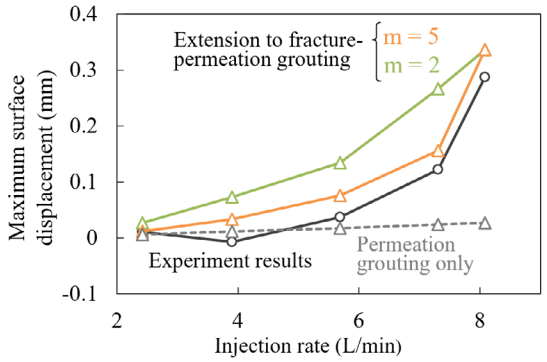
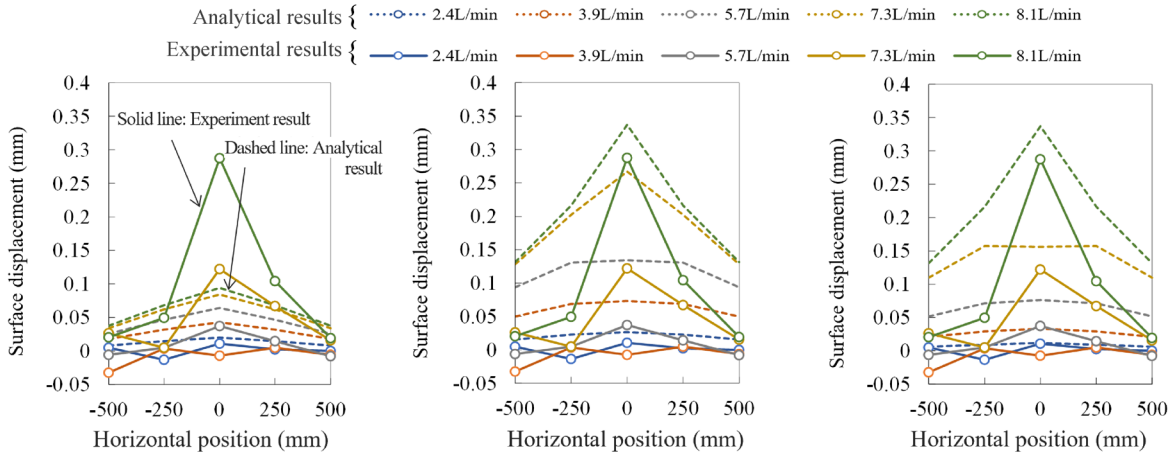


Fig. 11 Relationship between injection rate and maximum surface displacement



(a) Permeation grouting only (b) Extended to fracture-permeation grouting (m = 2) (c) Extended to fracture-permeation grouting (m = 5)

Fig. 10 Surface displacement versus horizontal distance from injection point

od, the results indicate that, under the conditions of the present water injection tests, the case of $m = 5$ shows better agreement with the experimental data than $m = 2$, which is commonly cited as a typical value for sandy soils in the literature.

These findings suggest that using $m = 2$, which corresponds to the typical range for sandy soils, provides conservative estimates. At the same time, selecting an appropriate value of m depending on the specific soil conditions may enable more accurate reproduction of ground behavior.

5. Parametric Study Using the Proposed Method

As shown in Table 2, five types of ground were defined by varying the coefficient of permeability. The selected permeability values generally correspond to soil types ranging from silt to gravel. The injection zone was arranged in a single vertical line from a depth of 2 m to 8 m, and a downward injection sequence was adopted. Two grouting methods were used: a double-pipe strainer method (injection rate: 12 L/min) and a double-pipe double-packer method (injection rate: 8 L/min).

First, the detailed analytical results are presented using the double-pipe strainer method with Ground C as an example (Figs. 12 and 13).

Figure 12 shows the relationship between injection pressure and depth, together with the fracture pressure. In this case, the injection pressure exceeds the fracture pressure at depths shallower than 6 m, indicating that fracture grouting is likely to occur in this zone.

Figure 13 shows contours of ground displacement, indicating that ground heave develops directly above the injection point. The maximum uplift is approximately 0.59 mm.

Figures 14 and 15 summarize the maximum surface heave directly above the injection point for all cases, from Ground A to Ground E. The results indicate that, as generally expected, smaller coefficients of permeability and higher injection rates tend to produce larger ground surface displacements.

Table 2 Ground conditions used in analysis

Unit weight (kN/m^3)		19
Deformation modulus (kN/m^2)		32800
Internal friction angle (°)		30
Coefficient of permeability (m/s)	Ground A	8.410×10^{-6}
	Ground B	2.103×10^{-5}
	Ground C	4.205×10^{-5}
	Ground D	6.308×10^{-5}
	Ground E	8.410×10^{-5}

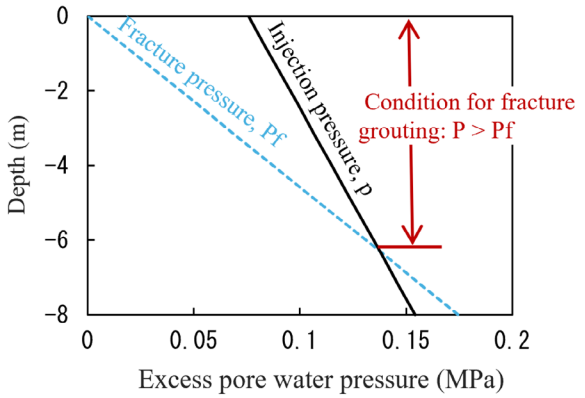


Fig. 12 Relationship between injection pressure and depth

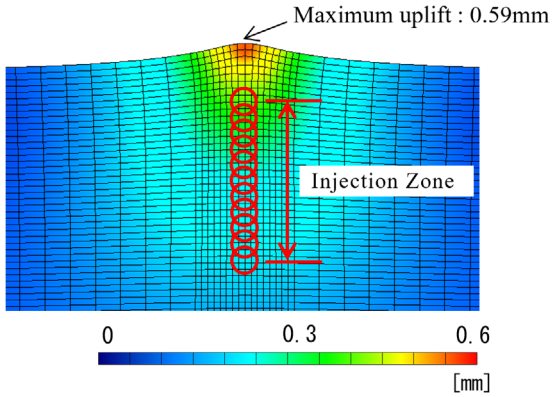


Fig. 13 Contour plot of ground surface displacement due to injection

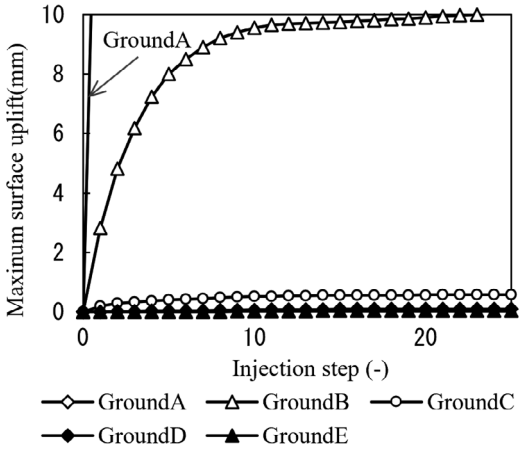


Fig. 14 Analysis results (Double-pipe strainer method)

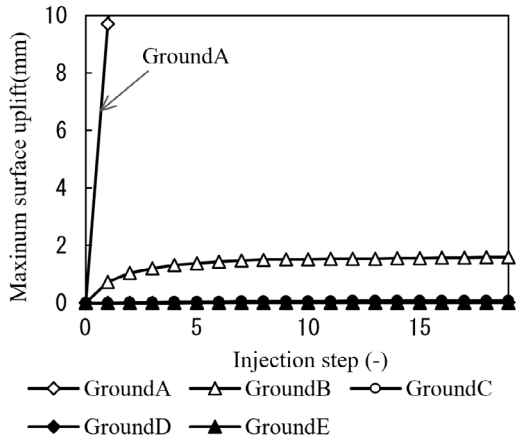


Fig. 15 Analysis results (Double-pipe double-packer method)

6. Conclusions

The main findings of this study are summarized as follows:

(1) By applying a simple and conservative correction to a simulation method originally developed for permeation grouting, the proposed approach successfully reproduces ground behavior over a range from permeation grouting to fracture-permeation grouting.

(2) A parametric study was conducted using the proposed method, which involved varying the coefficient of permeability for different chemical grouting methods. The results demonstrate that the proposed method allows for a straightforward estimation of ground surface displacement.

Future work will focus on comparing predicted surface heave with field measurements from actual construction sites in order to improve the accuracy of the proposed method.

References

[1] RTRI, "*Grouting Design and Construction Manual*," 2011 (in Japanese).

[2] Nakayama, T., Ushida, T., Yakita, S. and Okano, N., "*Case study of grouting and Indoor Test of Injected Soil*," Proceedings of Tunnel Engineering, JSCE, Vol. 21, pp. 136-148, 2011 (in Japanese).

[3] RTRI, "*Manual for Mitigation Measures for Construction Adjacent to Urban Railway Structures*," 2011 (in Japanese).

[4] Society of Soil Mechanics and Foundation Engineering, Construction in Close Proximity, pp. 362-374, 1989 (in Japanese).

[5] Mori, A., Tamura, M., Komine, H. and Ogawa, Y., "*A Method of Determine the Critical Injection Rate Considering Permeated Shapes in Chemical Grouting*," Journal of Geotechnical Engineering, Vol. 33, No. 3, pp. 159-169, 1993 (in Japanese).

[6] Fukui, Y., Mori, A. and Akagi, H., "*Study on improvement of the determining method for critical injection rate qcr in chemical grouting*," Japanese Journal of JSCE, No. 658, VI-48, pp. 81-92, 2000 (in Japanese).

[7] Akagi, H., Komiya, K., Suzuki, S., Yamazaki, H., and Mori, A., *"Deformation due to secondary injection during shield work in soft clayey ground and its sinkage prevention effects*," Japanese Journal of JSCE, No. 511, III-30, pp. 13-21, 1995 (in Japanese).

[8] Nakayama, T. and Okano, N., "*Prediction Method of Ground Displacement Caused by Chemical Grouting*," RTRI REPORT, Vol. 27, No. 6, pp. 19-24, 2013 (in Japanese).

[9] Ishii, T., Nakayama, T., Miwa, A., and Shimizu, T., "*Study on simulation method of ground behavior caused by chemical Injection*, "Proceedings of annual conference of the Japan society of civil engineers," Vol. 78, III-381, 2023 (in Japanese).

[10] Tsuboi, N., "*A Study on the Current State of Chemical Grouting Methods*," Kajima Institute Publishing Co., Ltd., pp. 155-173, 1976 (in Japanese).

[11] Karol, R.H., Chemical Grouting, Dekker, pp. 106-132, 1983.

[12] Mori, A., Tamura, M. and Hirano, M., "The injected shapes by long gel time grouts in sandy ground and the governing condition," Japanese Journal of JSCE, No. 388/III-8, pp. 61-70, 1987 (in Japanese).

[13] Research Report on the Evaluation and Development of Chemical Grouting, Public Works Research Institute, 1979 (in Japanese).

Authors



Yudai YAMASHITA
Researcher, Tunnel Engineering
Laboratory, Structures Technology Division
Research Areas: Tunnel Engineering



Takahiro ISHII
Researcher, Tunnel Engineering
Laboratory, Structures Technology Division
Research Areas: Tunnel Engineering



Takashi NAKAYAMA, Dr.Eng.
Head of Tunnel Engineering Laboratory,
Structures Technology Division
Research Areas: Tunnel Engineering

Fatigue Life Evaluation of Aged Rails Considering Actual Situation on Practical Track

Mitsuru HOSODA
Rail Maintenance and Welding Laboratory, Track Technology Division

Nobutaka TAKAHASHI
Rail Maintenance and Welding Laboratory, Track Technology Division (Former)

Tadashi DESHIMARU
Rail Maintenance and Welding Laboratory, Track Technology Division

This study analyzed the gap between sleepers and ballast beds and the surface irregularities of rail heads obtained from an operating railway track, which affect the fatigue life of aged rails. Fatigue tests were conducted on aged rails exhibiting surface irregularities and on aged rails with a history of being laid on hanging sleepers. On the basis of fatigue test results, we proposed a quantitative evaluation method for the safety and control value of aged rails for extended use. This method takes into account the probability of hanging sleepers, as well as the variability of the fatigue strength of aged rails.

***Key words*:** *rail, fatigue life, hanging sleeper, rail surface irregularity, fatigue test, replacement period*

1. Introduction

Some railway operators in Japan conduct periodic rail replacement work before the rail reaches its fatigue life to prevent fatigue fractures caused by rail bending. The rail replacement period is determined by the cumulative axle load, which is derived from the multiplication of the axle load and the number of axle-passings that the rail experiences [1].

Past findings indicate that rail fatigue life can be evaluated by combining rail usage history and remaining fatigue life as shown in Figure 1. The former is the cumulative tonnage of the aged rail over time in service, and the latter is the tonnage calculated from the S-N curve obtained by fatigue tests using aged rails collected from lines in revenue service [2, 3]. Figure 2 shows the S-N curve of aged rails with a usage history of 540 million gross tonnage (MGT). These rails were obtained from fatigue tests and were collected from welded rail joints, where excess bending stress occurs at the rail bottom, excited by irregularities in the rail-head surface. The result indicates that cracks initiate at the rusty surface of the rail bottom, and that fatigue lives are not different between types of rail welds. Additionally, there is no difference in fatigue strength from the test result for a rail exposed for five years without usage history [4].

The fatigue limit is a property of metal materials: the upper limit of fatigue stress where fatigue failure never occurs [5]. For example, the structural fatigue design guidelines set a "cut-off limit" based on the S-N curve for the material used [6]. Regarding aged railway rails, the fatigue limit is considered to be around 200 MPa in amplitude, and the rails seldom cause fatigue fractures at stresses below the limit [7]. However, there is very little data on fatigue tests conducted under conditions where the stress amplitude is lower than 200 MPa. Thus, we adopted a modification in which the S-N curve slope changes by half at the crossover point where the S-N curve and the fatigue limit of stress cross, as shown in Figure 2. We use the modified S-N curve to predict rail fatigue life.

Bending stress at rail bottoms on real operation lines, however, where there is no notable irregularity in rail top surfaces and gaps between sleepers and ballast bed, may be significantly lower than 200 MPa in amplitude [2]. Such aged rails, used under the conditions above and with cumulative large tonnages, have little fatigue accumulation; the fatigue bending test with such rails may indicate that there are very few differences between the results shown in Figure 2, such as fatigue test results and the S-N curve derived from the test results. As a result, railway operators can eliminate the periodic replacement of aged rails used under stress below a set "cut-off limit."

In this study, we conduct fatigue tests on aged rails collected from lines in revenue service and obtain data on their fatigue properties. Then, we propose a method for predicting rail fatigue life. This method considers parameters that affect fatigue life prediction, such as the irregularity of the rail top surface and gaps between sleepers and ballast bed. This method also uses estimated data from measurement data, including axle acceleration and vertical track irregularity obtained from track monitoring devices [8]. We then discussed the probability of extending the rail replacement period

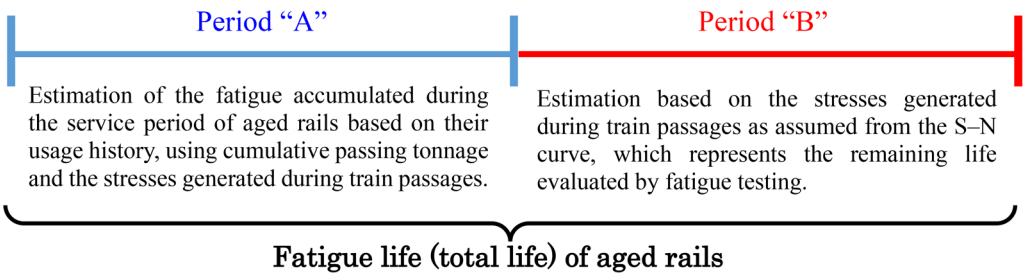


Fig. 1 Concept of fatigue life of aged rail

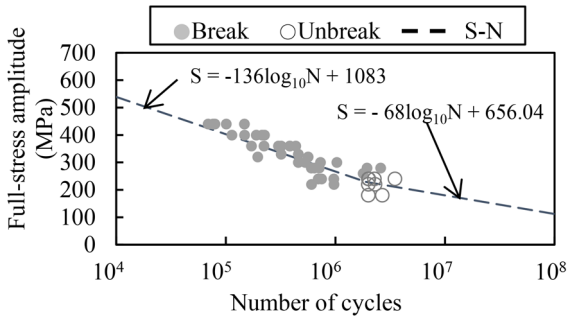


Fig. 2 S-N curve obtained from previous studies

Table 1 Overview of analyzed tracks

Track name	Track A	Track B
Rail	JIS 60kg rail	
Track length analyzed	150 km	
Track structure	Ballasted track	
Annual passing tonnage	10 million	40 million
Numbers of rail welds analyzed	16,328	9,036

based on the predicted rail fatigue life.

2. Occurrence of Hanging Sleepers and Rail Surface Irregularities

2.1 Hanging Sleepers on Lines in Revenue Service

We surveyed the actual gaps between sleepers and ballast beds using track measurement data from a continuously welded rail (CWR) section on an operating ballasted track. Table 1 shows the overview of the analyzed railway tracks. Both Track A, with an annual passing tonnage of 10 million, and Track B, with an annual passing tonnage of 40 million, adopt periodic replacement of aged rails. We estimated the gap between the sleepers and the track bed at 3 m in front of and behind each rail weld on these tracks: 16,328 for Track A and 9,036 for Track B. Using the previously proposed method [9] of dynamic track measurement data obtained via a track monitoring vehicle, we conducted a statistical analysis of the estimated gap. We applied this analysis to the track maintenance and management database system, LABOCS, developed by the Railway Technical Research Institute.

Figure 3 shows the probability histogram of gap occurrence. Both tracks show that 95% of welded rail joints have a gap of one mm or less. After fitting several distribution functions to the distribution shown in Figure 3, the Gamma distribution, which is used to model the probability of error occurrence in reliability engineering, is the most appropriate. From this result, we estimate the gap for any possible occurrence, assuming a Gamma distribution. The approximate formula of the possible density function of the Gamma distribution is shown as;

$$f(x)=\frac{1}{\beta^{\alpha}\Gamma(\alpha)}x^{\alpha-1}e^{-\frac{x}{\beta}} \quad (1)$$

where α is a shape parameter, and β is a scale parameter; these parameters are derived from the result of curve fitting to the accumulated possibilities shown in Figure 3. The approximate function derived shows good agreement with the distribution,

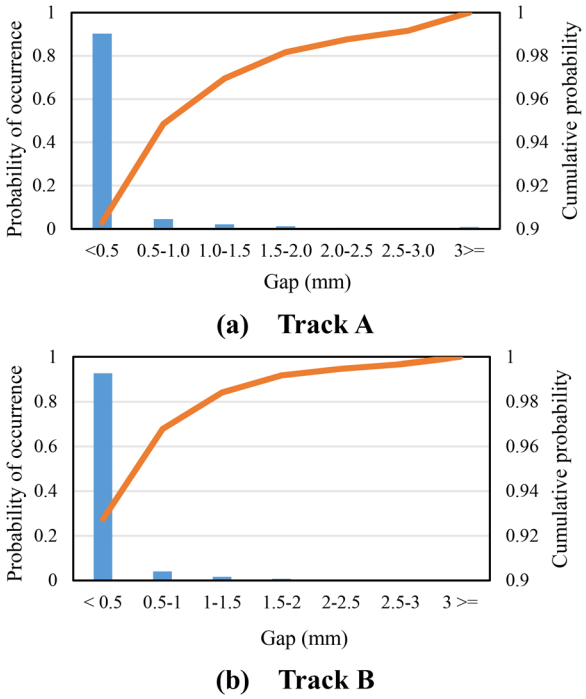


Fig. 3 Probability histogram of gap occurrence (at welded rails)

Table 2 Probability of occurrence of gaps at welded rails

Track name	Track A	Track B
Maximum value	5.8 mm	8.8 mm
Minimum value	0.11 mm	0.17 mm
α	0.11	0.11
β	1.12	1.60
Probability of occurrence of 1%	1.9 mm	2.6 mm
Probability of occurrence of 5%	0.8 mm	0.9 mm

with a coefficient of determination of 99%.

Table 2 shows the maximum and average gaps, α, β, and the estimated gap value at 5 and 1% probability of occurrence. Focusing on the cumulative probability of a Gamma distribution, the probability of occurrence of a gap of 1 mm or more is 5% or less. This result also means that the evaluation targets gaps occurring in 95% of all welded rail joints when the fatigue life is estimated at a 1mm gap. Though setting a lower probability of occurrence, such as 1%, is reasonable for a safer evaluation, it is notable that the comprehensive possibilities of occurrence are expressed as the multiplication of the probability of occurrence of rail surface irregularity and the deviation of the result of fatigue test with aged rail, as discussed later.

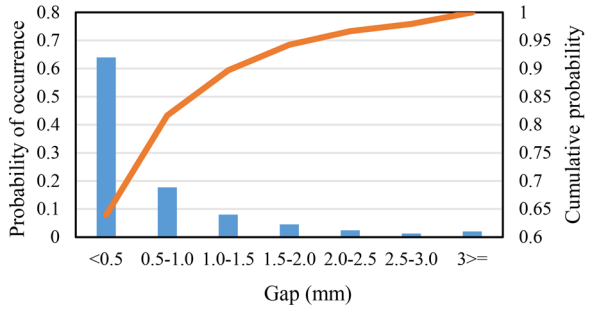
In contrast, the fatigue life of aged rails, except for welded rail joints, is expected to be shorter due to an increase in stress occurring at the bottom of the rail when the gap between the sleeper and the ballast bed is large. This occurs even though the rail surface irregularity is smaller than that occurring at welded rail joints. For this reason, we conduct the same statistical analysis for the maximum gaps obtained at every 25 m section, except for the 3 m in front of and behind each rail weld. Table 3 presents the statistical analysis of the gaps, and Figure 4 shows the histogram of the gap's probability of occurrence. Similar to the results for weld rail joints, fitting sev-

eral distribution functions showed that the gamma distribution provided a good fit for this distribution as well. Compared with the results obtained when only welded rail joints were considered, the target track is longer even within the same line section. Accordingly, the number of target locations is greater; as a result, the overall magnitude of rail uplift tends to be larger.

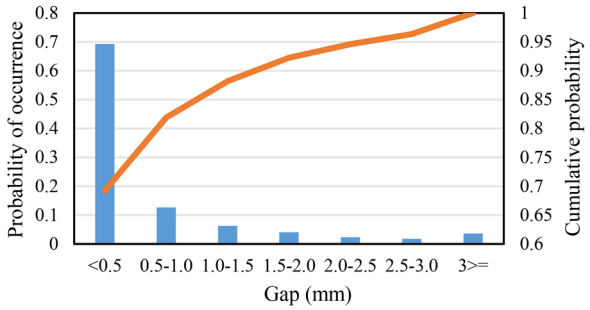
Furthermore, uplift values corresponding to occurrence probabilities of 5% and 1% are approximately 2.3-2.4 mm and 3.5-4.5 mm, respectively. Therefore, when evaluated at the same gap probability of with a 5% gap probability, similar to that of welded rail joints, the gap is estimated at around 2.5 mm. Accordingly, in the present results, the irregularity of the rail-head surface at welded rail joints does not appear to have a significant effect on the occurrence of uplifted sleepers. In this regard, factors other than head surface irregularities, such as the effects of maintenance work (e.g., tamping of the ballast bed) and variability in the sleeper support modulus along the longitudinal direction of the rail, are considered to be influential. Therefore, further investigation is required for these factors.

Table 3 Probability of occurrence of gaps except for welded rails

Track name	Track A	Track B
Maximum value	9.5 mm	12.7 mm
Minimum value	0.54 mm	0.57 mm
α	0.56	0.35
β	1.02	1.60
Probability of occurrence of 1%	3.5 mm	4.5 mm
Probability of occurrence of 5%	2.3 mm	2.4 mm



(a) Track A



(b) Track B

Fig. 4 Histogram of probability of occurrence of gaps (except for welded rails)

2.2 Rail Surface Irregularity on Operating Line

Similar to the analysis conducted for the gap between sleepers and the ballast bed, a statistical analysis is performed using data on rail-head surface irregularities. These irregularities were measured at welded rail joints in sections where such irregularities had never been removed by grinding or similar methods. These irregularities were measured using an irregularity measuring device (Fig. 5), which captures irregularities with 1 m and 100 mm chord lengths over a 1 m longitudinal range centered on the welded rail joint.

Table 4 summarizes the characteristics of the tracks analyzed for rail-head surface irregularities. The results are compiled from two tracks with annual traffic tonnages of approximately 30 million tons for Track C and 17 million tons for Track D. The analysis focuses on enclosed arc welded joints, based on previous findings [2] indicating that irregularities progress faster at enclosed arc welds than at other welding types. In addition, when calculating the fatigue life of aged rails described later, not only the magnitude of the head surface irregularities but also their progression rate (the 100 mm chord irregularity per 100 million tons of cumulative traffic at each weld) have conventionally been used as indices. Therefore, both indices were employed for evaluation in this study.

Figure 6 shows the correlation between the magnitude of rail-head surface irregularities and cumulative passing tonnage. Except for two points on Track C, it can be seen that the progression of irregularities is within 0.1 mm per 100 million tons. Next, the statistical values classified by irregularity magnitude and progression rate are presented in Table 5. In both track sections, approximately 95% of the welded joints exhibited head surface irregularities of 0.26 mm or less and progression rates of 0.05 mm per 100 million tons or less. In addition, since the frequency distribution of the rail head surface unevenness in Figure 6 showed a trend similar to that in Figure 3, the rail-head surface irregularities and the irregularity progression corresponding to a 1% occurrence probability were determined from the probability density function based on the fitted gamma distribution. The results showed that the rail head surface unevenness was 0.27-0.40 mm, and the irregularity progression was 0.08-0.09 mm per 100 million tons.



Fig. 5 Irregularity measuring device

Table 4 Overview of analyzed tracks

Track name	Track C	Track D
Annual passing tonnage	30 million	17 million
Cumulative passing tonnage	200 to 850 MGT	430 MGT
Track structure	Ballasted track	Ballastless track
Numbers of rail welds analyzed	164	100

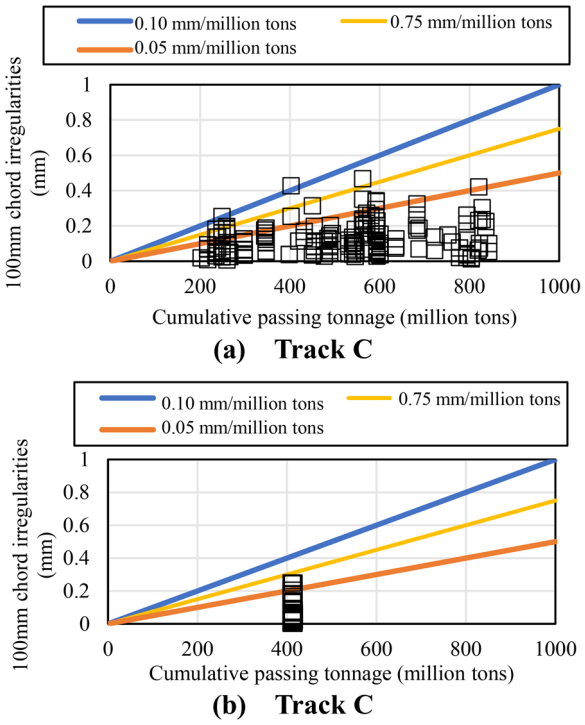


Fig. 6 Rail-head surface irregularities and cumulative passing tonnage

Table 5 Probability of occurrence of rail-head irregularities at welded rails

Sorting		By rail-head irregularity		By progression rate of irregularities	
Track name		Track C	Track D	Track C	Track D
Maximum value		0.47 mm	0.25 mm	0.11 mm	0.06 mm
Mean value		0.13 mm	0.07 mm	0.03 mm	0.02 mm
α		0.93	0.64	1.79	1.22
β		9.12	7.58	1.46	1.50
Probability of occurrence	1%	0.40 mm	0.27 mm	0.09 mm	0.08 mm
	5%	0.26 mm	0.17 mm	0.06 mm	0.05 mm

3. Fatigue Test of Aged Rails

From among the midsections of general tracks excluding welded joints and locations, where there is a large gap between sleepers and track bed—such as those with a gap magnitude of 7 mm reported in the previous chapter—, as well as welded rail joints where with significant rail-head surface irregularities, of about 0.5 mm, aged rails that have reached their replacement cycle are collected. Four-point bending fatigue tests are conducted on these rails to determine the relationship between stress and the number of cycles to failure, and to derive the S-N curves.

Table 6 summarizes the service history and specifications of the aged rails collected from operating lines. The test specimens consisted of rails with unground welded joints that had a cumulative passing tonnage ranging from 600 to 920 million tons, and rails sampled from midsections where a large gap had occurred. At the unground welded joints, the rail-head surface irregularities reached

Table 6 Specification of collected aged rails

Track type	Conventional line	High speed line
Rail type	JIS 50 kgN JIS 60 kg	JIS 60 kg
Cumulative passing tonnage	600-920 MGT	800 MGT
Position test piece obtained	Rail welds (GPW) Rail welds (TW)	Rail welds (GPW) Midsection (incl. hanging sleepers)
Maximum gap magnitude	Maximum 1.90 mm	Maximum 7 mm
Maximum rail-head irregularities	Maximum 0.50 mm	Maximum 0.03 mm (ground)

up to 0.5 mm. In addition, in midsections where significant gaps occur, the bending stress at the bottom of the rail during train passage can, in some cases, exceed its allowable magnitude at welded joints even in the absence of rail-head surface irregularities; therefore, aged rails laid in midsections with gap magnitudes of up to approximately 7 mm are also collected. A total of forty-five specimens were tested: twenty gas pressure welds (GPW), five thermit welds (TW), ten midsection rails, and ten expansion joints. Because the cross-sections of the collected tongue rails from expansion joints were asymmetric, a part of the rail head was machined to enable loading prior to fatigue testing.

Figure 7 shows an overview of the rail bending fatigue testing machine, and Table 7 summarizes the test conditions. The specimens are installed in the machine and undergo four-point bending fatigue tests under cyclic loading in dry conditions. The rail is supported over a span of 1,300 mm. To prevent fretting fatigue—fatigue failure caused by minute relative slip between the specimen and the contact points of the testing machine—each contact point at the loading and support locations is designed as a rotatable roller.

In this study, the tests were performed under cyclic loading with a minimum stress of 30 MPa; however, because the actual track condition corresponds to pulsating stress starting from 0 MPa, the

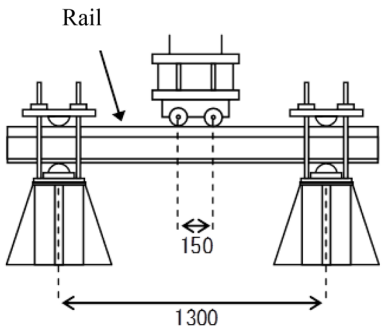


Fig. 7 Outline of bending fatigue test for collected rails (unit: mm)

Table 7 Test condition

Item	Test condition
Test type	Four-point bending (Repeated)
Minimum stress	30 MPa
Test environment	Normal temperature, in the atmosphere
Cycle frequency	2.0 to 5.0 Hz

stress range was converted to the full stress amplitude in accordance with previous studies. The conversion formula for the stress amplitude is given in (2). The values used here for the fracture strength of each type of welded joint were based on the failure strengths obtained from static bending tests and similar evaluations for each weld type, as shown in Table 8.

$$\sigma_A' = \sigma_T \cdot \frac{\sigma_A}{(\sigma_T - \sigma_{min})} \quad (2)$$

where σ'_A is the total stress amplitude after conversion, σ_T is the fracture strength of the test specimen, σ_A is the total stress amplitude during the test, and σ_{min} is the minimum stress during the test.

Figure 8 shows the test results; the arrows indicate specimens that did not fracture. Table 9 presents the intercepts of the S-N curves obtained from the test results. In the area where the full-stress amplitude exceeded 200 MPa, all specimens fractured. In this study, to evaluate the influence of parameters by focusing on changes in the intercept while fixing the slope, the slope of the S-N curve was set to a constant value of −136 for all conditions—in accordance with previous research [2]—and the intercepts were determined using the least squares method.

Table 8 Fatigue strength by welding type

Position	Fracture strength (unit: MPa)
Midsection	1,500
GPW	861
TW	683

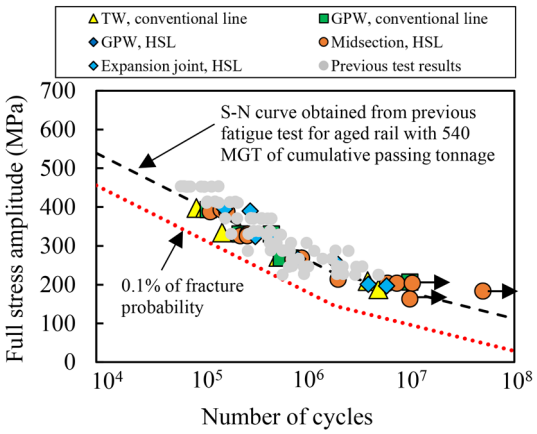


Fig. 8 Bending fatigue test results of collected rails

Table 9 Intercepts of S-N curves obtained fatigue test results

Considered test results	Intercept
Whole test results	1,085
Less than 750 MGT of cumulative passing tonnage (650 MGT as average)	1,075
Over 750 MGT of cumulative passing tonnage (800 MGT as average)	1,098
Previously obtained S-N curve of 50% probability of failure (540 MGT as average)	1,083
Previously obtained S-N curve of 0.1% probability of failure (540 MGT as average)	993

The S-N curves obtained from the present test results, when focusing on the intercept values, are almost identical to those for aged rails with 540 MGT reported in previous studies. It can also be confirmed that these values exceed the intercept corresponding to a failure probability of 0.1%, which is a lower bound previously proposed, given the variability in the fatigue strength of aged rails in conventional lines. Furthermore, the test results were classified into two groups—less than 750 million tons (average cumulative passing tonnage: 650 million tons) and 750 million tons or more (average cumulative traffic passing: 800 million tons). The intercepts were evaluated in the same manner, and the values were similar.

These results indicate that rails with large cumulative passing tonnages still possess sufficient fatigue strength for continued use even when in service under conditions where head-surface irregularities of up to about 0.5 mm or gaps up to about 7 mm are present. In addition, both fractured and unfractured specimens were observed in the range where the full stress amplitude is between 180 MPa and 200 MPa. However, at stress amplitudes below 180 MPa, all specimens remained unfractured. Three rail specimens, indicated by black arrows in Figure 8, did not fracture even after exceeding 10 million loading cycles. Among these specimens, one specimen subjected to a stress amplitude of 180 MPa was loaded up to 50 million cycles—significantly higher than the others—yet still did not fracture.

4. Fatigue Life Evaluation of Aged Rail

From the fatigue test results described in the previous chapter, the potential to extend the periodic replacement cycle was confirmed. Therefore, the fatigue life of rails was estimated using the life evaluation method based on the cumulative damage rule developed in previous studies [1, 2]. Figure 9 shows the flowchart for life prediction. This study focused on evaluating rail welded joints.

Because aged rails with an average cumulative passing tonnage

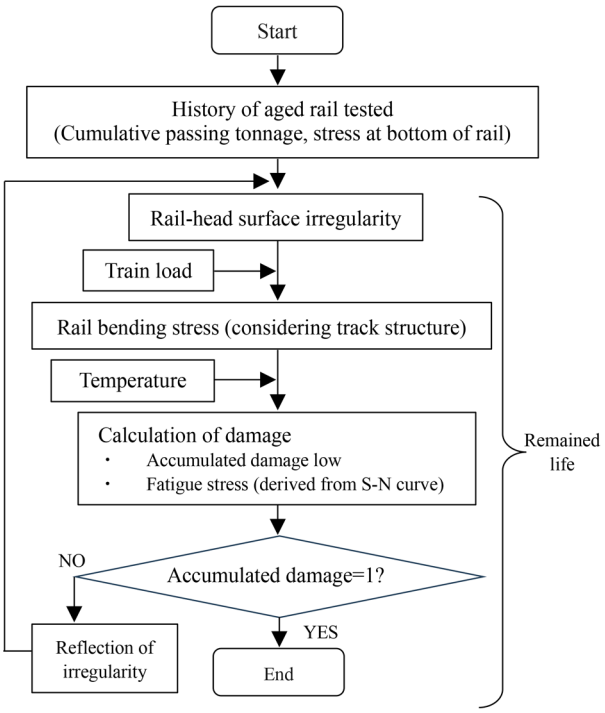


Fig. 9 Flowchart for fatigue life predictions

Table 10 Estimation equation of bending stress at bottom of rails

Rail type	Estimation equation (unit: MPa)
JIS 50kgN	30.8V+10.0W+6.8H+50.0
JIS 60kg	38.4V+5.0W+3.6H+41.9

V: 100mm chord rail-head surface irregularity, in mm;
W: 1m chord rail-head surface irregularity, in mm;
H: Magnitude of Gap between sleepers and ballast beds, in mm

Table 11 Fatigue life prediction result

Rail type	JIS 50kgN		JIS 60kg	
Existence of welds	Exist	Not exist	Exist	Not exist
Propagation of rail-head irregularity, in mm per 100 million tons	0.1	0	0.1	0
Gap, in mm	1.0	2.5	1.0	2.5
Estimated fatigue life, in 100 million tons	10.8	20.7	11.7	39.2

of 800 million tons—either unground or containing gaps—exhibited strength comparable to that represented by the previously obtained S-N curves, the remaining life was calculated using the loading history experienced up to the time of rail collection. This remaining life was then added to the prior loading history to estimate the overall fatigue life. The S-N curve shown in Figure 2, obtained from previous studies, was used as the fatigue strength characteristic curve.

Furthermore, as shown in Table 10, the bending stress at the bottom of the rail was estimated from the rail-head surface irregularities at the welded joints (V and W) and the gap magnitude between the sleepers and the ballast bed (H). The probability of failure was set as follows: 5% for the occurrence of gaps, 1% for the progression of rail-head surface irregularities, and 0.1% as the lower bound of the S-N curve. Previous investigations have indicated that the progression of rail-head surface irregularities at welded joints is at most approximately 0.10 mm per 100 million tons.

Table 11 presents the rail life estimation results. For welded joints, the progression of rail-head surface irregularities was set at 0.10 mm per 100 million tons, the occurrence probability of gaps was set at 5%, and the gap magnitude was set at 1.0 mm. For hanging sleeper locations in midsections, no progression of rail-head surface irregularities was assumed, while the occurrence probability of gaps was set at 5% and the gap magnitude at 2.5 mm.

Under the conditions for welded joints, the estimated fatigue life was approximately 1.08 billion tons for JIS 50 kgN rails and approximately 1.17 billion tons for JIS 60 kg rails. In addition, the estimated life of locations with a significant gap between sleepers and ballast beds in the midsections was longer than that of welded joints.

These results indicate that the fatigue life of rails exceeds the current replacement cycle of 600 to 800 million tons.

5. Extension of Replacement Interval and Determination of Stress Cut-Off Limit

As described in the previous chapter, it is estimated that the collected aged rails do not accumulate significant fatigue, provided that maintenance conditions similar to those during service are maintained. In addition, other fatigue tests conducted in the ultra-high-cycle regime, with loading cycles on the order of 10^8-10^9, have shown that fatigue failure does not occur when the applied stress is below the fatigue limit, suggesting that rails may be usable even after a cumulative passing tonnage of 1 billion tons [3]. This indicates that the probability of fatigue failure at the bottom of rails is low.

Furthermore, in the design of fatigue structures, it is generally accepted that when a structure is used under stress levels below a certain threshold, a stress cut-off limit can be defined, below which cumulative fatigue damage does not need to be considered. Depending on the judgment of each operator, if good track conditions are maintained, the bending stress at the rail bottom during train passage can be assumed to be small. In addition, if the rail surface is stable and has a rusted finish, it is unlikely that fatigue strength will deteriorate, and it may therefore be reasonable to forgo periodic replacement based on cumulative traffic tonnage.

Regarding the method for determining the cut-off limit of bending stress at the bottom of rails to eliminate the need for periodic replacement, it is considered appropriate to use the "soundness index" [9]—defined as the ratio of the fatigue limit of the rail to the bending stress at the bottom of rail, as proposed in previous studies—as an indicator. Maintenance management should then be conducted so as to maintain this soundness above a specified value.

The soundness of the rail is expressed by the following equation:

$$f_h = 1 - \frac{\sigma_m}{\sigma_c} \tag{3}$$

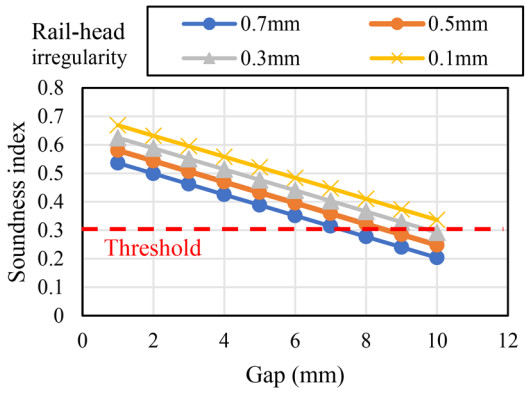
where, f_h denotes the soundness index, σ_m the bending stress at the rail base, and σ_c the fatigue limit. Assuming a fatigue limit of 200 MPa for the full stress amplitude, and considering on the safe side that a tensile axial force of 120 MPa acts in the rail due to temperature decreases in continuous welded rails, the correction formula for the fatigue limit obtained from the fatigue limit diagram [4] can be expressed as follows:

$$\sigma_c' = \sigma_c \left(1 - \frac{\sigma_{min}}{\sigma_T} \right) \tag{4}$$

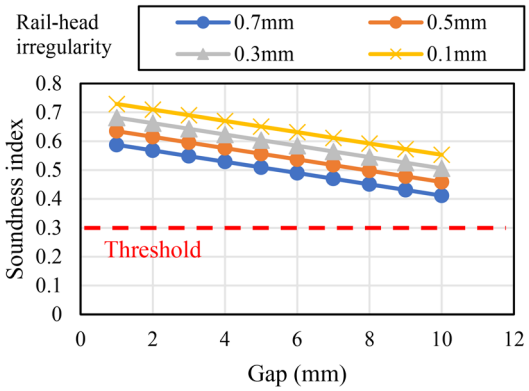
where, σ_c' denotes the corrected fatigue limit, σ_{min} the tensile axial stress (120 MPa), and σ_T the true fracture strength of the rail (1,500 MPa). From these values, σ_c' is calculated as 184 MPa, which is then substituted as the value of σ_c in (3) for determining the soundness index.

Next, the bending stress at the rail base, σ_m, is determined, for example, using the estimation equation shown in Table 10. This equation takes into account the magnitude of the rail-head surface irregularities and the gap. The resulting soundness index, calculated using this estimate, is shown in Figure 10. It can be observed that as head surface irregularity and the gap between sleepers and the ballast bed increase, the bending stress at the rail base increases, leading to a decrease in the soundness index.

For general metallic materials, the crack initiation limit is usually set 20% lower than the fatigue limit, where the fatigue limit represents the stress threshold for rail fracture, and the crack initiation limit represents the stress threshold for crack initiation. Therefore, it is considered that safe operation can be ensured if the soundness index is maintained at 0.2 or higher. In this study, values below a soundness index of 0.3 were examined to account for the margin of safety. Under conditions of a gap exceeding 8 mm and rail-head surface irregularity exceeding 0.5 mm for JIS 50 kgN rails, the



(a) JIS 50kgN rail



(b) JIS 60kg rail

Fig. 10 Relationship between soundness index, rail-head irregularity and gap

soundness index falls below 0.3.

In practical track maintenance management, when the soundness index falls to 0.3 or lower, measures become necessary, such as eliminating the gap between sleepers and the ballast bed through ballast tamping, correcting head surface irregularities through rail grinding, and ultimately rail replacement. Conversely, when the soundness index exceeds 0.3, the rail replacement interval can be extended up to the estimated fatigue life.

If rails continue to be used for extended periods, head surface irregularities may increase; therefore, appropriate maintenance measures, such as rail grinding or partial rail replacement, will be required at these locations. In addition, a significant gap between the sleepers and the ballast beds may occur. Because of the risk of rail fracture at such locations, it is desirable to use this rail management method.

Furthermore, JIS 60 kg rails are found to have a greater margin with respect to a soundness index of 0.3 than JIS 50 kgN rails. The fact that no reduction in strength was observed in fatigue tests of aged rails, even with a gap of around 7 mm, is attributed to their use under conditions where the soundness index remained above 0.3. However, since bending stress at the rail base increases with increasing wheel load, caution is required on lines where heavier trains operate.

It goes without saying that conventional rail replacement is the safest maintenance approach. In the future, decisions should take into account factors such as the condition of the rail head. In addition, verifying fatigue strength by testing rails with a certain service age and inspecting the corrosion condition at the bottom of the rail will continue to be effective.

6. Summary

This study focused on the gap between the sleeper and the ballast bed and on rail head surface irregularities. These factors are considered to affect the fatigue life of aged rails. A statistical analysis of their occurrence conditions on revenue lines was conducted. In addition, fatigue tests were carried out on aged rails under conditions with a significant gap and rail-head surface irregularities to obtain data on their fatigue characteristics. Furthermore, a quantitative evaluation method was proposed to ensure safety in rail maintenance management. This method accounts for the probabilities of gaps and rail-head surface irregularities, as well as the variability in the fatigue strength of aged rails. The main findings are summarized as follows.

(1) The occurrence conditions of gaps between sleepers and the ballast bed and of rail head surface irregularities were statistically evaluated using inspection data from revenue lines. These conditions influence the bending stress at the rail base, which is a key response parameter in estimating the fatigue life of aged rails. Probability density function analysis confirmed that the gamma distribution provides a good fit for both gap magnitude and head surface irregularity.

(2) Bending fatigue tests were conducted on aged rails that had reached the current replacement criterion of 800 million tons of cumulative passing tonnage. These rails exhibited significant gaps between sleepers and the ballast bed, as well as rail-head surface irregularities. The results showed that these rails have fatigue strength comparable to that of aged rails tested previously, with a cumulative passing tonnage of 540 million tons.

(3) From the fatigue test results in (2), it was found that the S-N curve for aged rails is independent of cumulative passing tonnage. Furthermore, extending the replacement interval is feasible provided that gaps between sleepers and the ballast bed, as well as head surface irregularities remain below certain levels.

(4) When the replacement interval is extended or eliminated, rails used in environments with significant gaps between sleepers and the ballast bed, and head surface irregularities, may experience increased bending stress at the bottom of the rails. This can potentially lead to rail fracture. To prevent such failures, a rail soundness index—derived from the bending stress at the rail base and the fatigue limit—was proposed as an indicator.

(5) It is believed that fatigue failure originating from the bottom of rails can be prevented by properly managing head surface irregularities and gaps between sleepers and the ballast bed so that the rail soundness index does not fall below 0.3. It should be noted that damage, such as corrosion and rail head defects, must be managed separately using methods such as ultrasonic inspection.

References

[1] Editorial Committee for the New Edition of Track Materials, New Edition of Track Materials. Tetsudo Genyo-sha, Japan, pp. 41-42, 2011 (in Japanese).

[2] Deshimaru, T., Kataoka, H., Abe, N., Ohno, M., "Life Prediction of Used Continuous Welded Rail," RTRI Report, Vol. 20, No. 4, pp. 5-10, 2006 (in Japanese).

[3] Mizutani, J., Hosoda, M., Kataoka, H., Yamamoto, R., "Study on the Extension for Replacement Period of Rail Taking into Account the Fatigue Life of Rail on the Region of the High Cycle Number," Journal of Railway Engineering, JSCE, Vol. 22, pp. 173-177, July 2019 (in Japanese).
[4] Ministry of Land, Infrastructure, Transport and Tourism, Design Standards and Commentary for Railway Structures - Track Structures, pp. 33-40, 2011 (in Japanese).
[5] Murakami, Y., Metal Fatigue: Effects of Small Defects and Nonmetallic Inclusions, Yokendo Ltd., 2004 (in Japanese).
[6] Japan Society of Steel Construction, Fatigue Design Recommendations for Steel Structures with Commentary, pp. 33-50, 2012 (in Japanese).
[7] Hosoda, M., Mizutani, J., Iwasaki, M., and Yamamoto, R., "Prediction Method of Fatigue Limit Based on Surface Roughness at the Rail Bottom of Aged Rails and Influence of Fatigue Test Methods," Transactions of the JSME, Vol. 86, No. 888, p. 20-00147, 2020 (in Japanese).
[8] Hosoda, M., Takahashi, N., Deshimaru, T., "Evaluation of Fatigue Life of Aged Rails Focusing on the Occurrence of Hanging Sleepers and Rail Irregularities Affecting Rail Stress," Transactions of the JSME, Vol. 91, No. 941, p. 24-00154, 2025 (in Japanese).
[9] Aizawa, H., Hosoda, M., Yamamoto, R., Tanaka, H., Mori, K., "Estimation of Causes of Rail Breakage at Tunnel Water Leakage Locations and Proposal of Management Methods," RTRI Report, Vol. 37, No. 8, pp. 11-12, 2023 (in Japanese).

Authors



Mitsuru HOSODA, Dr.Eng.
Senior Researcher, Rail Maintenance and Welding Laboratory, Track Technology Division
Research Areas: Railway Rails, Non-Destructive Inspection



Nobutaka TAKAHASHI
Researcher, Rail Maintenance and Welding Laboratory, Track Technology Division (Former)
Research Areas: Rail Welding, Non-Destructive Inspection, Track Maintenance



Tadashi DESHIMARU, Dr.Eng
Head of Rail Maintenance and Welding Laboratory, Track Technology Division
Research Areas: Railway Rails, Rail Fastening Systems, Track Components

Formalization of Field Engineers' Tacit Knowledge Using a Statistical Occurrence Prediction Model for Lock Adjustment of Point Machines

Wataru INABA
Data Analytics Laboratory, Information and Communication Technology Division (Former)

Kodai MATSUOKA
Data Analytics Laboratory, Information and Communication Technology Division

Shigeyuki TAMEHIRO
Shikoku Railway Company

Mitsuyoshi FUKUDA
Signalling and Operation Systems Technology Division

Kiyoyuki KAITO
Osaka University

This study aims to estimate a mixed Weibull hazard model from inspection records of lock adjustment, which are a major inspection item for point machines with a short inspection cycle. Furthermore, we conducted a survey to verify whether tacit knowledge regarding lock adjustment held by engineers has been formalized through the statistical model. The result clarified that the characteristic of lock adjustment, wherein one adjustment triggers the next, has been formalized as acceleration parameter estimates for an initial defect type. In addition, ten of the sixteen point machines with large heterogeneity parameter estimates, were identified as being subject to frequent adjustments. This confirms that some of the explicit knowledge is consistent with the tacit knowledge of engineers.

Key words**:** *mixed Weibull hazard model, points, lock adjustment, tacit knowledge*

1. Introduction

Amid the expected rapid decline in the working-age population, transferring inspection technologies has become a crucial issue for sustainable railway operation. This issue must be addressed alongside labor-saving measures and efficiency improvements in maintenance management, all while ensuring safety. Previous survey results indicate that, in the maintenance management of railway electrical systems, large amounts of labor are required for the periodic inspection of point machines and rail bonds (track circuits). The principal inspection judgments for point machines often needs tacit knowledge of experimental engineers [1]. However, methods for formalizing the experimental knowledge of engineers, hereinafter referred to as tacit field knowledge, have not yet been sufficiently discussed, and the verification of tacit knowledge related to point machine inspection has not progressed.

In recent years, advances in asset management methods in the field of civil engineering have enabled the sophistication of management processes by formalizing tacit knowledge [2]. In particular, studies have developed statistical deterioration prediction models that estimate the probability of defect occurrence and damage levels in civil engineering structures, such as tunnels and pavements, using inspection records [3]. Developing and applying similar statistical models to railway electrical systems could formalize engineers' tacit knowledge contained in inspection records. This in turn could enable quantitative discussion on optimizing inspection intervals. Additionally, if formal knowledge derived from statistical models is consistent with tacit field knowledge, it can serve as a basis for facilitating transmission of technical knowhow, thereby contributing to maintaining and improving technical capabilities.

Based on the background described above, this study focuses on inspection records of point-machine lock adjustment. A mixed Weibull hazard model, a type of statistical defect occurrence model, is applied to estimate the adjustment occurrence period (defined as the number of days until lock adjustment is required), along with its influencing factors and the variability among individual point machines. Furthermore, this study clarifies that the characteristics of lock adjustment occurrence obtained from the mixed Weibull hazard model, hereinafter referred to as adjustment occurrence characteristics, which constitute formalized tacit field knowledge. Specifically, the consistency between tacit field knowledge and two parameters of the model was verified through engineer questionnaires and on-site interviews [4].

2. Statistical predictive model for lock adjustment

2.1 Lock adjustment of point machines

Figure 1 shows the structure of the locking mechanism in a point machine. As shown in Fig. 2, when a point machine is locked, a total clearance of 3 mm remains to ensure smooth insertion of lock piece into the lock rod. This clearance varies due to factors such as temperature changes and train-induced vibration. An excessive imbalance between the left and right clearances hinders lock piece insertion, causing a switching failure. Therefore, field engineers periodically check point machines and perform adjustments as necessary.

When a lock adjustment is performed, it is recorded as an "adjustment" in the inspection log. Otherwise, when no adjustment is performed, this is recorded as a "normal operation". From inspec-

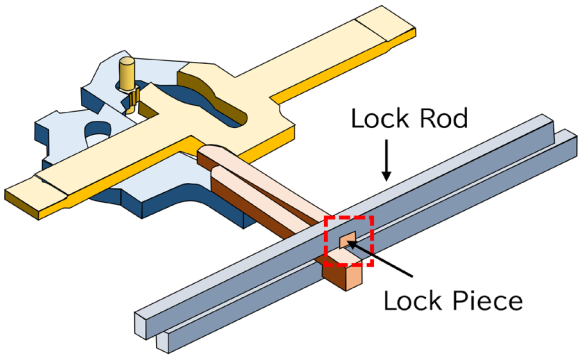


Fig. 1 Structure of locking mechanism in point machine

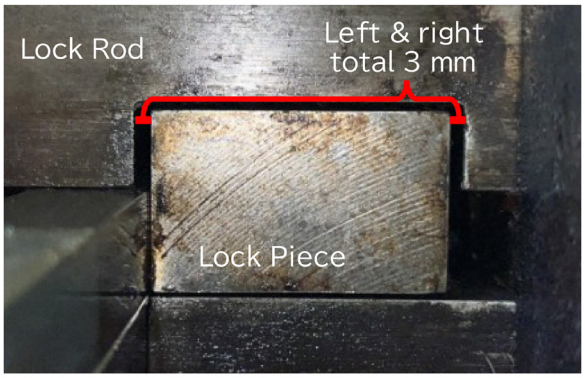


Fig. 2 Enlarged view of lock piece and lock rod (red box in Fig. 1)

tion records consisting of such discrete evaluation values, statistical properties related to lock adjustment are derived from inspection records consisting of such discrete data. In this study, we focus on the mixed Weibull hazard model [5], which is applicable to binary evaluation data to estimate the adjustment intervals. The point machines investigated in this study are limited to motor-driven electric point machines.

2.2 Mixed Weibull hazard model

In this paper, the survival probability is defined as the probability that a point machine is in a normal state, where lock adjustment is not required at a given time. The adjustment occurrence probability is defined as the probability that a point machine requires adjustment within a given period. Considering a discrete time axis $[0,\infty)$ where the start time of use of point machine k(k = 1, . . ., K) is denoted by τ_0, and the unit time is one day, the survival function at time t_k (t_k days after the start of operation) is expressed by Equation (1) in the mixed Weibull hazard model.

$$S(t_k)=\exp(-\epsilon_k \gamma_k t_k^m) \quad (1)$$

Here, m is the acceleration parameter, γ_k is the scale parameter of point machine k, and ε_k is the heterogeneity parameter of point machine k. The acceleration parameter is common to all point machines, whereas the scale parameter is defined by Equation (2) to reflect the individual characteristics of each point machine, available from equipment ledgers and similar records. As shown in Equation (3), the heterogeneity parameter is assumed to follow a gamma distribution with a mean of 1 and a variance of $1/\phi$, allowing for the consideration of individual differences among equipment that cannot be explained solely by their characteristics.

$$\gamma_k=\exp(x_k\beta') \quad (2)$$

$$p(\varepsilon_k)\sim g(\varepsilon_k|\phi) = \frac{\phi^\phi}{\Gamma(\phi)}(\varepsilon_k)^{\phi-1}\exp(-\phi\varepsilon_k) \quad (3)$$

Here, $x_k=(x_k^1,\ldots,x_k^J)$ is a row vector of the characteristic variables for point machine k, and $\beta=(\beta^1,\ldots,\beta^J)$ is a row vector of the corresponding regression parameters β^j (j = 1, . . ., J). The above formulation represents a mixed distribution model in which the gamma distribution of the heterogeneity parameter is incorporated into the Weibull distribution.

The occurrence of defects or adjustments in field equipment such as point machines is considered to be affected not only by explicitly identifiable characteristics, but also by factors that are difficult to quantify, such as the installation environment. Even in such cases, assigning a heterogeneity parameter to each point machine makes it possible to account for the actual individuality of each machine.

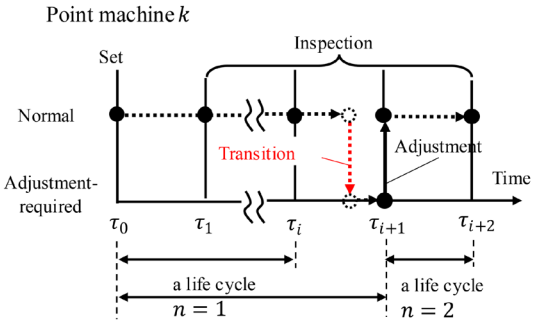


Fig. 3 Schematic diagram of a point machine for periodic inspections

2.3 Parameter estimation

This section describes the dataset and likelihood function used to estimate the mixed Weibull hazard model. Figure 3 is a schematic diagram of a point machine which can be monitored through periodic inspections. Periodic inspections only check status of the point machine at the time of inspection, i.e., whether adjustments are required, is observed. In the figure, black dots indicate the observable states of the point machine, whereas white dots indicate unobserved states. It is supposed that the point machine starts operation at time τ_0, and that the state requiring adjustment is first observed at the (i+1)-th inspection time τ_{i+1}. At this point, although the exact time of the state transition from normal to requiring adjustment is unknown, it is known that the transition occurred within the interval $(\tau_i, \tau_{i+1}]$. In other words, two types of information are obtained simultaneously: the information that point machine k remained in a normal state during the period $\bar{T}_{1k,1}=\tau_i-\tau_0$ and the information that it entered the adjustment-required state during the period $\bar{T}_{2k,1}=\tau_{i+1}-\tau_0$. Since adjustment is performed on the same day that adjustment-required state is observed and the machine is returned to a normal state, time τ_{i+1} becomes the new starting point of operation. When the inspection is completed at time τ_{i+2}, the second state transition is not observed, and it is possible to use only the data indicating that the point machine remained normal during the period $\bar{T}_{2k,2}=\tau_{i+2}-\tau_{i+1}$.

A dataset spanning from the start of operation of a point machine to the time of adjustment is treated as a single unit and is referred to as a life cycle. If the occurrence of adjustment in a point machine follows the mixed Weibull hazard model, the likelihood of a certain life cycle occurring is formulated by taking the difference of Equation (1). If the adjustment-required state is not observed by the final inspection time, the life cycle is treated as a partially ob-

served life cycle, and the likelihood is formulated solely on the basis of the fact that the point machine remained normal during that period. Based on the above, the likelihood for an arbitrary life cycle of a point machine is expressed by Equation (4).

$$\ell_{k,n}(\bar{d}_{k,n}, \bar{T}_{1k,n}, \bar{T}_{2k,n}, \bar{x}_k | m, \beta, \varepsilon_k) = \{S(\bar{T}_{1k,n}) - S(\bar{T}_{2k,n})\}^{\bar{d}_{k,n}} \cdot S(\bar{T}_{1k,n})^{1-\bar{d}_{k,n}} \quad (4)$$

Here, $\bar{T}_{1k,n}$ is the period from the start of operation to the time when the point machine was last observed as normal in the *n*-th life cycle of point machine *k*. $\bar{T}_{2k,n}$ is the period from introduction into service to the time at which the point machine was observed to be in the adjustment-required state in the *n*-th life cycle of point machine *k*. $\bar{d}_{k,n}$ is a dummy variable indicating whether the adjustment-required state was observed in the *n*-th life cycle of point machine *k*, and $\bar{x}_k$ denotes the observed values. Assuming that all life cycles across all point machines are mutually independent, the likelihood function for all target point machines is given by Equation (5).

$$\mathcal{L}(m, \beta, \varepsilon | \bar{d}, \bar{T}_1, \bar{T}_2, \bar{x}) = \prod_{k=1}^{K} \prod_{n=1}^{N_k} \ell_{k,n}(\bar{d}_{k,n}, \bar{T}_{1k,n}, \bar{T}_{2k,n}, \bar{x}_k | m, \beta, \varepsilon_k) \quad (5)$$

Here, K is the total number of point machines N_k is the number of life cycles of point machine k, $\bar{d} = (\bar{d}_{1,1}, \ldots, \bar{d}_{K,N_K})$, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_K)$, $\bar{T}_1 = (\bar{T}_{1,1,1}, \ldots, \bar{T}_{1,K,N_K})$, $\bar{T}_2 = (\bar{T}_{2,1,1}, \ldots, \bar{T}_{2,K,N_K})$, and $\bar{x} = (x_1, \ldots, x_K)$.

As shown in Equation (5), ε_k, is a random variable that follows the gamma distribution $g(\varepsilon_k | \phi)$. Therefore, the conditional probability of occurrence with respect to the gamma distribution $g(\varepsilon_k | \phi)$ the likelihood l', which corresponds to $p(\varepsilon_k | \phi)$ shown in Equation (3). In this case, the likelihood function with respect to ϕ is expressed as follows.

$$\mathcal{L}'(\phi | \varepsilon) = \prod_{k=1}^{K} \frac{\phi^{\phi}}{\Gamma(\phi)} (\varepsilon_k)^{\phi-1} \exp(-\phi \varepsilon_k) \quad (6)$$

Based on the above, the mixed Weibull hazard model can be estimated by finding the parameters (m, β, ε) that maximize the likelihood given the observed data $(\bar{d}, \bar{T}_1, \bar{T}_2)$. However, since the heterogeneity parameter ε is subject to the constraint given by Equation (6), the parameter that maximizes L' must be estimated simultaneously. Taking into account this hierarchical structure of the parameters, prior distributions were assigned to all parameters (m, β, ε, ϕ), and a hierarchical Bayesian method [6] was applied. As a numerical approximation method, the Markov chain Monte Carlo (MCMC) method [7] was used to estimate the posterior distributions of (m, β, ε, ϕ), and the estimates were evaluated using statistics of the posterior distributions. The number of iterations in the MCMC method was set to 20,000, with the first 5,000 discarded as burn-in. The remaining 15,000 samples were used to construct the posterior distribution.

3. Example of model application

3.1 Data overview

In this study, approximately 15 years of inspection records for 280 point machines managed by Shikoku Railway Company (JR Shikoku) were used to estimate the parameters of the mixed Weibull hazard model. The inspection records categorize conditions as either normal (indicated by '○') or requiring adjustment (indicated by '△'). These records include not only periodic maintenance inspections, conducted approximately every 90 days to check the locking condition, but also special inspections performed when lock deviation is detected by remote monitoring systems. The target point machines were mainly NS-type, accounting for more than 90% of the total. Although the inspections covered normal (N) and reverse (R) side locking conditions, due to space constraints, this paper only presents results of analyses for the normal side.

Lifecycle data were constructed from the inspection records and incorporated with features representing the operating conditions and structural characteristics of the point machines obtained from equipment ledgers. Specifically, the installation data for each point machine was set as its introduction into service, and life cycles were tracked using inspection results from that the starting point to the final inspection. When an adjustment was made, it was treated as a new starting date of operation; therefore, multiple life cycles were recorded for the same point machine. The characteristic values used were the type of switch operated by the point machine, specifically, elastic switches for high-speed operation or other ordinary switches, and the number of daily operations. In total, the dataset created using this method contained 3,025 cycles.

3.2 Estimation results of model parameters

Table 1 presents the estimated model parameters using the hierarchical Bayesian method. In MCMC, posterior distributions of parameters are estimated by repeated sampling; therefore, the statistical significance of the estimation results can be tested by hypothesis testing using posterior distributions. Specifically, if the 5th and 95th percentiles of the posterior distribution for a parameter share the same sign, the 90% credible interval does not include zero, indicating that the parameter is statistically significant.

As a result, the type of switch β^2 was non-significant, while the constant term β^1 and the number of operations β^3 were significant. Since β^3 has a positive value, the adjustment occurrence period becomes shorter as the number of operations increases. This is consistent with general physical mechanisms.

The failure characteristics of the target equipment can be classified based on whether the acceleration parameter is greater than, equal to, or less than 1. When m is greater than 1, the equipment is classified as a wear-out failure type, in which failures are more like-

Table 1 Estimation results of each parameter

Parameter	Description	Expected value	5th percentile	95th percentile	Significance	Adopted
$\hat{m}$	Acceleration parameter	0.814	0.783	0.843	—	○
$\hat{\beta}^1$	Constant term	−4.771	−5.060	−4.490	○	○
$\hat{\beta}^2$	Switch type	−0.001	−0.128	0.124	×	×
$\hat{\beta}^3$	Number of operations	0.554	0.349	0.793	○	○
$\hat{\phi}$	Variance of heterogeneity	7.875	6.165	9.781	—	○

ly to occur as the usage period becomes longer. When m is equal to 1, it is classified as a random failure type, in which the failure rate does not change over time. When m is less than 1, it is classified as an initial failure type, in which failures are most likely to occur immediately after the start of use. Based on the estimation results, the lock adjustment of point machines is classified as an initial failure type. This knowledge is formalized through statistical analysis of inspection records, and its consistency with tacit field knowledge is verified in Chapter 4.

Figure 4 shows the distribution of the estimated heterogeneity parameters. While the estimated values for many point machines are distributed near 1.0, those for some point machines are relatively large, exceeding 1.5. In particular, the 16 point machines whose 5th percentiles of the heterogeneity parameter exceeds 1.0, are considered to undergo frequent adjustments due to factors other than the number of operations, indicating significantly large heterogeneity. This knowledge is also derived from the statistical model of inspection records, and its consistency with tacit field knowledge is verified in Chapter 4.

3.3 Estimation results of the survival function

Figure 5(a) shows the survival functions of point machines when the number of operations for the minimum (1 operation/day), average (43 operations/day) and maximum (116 operations/day) values of the number of operations, which was identified as a significant characteristic variable. The heterogeneity parameter was set to 1.0. The periods until the survival probability decreased to 0.5 were 222, 136, and 112 days for the minimum, average, and maximum number of operations, respectively, all of which exceed 90 days.

Figure 5(b) shows the survival functions for 280 point machines, calculated using the estimated parameters of the mixed Weibull hazard model. The thin blue lines in the figure represent the survival functions for individual point machines, while the thick red line represents that of the average point machine, with the number of operations and the heterogeneity parameter set to the average value and 1.0, respectively. Although the variability among point machines is expressed by the heterogeneity parameter, this variability is markedly larger than that obtained by only the characteristic variables. As a result, the minimum period until the survival probability decreases to 0.5 is 54 days, while the maximum period is 399 days.

At 90 days, the minimum survival probability among the point machines is 0.35, which corresponds to an adjustment occurrence probability of 0.65. In other words, the lower limit of the survival probability traditionally accepted in maintenance inspections to date is 0.35, with the corresponding upper limit of adjustment occurrence probability—or risk—being 0.65. If this management standard is accepted, the survival probability of 0.35 can be used as a reference value for one management standard. Furthermore, for point machines whose survival probability remains above 0.35 at 90 days under the current inspection interval, the interval can be extended while maintaining the management standard, as long as the survival probability at the time of inspection does not fall below 0.35.

3.4 Validation of model adequacy

To verify whether the variability in the survival functions of the point machines reflects differences in their adjustment frequency in the inspection records, the expected adjustment occurrence period was calculated using Equation (7).

$$E[T_k]=\int_0^\infty S_k(u)\,du=\frac{\Gamma\left(1+\frac{1}{m}\right)}{\varepsilon_k\{\exp(x_k\beta')\}^{1/m}} \tag{7}$$

Here, Γ is the gamma function. Substituting the estimated parameter and characteristic values into Equation (7) yields the expected value for each point machine. Since the true values of the adjustment occurrence periods are unknown, the number of days from the start of operation of the time of adjustment (hereinafter referred to as the adjustment period) was used as the true value. The mean absolute error (MAE) for this true value was then calculated

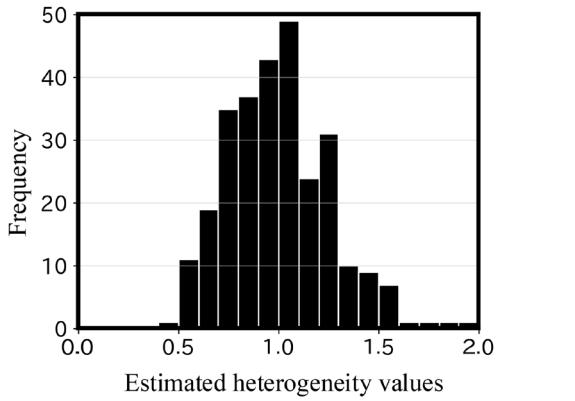
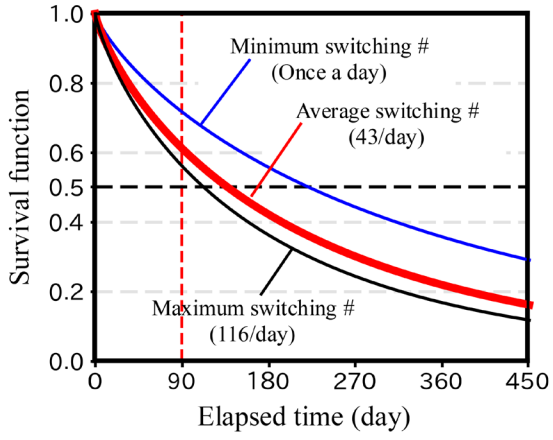
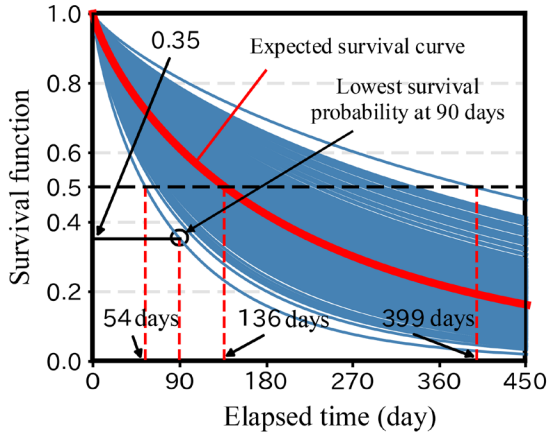


Fig. 4 Estimation results of heterogeneity parameters



(a) Effect of number of operations



(b) Effect of individual differences

Fig. 5 Survival functions for lock adjustment of point machines

using Equation (8) to evaluate the estimation accuracy of the model.

$$MAE = \frac{1}{K'} \sum_{k=1}^{K'} \left| E[T_k] - \frac{1}{N_k} \sum_{n=1}^{N_k} \bar{T}_{2k,n} \right| \quad (8)$$

where K' is the total number of target point machines, and $\bar{T}_{2k,n}$ is the adjustment period in the n-th life cycle of point machine k.

Figure 6 shows the comparison between the expected and observed adjustment occurrence periods for each point machine. For point machines with observed adjustment periods of less than 90 days, about half of the data points are concentrated within the range of the estimated values. The MAE for the point machines within this range was 37 days. This indicates that for these point machines, the prediction error is within approximately half of the 90-day inspection interval, demonstrating good accuracy. In the figure, the 16 point machines with significantly large estimated heterogeneity parameter are denoted by black squares. Their MAE was 22 days, indicating that more accurate prediction was achieved.

The above results demonstrate that the estimated mixed Weibull hazard model can predict observed values with a prediction error well within the inspection interval, particularly in the range where adjustment occurrences are concentrated for approximately half of the point machines. In particular, point machines with significantly large heterogeneity parameters tended to exhibit excellent good agreement between their predicted and observed values.

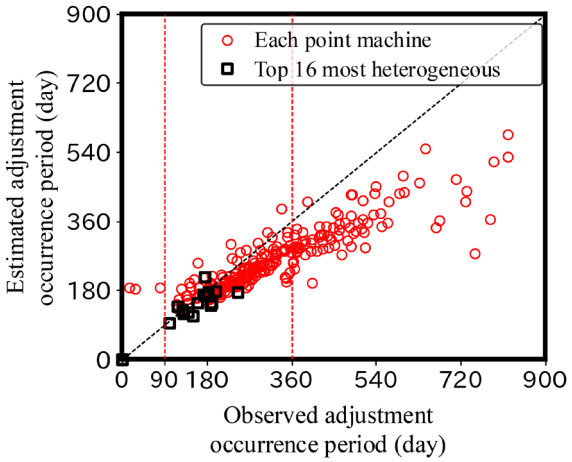


Fig. 6 Verification of prediction accuracy

Table 2 Experience of lock deviation immediately after lock adjustment, i.e., failure experience

Item	Content
Question 1	Have you ever experienced lock deviation immediately after lock adjustment?
Response	Yes / No
Percentage	Yes: 72% (73 respondents) / No: 28% (29 respondents)
Reasons given by respondents who answered "Yes"	Multiple answers allowed
Reason	Temperature changes, such as air temperature and season: 25% (18 respondents)
Reason	Dirt and grime: 41% (30 respondents)

4. Discussion on formalizing field tacit knowledge

4.1 Discussion on the initial failure type

The estimated value of the acceleration parameter was less than 1.0, indicating that lock adjustment tends to occur in the early stages. To comprehensively understand tacit field knowledge concerning point-machine inspections as a whole, including this tendency, a questionnaire survey, (the "engineer questionnaire") was conducted among 102 active and former field engineers at JR Shikoku. Based on the responses, comparisons were also made based on differences in the years of experience in both maintenance work and inspection work.

In this study, it is assumed that performing lock adjustment restores the point machine to its initial state; therefore, the period immediately after adjustment corresponds to the initial stage of use. Accordingly, the finding that lock adjustment tends to occur in the early stage can be interpreted as indicating that an adjustment tends to induce the next. To confirm whether this interpretation is consistent with field tacit knowledge, Question 1 investigated whether respondents had experienced lock deviation immediately after lock adjustment, that is, failure experience. In addition, Question 2 investigated factors other than the amount of deviation that are considered when determining whether lock adjustment is necessary, thereby examining the causes of failures.

Table 2 summarizes the responses to Question 1. More than 70% of the respondents (73 engineers) answered that they had experienced such a failure. As causes of the failures, 30 respondents cited insufficient consideration of dirt and grime, while 19 respondents cited insufficient consideration of temperature changes, such as air temperature and season. Therefore, a considerable number of field engineers tend to recognize that excessive adjustment induces a condition in which the subsequent adjustment is required. This tacit field knowledge can be considered formalized as an acceleration parameter less than 1.0.

Table 3 summarizes the responses to Question 2. Seventy-five engineers answered that additional factors to be considered; among these, temperature changes (such as air temperature and season) were cited by 53 respondents, and dirt and grime by 22. These responses reflect the common factors that were overlooked when lock adjustment failed.

Based on the above survey results, the appropriate clearance between the lock rod and the lock piece must be carefully determined during adjustment. This decision should account for not only the deviation amount but also environmental factors such as dirt, grime, and temperature changes. If this judgment is incorrect and excessive ad-

Table 3 Factors other than the amount of deviation use to determine need for lock adjustment

Item	Content
Question 2	Are there other factors other than the amount of deviation that you take into account to determine need for adjustment?
Response	Yes / No
Percentage	Yes: 74% (75 respondents) / No: 26% (27 respondents)
Reasons given by respondents who answered "Yes"	Multiple answers allowed
Reason	Temperature changes, such as air temperature and season: 71% (53 respondents)
Reason	Dirt and grime: 29% (22 respondents)

justment is performed, another intervention will be required again immediately afterward, consequently leading to highly frequent.

4.2 Analysis of high heterogeneity

There were 16 point machines for which the estimated heterogeneity parameter was significantly large, indicating that their high adjustment frequency was driven by factors other than the number of operations. To confirm whether this tendency is consistent with field tacit knowledge, interviews were conducted with members of field groups responsible for these point machines. Although the interviewees were from the electrical systems division, the survey results also incorporate findings from the track maintenance division. Table 4 shows the interview results.

Of the 16 point machines investigated, 10 were recognized as having high resistance or excessive tightness, resulting in a high frequency of lock adjustment. This indicates that tacit field knowledge was formalized to some extent through the heterogeneity parameter. The heterogeneity parameter captures influential factors that cannot be accounted for by characteristic variables alone. The characteristic variables were defined based solely on equipment ledger data from the electrical system. Consequently, point machines noted for high field resistance or tightness may have undergone frequent adjustment due to factors outside beyond the electrical division's management scope. In fact, several point machines required fewer adjustment after the track maintenance division investigated and addressed track conditions at the request of the electrical systems division.

As a result of the interviews, factors outside the management scope of the electrical system, namely, track-related factors such as deterioration of the condition of turnouts and rails, were confirmed in several point machines. Despite their high heterogeneity parameters, six point machines (T-8, T-11, T-15, T-16, T-3 and T-14) were not perceived by the field group in charge of these machines as having high lock adjustment frequency. Among these, T-8, T-11, and T-15 had fewer operations than the other machines. Therefore, even if their adjustment frequency was similar to that of other point machines, their standard hazard rates were smaller, and the heterogeneity parameters may consequently have become relatively large. In this case, because the adjustment frequency itself does not differ significantly from that of other point machines, it may not have been particularly remarkable for the field groups responsible for these point machines.

For the remaining point machines (T-3, T-14, and T-16), the original data confirms that adjustments were performed almost every time. However, the field groups responsible for monitoring did not perceive these machines as having a high lock adjustment frequency. In this study, the reason for this lack of field awareness regarding the three point machines, despite their frequent adjustments, could not be clarified and remains an issue for future study.

Furthermore, achieving sustainable efficiency improvement in maintenance management and elevating inspection quality is critical. This requires not only formalizing tacit field knowledge as described above, but also generating new tacit knowledge—or practical awareness—by analyzing and discussing the discrepancies between the two.

5. Conclusion

This paper describes the development of a mixed Weibull hazard model using inspection records to estimate the adjustment intervals. We focused on lock adjustment, a critical inspection item for point machines. Then, we verified the consistency between adjustment occurrence derived from a statistical model and tacit knowledge of field engineers through questionnaires and on-site interviews.

The estimation result of the mixed Weibull hazard model revealed that lock adjustment of point machines is classified as an initial failure type, with an acceleration parameter less than 1.0. Furthermore, unobserved heterogeneity among the point machines was found to have a significant influence that cannot be explained by characteristic variables. Based on these results, the characteristics of adjustment occurrences were interpreted and formalized as follows: frequent lock adjustments induce subsequent lock adjustment, and

Table 4 Field recognition of point machines with significantly large heterogeneity parameters

Point machine ID	Estimated heterogeneity value	Recognition of strong resistance/ tightness	Operations/day	Remarks
T-1	1.91	Yes	70	Frequent lock detection after adjustment
T-2	1.87	Yes	16	Frequent adjustment due to track-related factors
T-3	1.65	No	80	—
T-4	1.58	Yes	48	Frequent adjustment due to track-related factors
T-5	1.58	Yes	62	Locking failure occurred
T-6	1.56	Yes	26	Locking failure occurred
T-7	1.53	Yes	101	Frequent lock detection after adjustment
T-8	1.52	No	8	—
T-9	1.51	Yes	2	Switching movement is heavy
T-10	1.50	Yes	45	Frequent lock detection after adjustment
T-11	1.46	No	16	—
T-12	1.43	Yes	11	Frequent adjustment due to track-related factors
T-13	1.41	Yes	45	Frequent adjustment due to track-related factors
T-14	1.41	No	48	—
T-15	1.40	No	16	—
T-16	1.25	No	50	—

certain point machines experience recurring adjustments due to factors outside the scope of the electrical system management.

Results of the engineer questionnaire, however, confirmed that 74% of respondents had experienced readjustments due to lock deviation immediately after adjustment. Furthermore, interviews revealed that 10 of the 16 point machines with high heterogeneity were recognized by the relevant field groups as having strong resistance or tightness. As described above, the interpretations obtained from the statistical model were consistent to some extent with tacit field knowledge. Furthermore, the results demonstrated that parts of tacit field knowledge can be formalized using a statistical model estimated from inspection records.

Future areas requiring researcher are as follows. Firstly, it is necessary to establish adjustment rules that reflect the extracted tacit field knowledge. When formulating the adjustment rules, it is necessary to suppress excessive adjustment by taking into account lock characteristics that may induce subsequent adjustment. It is also necessary to deepen the analysis of the engineer questionnaire results to identify technical and psychological field engineer factors related to when excessive adjustment is performed. Based on these findings, we will establish rules to determine need for adjustment that minimize unnecessary adjustment, and verify their effectiveness.

Secondly, it is necessary to incorporate the effects of both rail and track conditions. This study revealed that several point machines with high heterogeneity may have been affected by track conditions. Since track conditions are measured by the track maintenance division, confirming the hypothesis that these conditions reflect machine heterogeneity could allow us to directly assess the resistance or tightness of point machines. This could make it possible to prevent unnecessary adjustments; specifically, the electrical systems division could request track maintenance from the relevant division before repeatedly performing adjustment of point-machines.

Acknowledgment

This study was conducted as part of a joint research project between Shikoku Railway Company (JR-Shikoku) and the Railway Technical Research Institute (RTRI) on the maintenance of facilities and equipment using digital technologies.

References

[1] Takahashi, K., Shimoki, K. et al., "A survey and perspective for improving maintenance efficiency of railway electrical facilities," *Journal of Railway Engineering*, Vol. 28, Issue 1, pp. 136-143, 2024 (in Japanese).

[2] Kobayashi, K. and Kaito, K., "Big data-based deterioration prediction models and infrastructure management: towards assetmetrics," In *Life-cycle of Structural Systems.* Routledge., London, U.K., pp. 84-93, 2018.

[3] Prakash, G., Yuan, X.X., Hazra, B. and Mizutani, D., "Toward a big data-based approach: A review on degradation models for prognosis of critical infrastructure," *Journal of Nondestructive Evaluation, Diagnostics and Prognostics of Engineering Systems*, Vol. 4, No. 2, ID. 021005, 2021.

[4] Inaba, W., Matsuoka, K., Tamehiro, S., Fukuda, M. and Kaito, K., "Consistency of adjustment cycle prediction model of railroad point machines estimated from inspection records with engineers' tacit knowledge," *Japanese Journal of JSCE*, Vol. 81, Issue 9, ID. 24-00301, 2025 (in Japanese).

[5] Kobayashi, K. and Kaito, K., "Random proportional Weibull hazard model for large-scale information systems," *Facilities*, Vol. 29, Issue 13-14, pp. 611-627, 2011.

[6] Mizutani, D., Matsuoka, K. and Kaito, K., "Statistical deterioration prediction model considering the heterogeneity in deterioration rates by hierarchical Bayesian estimation," *Structural Engineering International*, Vol. 23, No. 4, pp. 394-401, 2013.

[7] Yotsui, H., Matsuoka, K. and Kaito, K., "Dual sampling method for evaluating uncertainty when updating a Bayesian estimation model of a high-speed railway bridge," *Reliability Engineering & System Safety*, Vol. 259, ID. 110901, 2025.

Authors



Wataru INABA
Researcher, Data Analytics Laboratory, Information and Communication Technology Division (Former)
Research Areas: Statistical Degradation Prediction, Asset Management, Data Analysis



Kodai MATSUOKA, Ph.D
Senior Researcher, Data Analytics Laboratory, Information and Communication Technology Division
Research Areas: Structural Health Monitoring, Bayesian Statistics, Signal Processing



Shigeyuki TAMEHIRO
Technical Manager, Shikoku Railway Company
Research Areas: Point Machine, Inspection



Mitsuyoshi FUKUDA
Director, Head of Signalling and Operation Systems Technology Division
Research Areas: Signalling System, Safety Engineering



Kiyoyuki KAITO, Ph.D
Professor, Osaka University
Research Areas: Asset Management, Social Infrastructure

Survey of Load Conditions and Evaluation of Bending Load Capacity of Aged Shinkansen Prestressed Concrete Sleepers

Tsutomu WATANABE **Shintaro MINOURA**
Track Dynamics Laboratory, Railway Dynamics Division

Keiichi GOTO
Structural Mechanics Laboratory, Railway Dynamics Division

Field measurements and numerical analyses were performed to evaluate the actual load conditions of prestressed concrete (PC) sleepers used for high-speed trains. The results showed that the applied loads had a sufficient safety margin relative to the design bending crack limit. Furthermore, bending tests conducted in accordance with Japanese Industrial Standards indicated that the load capacity of PC sleepers generally decreases with age. Although sleepers over 50 years old still satisfied the standard bending fracture load, some specimens exhibited flexural cracking loads lower than the specified proof load.

***Key words*:** *prestressed concrete sleeper, track, maintenance, numerical analysis*

1. Introduction

Prestressed concrete sleepers (hereafter referred to as PC sleepers) are essential structural components for achieving higher railway speeds, improved ride comfort, and stable transportation. In Japan, PC sleepers were first installed on a trial basis in 1951. Since then, on Japanese National Railways and across the Japan Railway Group alone, excluding private railway companies, more than 40 million PC sleepers have been installed.

Although the design service life of PC sleepers is generally considered to be 50 years [1] an increasing number of them are being used beyond this period. Consequently, research on the evaluation of the remaining service life and maintenance management of PC sleepers has been actively promoted [2]. However, in Japan, most of these studies [3, 4] have focused on railways with a maximum operating speed of 160 km/h or less (hereafter referred to as conventional lines). In contrast, only a limited number of investigations and studies have addressed PC sleepers used for high-speed Shinkansen lines, and it is difficult to say that the latest findings have been sufficiently organized and consolidated. In addition, Shinkansen lines are inspected in a more detailed manner than conventional lines, as maintenance vehicles operate on them every day before revenue service begins. Nevertheless, evaluation of the structural soundness of PC sleepers is still primarily based on visual inspections of their surface conditions [5]. These evaluation methods are not necessarily suitable for assessing the load-bearing performance of the sleepers. Under these circumstances, recent studies have begun to focus on the relationship between cracks occurring on the upper surface of PC sleepers and the pull-out resistance of rail fastening system anchor bolts. These studies explore methods for automatically detecting crack lengths using deep learning and for evaluating the pull-out resistance of embedded anchors [6].

Furthermore, given the large number of installed PC sleepers, only a limited number can be replaced in a single nighttime maintenance period. Therefore, if a large number of existing aged PC sleepers were to rapidly and continuously fail to meet load-bearing performance during a certain period, replacement work could fail to keep pace, raising concerns about potential impacts on railway safety. One way to avoid such a situation is to quantitatively define the practical service life of PC sleepers and manage them in a planned manner.

Accordingly, this study aims to establish a new maintenance management framework for PC sleepers by focusing on the following points:

(1) Investigate actual load conditions by measuring the bending moments and wheel loads acting on PC sleepers installed on revenue lines.
(2) Examine, through numerical analysis, the effects of various parameters—namely, the Young's modulus of concrete, the spring constant of rail pads, the spring constant of the ballast bed, and the passenger load factor—on the bending moment of PC sleepers.
(3) Evaluate the bending load-carrying capacity of existing aged PC sleepers by conducting bending tests specified in the Japanese Industrial Standards (JIS).

2. Investigation of PC sleeper load environment

2.1 Investigation methods

2.1.1 Outline of a PC sleeper

Figure 1 shows the field measurement site. The site is a straight ballast track section on a reinforced concrete viaduct, with operating speeds ranging from approximately 130 to 300 km/h. Tracks use 60kg rails specified in JIS E1101, and PC sleepers are spaced 0.58 m apart.

Figure 2 illustrates the PC sleeper investigated in this study. The target sleeper is a PC sleeper specified in JIS E1201 (type: 3T). The sleeper has a length of 2,400 mm, and the heights of the cross-sections at the rail seat and at the center of the sleeper are 190 mm and 175 mm, respectively. The design compressive strength of the concrete is 49.1 N/mm², and prestressing steel tendons consisting of three deformed steel wires with a diameter of 2.9 mm each are used.

The design wheel load is defined as 160 kN, which is obtained by multiplying the static wheel load of 80 kN by a dynamic wheel load factor of two to account for dynamic effects due to train operation. The sleeper is designed to satisfy full prestressing under this load. The design decompression moment at the rail seat section is 10.95 kN·m, and the yielding moment of the prestressing steel is

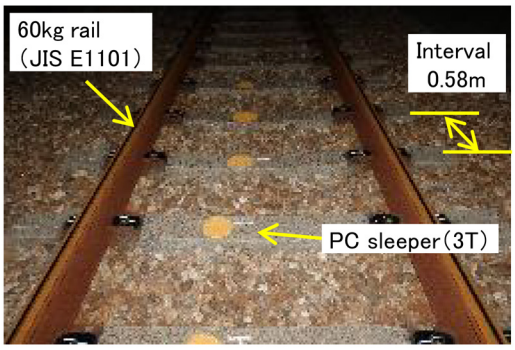


Fig. 1 Field measurement site

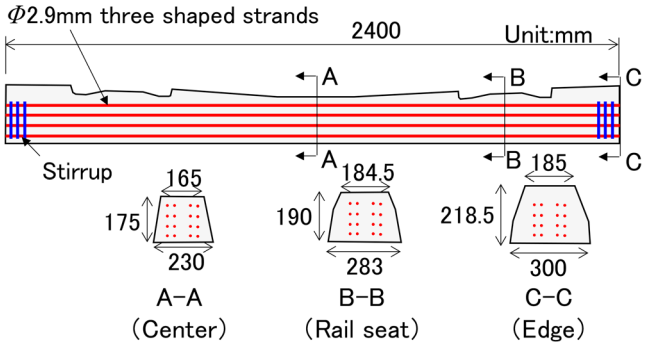


Fig. 2 Outline of PC sleeper (classified as 3T in JIS E 1201)

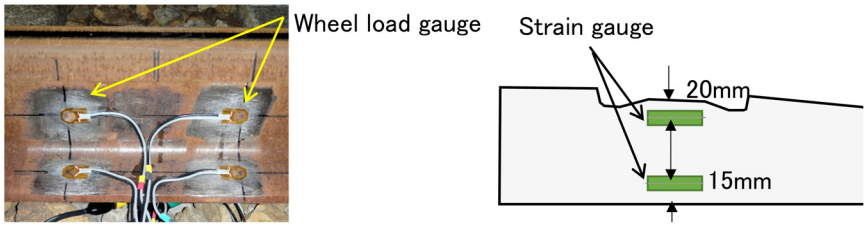


Fig. 3 Installation of wheel load gauges and concrete strain gauges

32.36 kN·m.

2.1.2 Methods for measuring wheel loads and bending moments of PC sleepers

Figure 3 shows the installation of the wheel load gauges and concrete strain gauges. The concrete strain gauges were attached to the rail seat section, approximately 20 mm inward from the upper edge and approximately 15 mm inward from the lower edge of the cross-section. The measured strains were converted into bending moments.

A total of seven consecutive PC sleepers were measured. All gauges were connected to data acquisition modules, and the data were recorded on a personal computer via a LAN cable. A data acquisition program was developed using LabVIEW, which enabled automatic recording triggered by the vibration acceleration of the PC sleepers. The sampling frequency was set to 10 kHz. Over a period of three days, a total of 204 train passages were measured.

2.2 Results of field measurements

Figure 4 shows the measurement results of the wheel loads and bending moments. For the bending moments, the average values of the five PC sleepers exhibiting the largest bending moments are highlighted in five different colors. The maximum and average wheel loads were 68.81 kN and 60.56 kN, respectively. The bending moment at the rail seat section of the PC sleepers reached a maximum of 3.81 kN·m and an average of 1.99 kN·m.

When compared with the design values, the maximum and average wheel loads correspond to 43% and 38% of the design wheel load of 160 kN, respectively. Similarly, the maximum and average bending moments correspond to 35% and 18% of the design decompression moment (DDM) of 10.95 kN·m. These results clearly indicate that the measured values are significantly lower than the design values. In addition, no clear dependence of wheel load or bending moment on train speed was observed.

3. Verification of bending moments of PC sleepers by numerical analysis

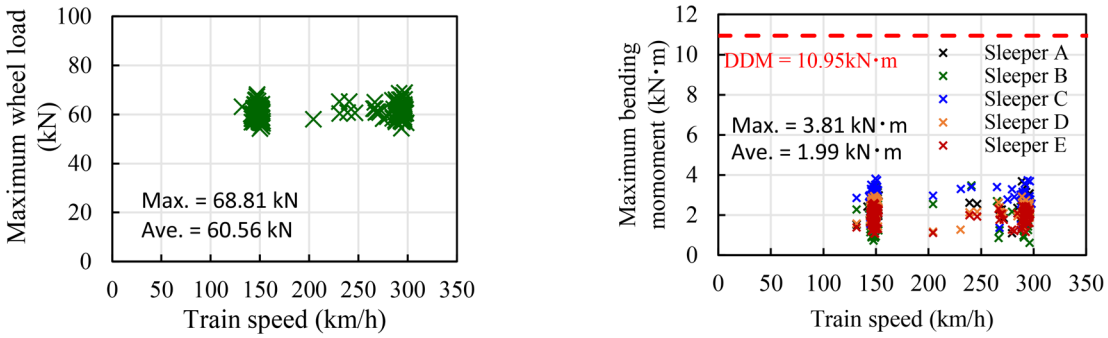
3.1 Numerical analysis method

Figure 5 presents an overview of the numerical analysis model. The PC sleepers analyzed in this section are identical to those investigated in Chapter 2. DIASTARS III was employed for the numerical analysis. DIASTARS III is a dynamic interaction analysis program for railway vehicles and track structures developed by the Railway Technical Research Institute [7].

The vehicle was modeled as a three-dimensional mechanical system in which the car body, bogies, and wheelsets were treated as rigid bodies interconnected by springs and dampers. Each vehicle model has 31 degrees of freedom. The vehicle specifications were set to match those of the trains passing through the measurement site, with a static wheel load of 56.6 kN. The passenger load factor, defined as the ratio of passenger load to the vehicle's nominal capacity, was set to 50%. In the analysis cases described later, changes in the passenger load factor were implemented by increasing or decreasing the mass of the car body in the vehicle model.

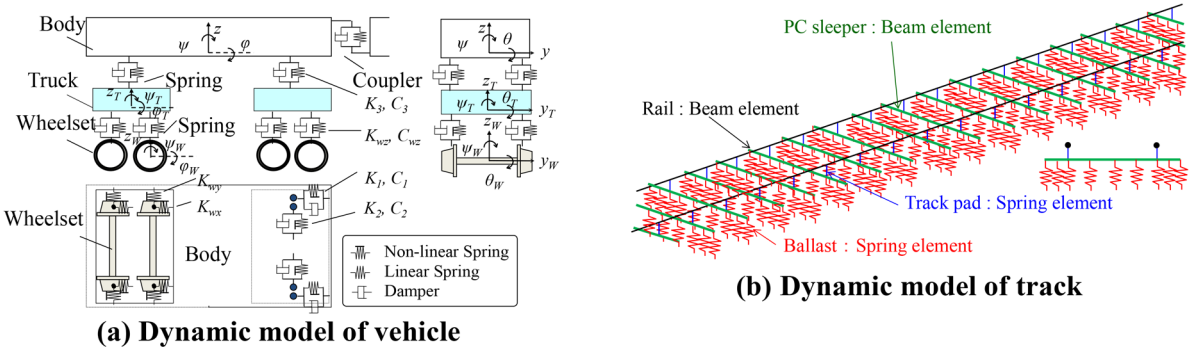
The track structure was modeled using the finite element method. Specifically, the rails and PC sleepers were modeled as beam elements, while the rail pads and ballast bed were modeled as spring elements. The rail was divided into four elements between adjacent rail fastening points connected by rail pads, and each PC sleeper was divided into eight elements along its longitudinal direction. As the measurement site is located on a viaduct, the concrete members of the viaduct were assumed to be sufficiently rigid, and the boundary condition at the bottom of the spring elements representing the ballast bed was fixed.

To reproduce the fluctuating forces generated at the wheel–rail interface, measured rail surface irregularities were applied to the rail head surface along the wheel running path, as illustrated in Fig. 5(c). The rail irregularities were measured using a rail head surface roughness measuring device with a measurement length of 1 m.

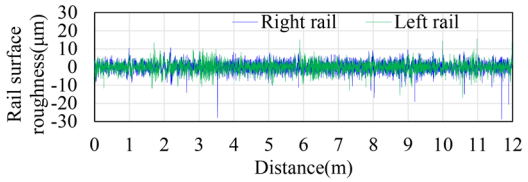


(a) Wheel loads (b) Bending moments

Fig. 4 Measurement results



(a) Dynamic model of vehicle (b) Dynamic model of track



(c) Rail surface roughness

Fig. 5 Overview of numerical analysis model

Table 1 Material properties used in the numerical analysis

	Material constants
Rail	60 kg rail (JIS E 1101; Young's modulus E_S = 200 GPa Track gauge = 1,435 mm
Rail pad	Spring constant D_p = 50 MN/m
Sleeper	Type 3T (JIS E 1201), PC steel strand = φ2.9 mm, N = 16 Length L_p = 2,400 mm Bottom width B_p = 283 mm (rail seat section), 230 mm (center section) Height H_p = 190 mm (rail seat section), 175 mm (center section) Young's modulus E_C = 33 GPa
Ballast	Ballast thickness h = 250 mm Support spring constant D_B = 200 MN/m (per rail)

Table 1 lists the material properties used in the numerical analysis. The material constants were assigned standard values specified in the Design Standards for Railway Structures and Commentary [8, 9].

Table 2 summarizes the analysis cases considered in this study. Parameters for which variations are expected in practice were selected for analysis. Specifically, the Young's modulus of PC sleepers concrete, the spring constant of rail pads, the spring constant of the ballast bed, and the passenger load factor of the vehicle were selected as parameters. The ranges of variation for these parameters were determined with reference to previous research findings. In addition, the rail pads were assumed to stiffen with aging [10].

It should be noted that, in order to ensure the realistically conceivable ranges of each parameter were fully encompassed, the up-

Table 2 Analysis cases

Parameter	Constant
Young's modulus of concrete E_C (kN/mm²)	15, 24, 33, 40, 48, 57
Spring constant of rail pad D_P (MN/m)	50, 100, 150, 250, 500
Spring constant of ballast bed D_B (MN/m)	50, 100, 200, 300, 500
Passenger load factor of the vehicle F_V (%)	0, 50, 100, 150

* The underlined constants constitute the basic case.

per and lower bounds of the parameters were intentionally set to values that may exceed the strictly anticipated ranges.

3.2 Results of verification of the effects of various parameters based on numerical analysis

Figure 6 compares the analytical and measured results for the bending moments at the rail seat section of the PC sleepers. Among the 204 measured data sets, those corresponding to train passages at a speed of 300 km/h were extracted such that the average peak bending moments of one train was approximately equal to the average value of about 2.0 kN·m shown in Fig. 4(b). As shown in Fig. 6, the analytical results are generally in good agreement with the measured results.

Figure 7 presents the results of the investigation into the effects of various parameters on the bending moment of PC sleepers. The red plots in the figure indicate the base case. As shown in Fig. 7(a), relative to the base-case Young's modulus of 33 kN/mm², the bending moment decreased by 12% when the Young's modulus was reduced to 15 kN/mm², and increased by 6% when it was increased to 57 kN/mm². As shown in Fig. 7(b), relative to the base-case rail pad spring constant of 50 MN/m, the bending moment increased by 50% when the rail pad spring constant was increased to 500 MN/m. Furthermore, as shown in Fig. 7(c), relative to the base-case ballast bed spring constant of 200 MN/m, the bending moment decreased by 4% when the spring constant was reduced to 50 MN/m and decreased by 9% when it was increased to 500 MN/m. Finally, as shown in Fig. 7(d), relative to the base-case passenger load factor of 50%, the bending moment decreased by 18% when the passenger load factor was reduced to 0% and increased by 11% when it was increased to 150%.

From these results, it was confirmed that the bending moment generated in PC sleepers varies depending on the effects of the individual parameters. Under the conditions considered in this analysis, even if all the conditions indicated by the parameters that tend to increase the bending moment relative to the base case were overlapped, the resulting bending moment would still be approximately 4.4 kN·m. Consequently, it is considered unlikely that bending moments exceeding the DDM of 10.95 kN·m will occur in PC sleepers.

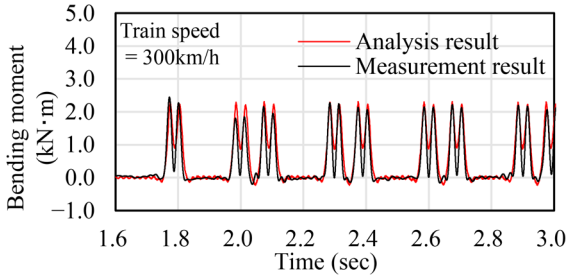


Fig. 6 Analytical and measured results for the bending moments at the rail seat section of the PC sleepers

4. Evaluation of bending load-carrying capacity of PC sleepers for Shinkansen lines

4.1 Test method

4.1.1 PC sleepers investigated

The PC sleepers investigated for the evaluation of bending load-carrying capacity included not only the type 3T PC sleepers shown in Fig. 2, but also type 3H PC sleepers specified in JIS E1201, as shown in Fig. 8, which are similar to type 3T.

The type 3H PC sleepers were designed with a design wheel load of 300 kN based on the results of speed-up tests of test vehicles conducted in 1972 [11]. Compared with type 3T sleepers, type 3H sleepers have a larger cross-sectional size and incorporate four additional prestressing steel tendons.

Table 3 lists the numbers of PC sleepers that were investigated, classified by service age. The service age of sleepers ranged from 10

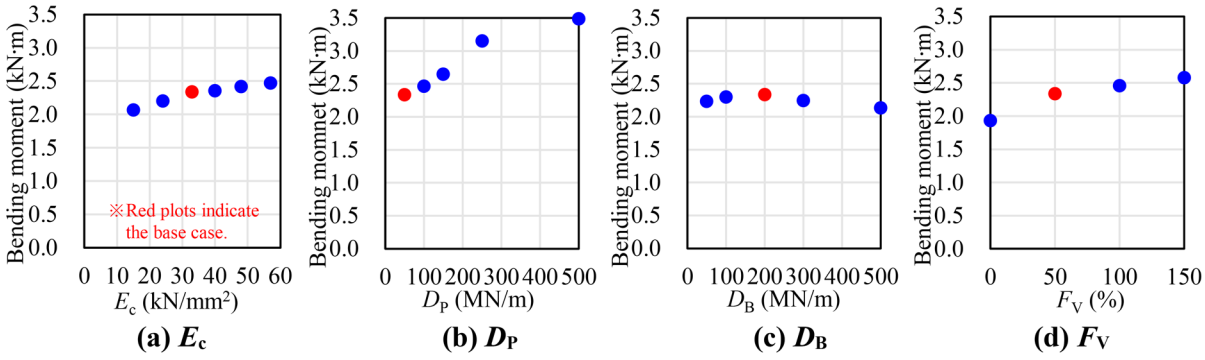


Fig. 7 Results of investigation into effects of various parameters on bending moment at rail seat section of PC sleepers

to 60 years. Although all of the sleepers were installed on the same railway line, they were located at different positions along the line. The number of type 3T PC sleepers was 31, while that of type 3H PC sleepers was 49.

4.1.2 Bending test specified in JIS E1201

Figure 9 illustrates an overview of the bending test specified in JIS E1201. This test involves a positive bending test at the rail seat section and a negative bending test at the center section of the sleeper. The loading span is 700 mm.

This test was originally conducted to verify the performance of PC sleepers immediately after they were manufactured. However, apart from this test, no other test method exists that can quantitatively evaluate the mechanical performance of PC sleepers and has been widely applied in practice. Therefore, in this study, this test method was adopted for the evaluation of PC sleepers.

Markings were made at the loading points and support positions on the PC sleepers to ensure consistent placement and minimize variability in the setup during testing. JIS E1201 specifies the bending proof load P_{cr} and the bending fracture load P_u as reference values, which are calculated using Eqs. (1) and (2), respectively.

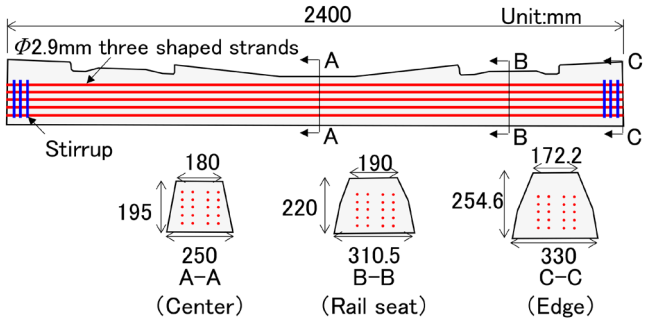


Fig. 8 Outline of PC sleeper (classified as 3H in JIS E 1201)

Table 3 Numbers of PC sleepers

Age	~10	~20	~30	~40	~50	~60	Total
3T	0	3	2	14	8	4	31
3H	0	6	6	18	19	0	49

The test verifies that no cracking occurs when the bending proof load P_{cr} is applied, and that the PC sleeper does not fracture when the bending fracture load P_u is applied.

$$P_{cr}=4\times(\sigma_{CPr}\times 0.9+f_{ta})\times Z/L \quad (1)$$

$$P_u=4\times 3\sigma_{Pe}\times Z/L \quad (2)$$

Here, σ_{CPr} denotes the concrete stress due to the introduced prestressing force, f_{ta} is the allowable tensile stress of concrete (3 N/mm²), and σ_{pe} is the concrete stress due to the effective prestressing force, with an effective prestress ratio of 65%. Z represents the section modulus, and L is the loading span, which is 700 mm.

For type 3T PC sleepers, the specific values of the bending proof load P_{cr} and bending fracture load P_u are 114 kN and 190 kN, respectively, for positive bending at the rail seat section, and 95 kN and 171 kN, respectively, for negative bending at the center section of the sleeper. For type 3H PC sleepers, the corresponding values for positive bending at the rail seat section are 190 kN and 300 kN, respectively, while those for negative bending at the center section are 120 kN and 203 kN. These values indicate that the bending capacities of type 3H PC sleepers are greater than those of type 3T PC sleepers.

The test results were organized in terms of service age and cumulative passed tonnage (hereafter referred to as cumulative tonnage) for both the cracking load and the ultimate load. This approach was adopted because, even for PC sleepers with the same service age, the number of trains passing over them differs depending on the line section; therefore, an evaluation based on both service age and cumulative tonnage was considered necessary. The cracking load was defined as the load at which cracking was visually identified, while the ultimate load was defined as the maximum load at which the PC sleeper could no longer sustain the applied load due to bending failure.

4.2 Test results

Figures 10 to 13 present the results of the bending tests. These figures also include the previously specified bending proof load and bending fracture load. As can be seen from the figures, both the cracking load and the ultimate load tend to decrease with increasing service age and cumulative tonnage. This tendency is consistent with that reported for PC sleepers used on conventional railway lines [2].

All of the type 3T PC sleepers tested in this study satisfied the bending proof load and bending fracture load specified in JIS. For

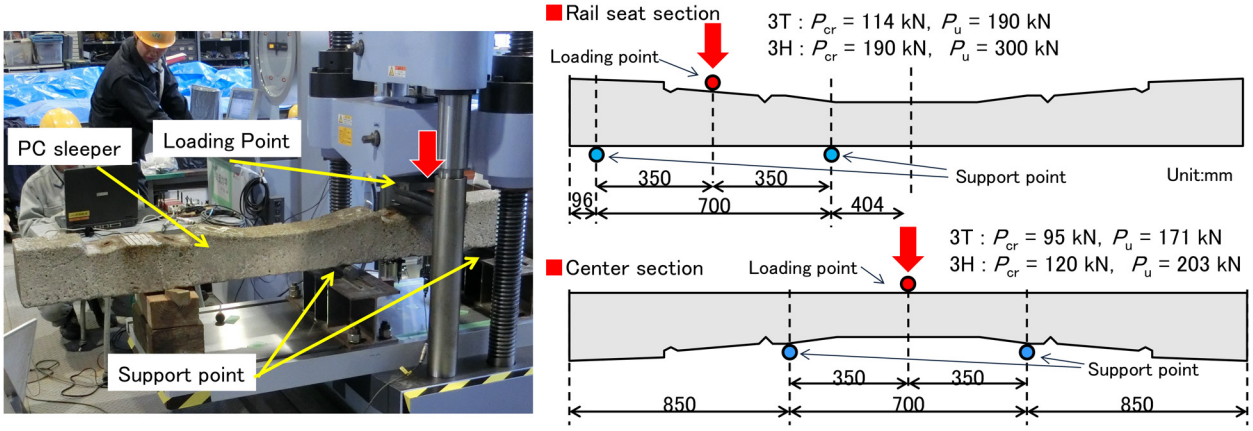


Fig. 9 Overview of the bending test specified in JIS E1201

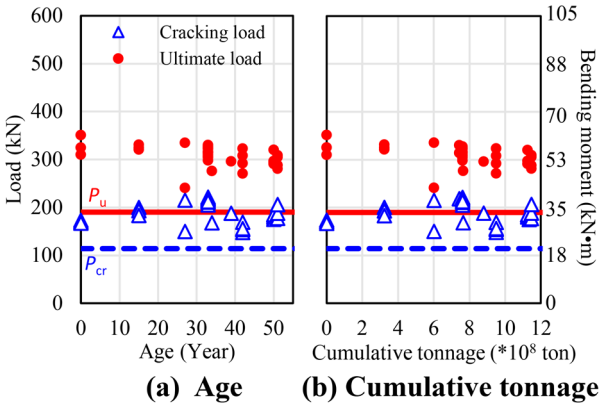


Fig. 10 Positive bending test (3T)

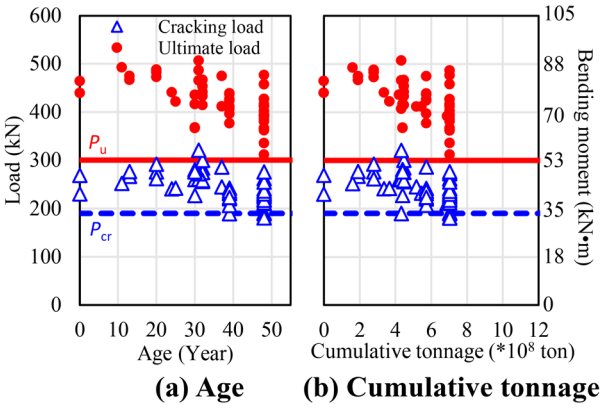


Fig. 11 Negative bending test (3H)

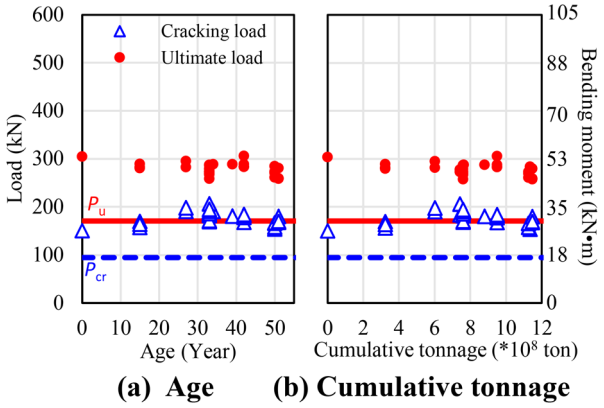


Fig. 12 Positive bending test (3T)

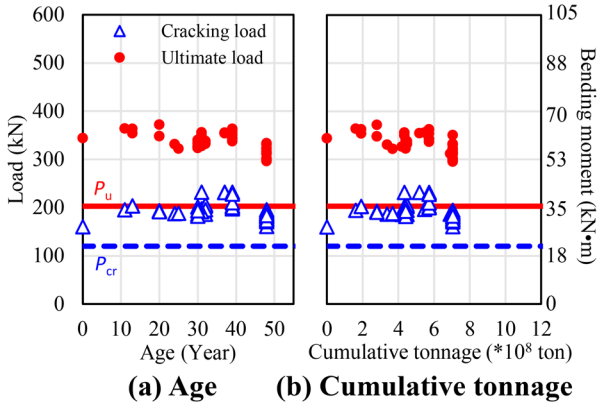


Fig. 13 Negative bending test (3H)

type 3H PC sleepers, although all specimens satisfied the JIS-specified bending fracture load with respect to the ultimate load, two out of the 49 sleepers (approximately 4%), both with a service age of 48 years, failed to satisfy the bending proof load in the positive bending test at the rail seat section.

In addition, the right-hand axes of these figures indicate the bending moments generated in the PC sleepers. The bending moments acting on PC sleepers during actual train passages, as shown in Fig. 6, are approximately one-tenth of those corresponding to the cracking load in the bending tests and approximately one-sixteenth of those corresponding to the fracture load.

From these results, it can be concluded that, at present, the PC sleepers satisfy the JIS requirements for bending fracture and are not in a condition that immediately affects railway safety. However, since the bending load-carrying capacity tends to decrease with increasing service age and cumulative tonnage, it is desirable to quantitatively define the practical service life and shift toward a planned maintenance management framework. This will help avoid a situation in which a large number of sleepers fail to meet load-bearing performance requirements simultaneously and continuously, as discussed earlier.

Furthermore, although this assessment is qualitative, type 3H PC sleepers exhibit greater scatter in ultimate load values, particularly in the positive bending tests and more degradation than type 3T PC sleepers. Differences in sleeper type or variations in installation environments may influence deterioration of bending load-carrying capacity; however, verification of these effects remains a subject for future study.

5. Conclusions

The findings obtained in this study can be summarized as follows:

(1) Field measurements of wheel loads and bending moments acting on PC sleepers revealed that, within the range investigated, the maximum and average wheel loads were 68.81 kN (43% of the design value) and 60.56 kN (38%), respectively. The maximum and average bending moments were 3.81 kN·m (35% of the design decompression moment) and 1.99 kN·m (18%), respectively. These results indicate that the measured values are significantly lower than the corresponding design values. In addition, no dependence on wheel load or bending moment on train speed was observed.

(2) Numerical analyses examining the effects of various parameters on the bending moment of PC sleepers showed that it is unlikely that bending moments exceeding the design decompression moment of 10.95 kN·m will occur in PC sleepers under the conditions considered in this study, even if all the conditions indicated by the parameters that increase the bending moment relative to the base case were to occur simultaneously.

(3) Bending tests conducted on existing aged PC sleepers in accordance with JIS demonstrated that both the cracking load and the ultimate load tend to decrease with increasing service age and cumulative tonnage. These results are consistent with previous studies on PC sleepers used on conventional lines. Although approximately 2% of the tested sleepers did not satisfy the JIS requirement for the cracking proof load in the positive bending test at the rail seat section, all of the sleepers satisfied the JIS

requirement for bending fracture, indicating that they are not in a condition that would immediately affect railway safety. Furthermore, the bending moments acting on PC sleepers during actual train passages were found to be approximately one-tenth of those corresponding to the cracking load and approximately one-sixteenth of those corresponding to the fracture load obtained in the bending tests.

(4) On the basis of these results, it is concluded that even PC sleepers with a service age of approximately 50 years currently satisfy the JIS requirements for bending fracture and are not in a condition that immediately affects safety. In future work, investigations and tests focusing on aged PC sleepers will be continued to monitor changes in load-carrying capacity associated with service age and cumulative tonnage. In addition, further studies will be conducted to clarify the mechanisms underlying the aging-related deterioration of PC sleepers.

Acknowledgment

The authors wish to express their sincere thanks and acknowledge the significant support received from Mr. Shuji Kano, Mr. Shuhei Kikkawa, and Mr. Takuya Inoue of the West Japan Railway Company, as well as all other related personnel, in conducting the field measurements.

References

[1] FIP Commission on Prefabrication, FIP State of the Art Report (Concrete Railway Sleepers). Thomas Telford Ltd.: London, UK, 1987.

[2] Watanabe T., Minoura S., Goto K., Ikeda M., "Actual Condition Survey of the Effects of Aging and Accumulated Tonnage on the Load-bearing Capacity of PC Sleepers," Quarterly Report of RTRI, Vol. 65, No. 1, pp. 55-60, 2024.

[3] Japan Railway Civil Engineering Association, Research Report on the Service Life of PC Railway Sleepers, 1986 (in Japanese).

[4] Maeda M., Funada T., "Strength evaluation of aged PC sleepers," 60th Annual Meeting of JSCE, 4-129, pp. 257-258, 2005 (in Japanese).

[5] Railway Technical Research Institute, Maintenance Standard for Railway Steel Structures and Commentary (Track Structures). Maruzen Co., Ltd.: Tokyo, Japan, 2007.

[6] Minoura S., Watanabe T., "Fundamental Study on Crack Detection Method for Prestressed Concrete Sleepers Using Deep Learning Model," Quarterly Report of RTRI, Vol. 67, No. 2, pp. 1-8, 2026.

[7] Wakui, H., Matsumoto, N., Tanabe, M., "A Study on Dynamic Interaction Analysis for Railway Vehicle and Structures," Quarterly Report of RTRI, Vol. 35, No. 2, pp. 96-104, 1994.

[8] Railway Technical Research Institute, Design Standards for Railway Structures and Commentary (Concrete Structures). Maruzen Co., Ltd.: Tokyo, Japan, 2023.

[9] Railway Technical Research Institute, Design Standards for Railway Structures and Commentary (Track Structures). Maruzen Co., Ltd.: Tokyo, Japan, 2012.

[10] Moriya T., Tanaka Y., Yokoyama H., Iwata N., "Influence of low-spring-constant track pad installation on ground vibrations," 60th Annual Meeting of JSCE, 4-111, pp. 221-222, 2005 (in Japanese).

[11] Sato Y., "Cause and Countermeasure of Extremely Large Wheel-Load Cased in 951-Type Test Electric Car Running on the Tokaido Shin Kansen," RTRI Report, No. 824, 1972 (in Japanese).

Authors



Tsutomu WATANABE, Dr.Eng.
Head of Track Dynamics Laboratory, Railway Dynamics Division
Research Areas: Concrete Sleeper, Numerical Simulation, Structure-Borne Noise, Vibration



Shintaro MINOURA, Dr.Eng.
Senior Researcher, Track Dynamics Laboratory, Railway Dynamics Division
Research Areas: Concrete Sleeper, Machine Learning



Keiichi GOTO, Dr.Eng.
Head of Structural Mechanics Laboratory, Railway Dynamics Division
Research Areas: Dynamic Interaction between Vehicle and Bridge, Post-derailment Vehicle Behavior

Estimation of Newly Fallen Snow Density Using Disdrometer Data

Kazuya TAKAMI
Meteorological Disaster Prevention Laboratory, Disaster Prevention Technology Division

This study proposes a method for estimating newly fallen snow density using disdrometer data. By analyzing the particle size and fall velocity distributions obtained from a disdrometer, the degree of riming is quantified and used to derive snow density. The proposed approach shows a stronger correlation (R = 0.78) and significantly lower RMSE (10.5 kg m^{-3}) compared with conventional temperature-based methods. This technique enables automated, high-frequency estimation of snow density, improving snowfall depth calculations and snow hazard assessments, especially under sub-zero conditions where traditional methods are less accurate.

Key words**:** *snowfall, accretion, newly fallen snow density, disdrometer*

1. Introduction

To ensure the safety and stability of winter railway operations railway companies engage in snow clearing work depending on snow depth and related conditions. Accurately identifying "where", "how much", and "what type" of snow has fallen along railway lines is therefore essential for making appropriate operational decisions. Some meteorological stations operated by the Japan Meteorological Agency provide snow-depth measurements. However, snow conditions on railway lines differ from those in surrounding areas. Snow accumulating between tracks is lifted up or compacted by passing trains, resulting in differences both in quantity and physical properties of the snow compared to that outside the railway line. Consequently, it is necessary to continuously evaluate the amount of snow that falls between successive train passages at arbitrary locations along the line.

In recent years, the widespread availability of high-resolution weather radar has enabled the use of spatial precipitation data. By multiplying radar-derived precipitation intensity by the ratio of newly fallen snow density to the known density of liquid water (i.e., snow-to-liquid ratio), it is possible to convert precipitation into snowfall depth. Moreover, newly fallen snow density itself is an important parameter for assessing snow-related hazards; for example, the amount of snow accreting beneath rolling stock increases as the density of the surface snow layer decreases [1].

Newly fallen snow density is determined by measuring the volume and weight of the snow that accumulates over a fixed time interval. To eliminate the influence of compaction from snow deposited prior to the observation period, a snow board is typically used, and only the snow that accumulates on the board during the target interval is measured [2, 3, 4, 5, 6]. This procedure requires manual operation, making high-frequency, continuous observations difficult. Consequently, even at sites staffed by observers, measurements are generally limited to daily snowfall, and only a few studies have conducted observations at intervals of approximately one hour or less [4, 6, 7].

Although automated instruments such as rain gauges and snow-depth sensors can, in principle, be used to estimate snow volume (from differences in snow depth) and weight (from precipitation measurements), snow-depth sensors cannot distinguish changes caused by new snowfall from those caused by compaction of pre-existing snow. Methods have been proposed to estimate and correct for compaction [7, 8, 9]; however, in heavy-snow regions where a persistent snowpack develops, accurately calculating the settlement of existing snow is difficult, and the associated errors increase with snowfall amount and snow depth [8]. Furthermore, because snowfall intensity is generally lower than rainfall intensity, the resolution of sensors strongly affects the accuracy of precipitation and snow-depth measurements. For example, the resolutions of snow-depth sensors (1 cm) and rain gauges (0.5 mm) at Japan Meteorological Agency sites are not enough to capture event-scale variability. In addition, the catch efficiency of rain gauges for solid precipitation is highly sensitive to wind speed, often resulting in substantial underestimation [10]. Snow-depth sensors are also far less widely installed than rain gauges, particularly in regions with infrequent snowfall such as the Tokyo metropolitan area. Given these limitations, continuous and reliable in situ measurements of newly fallen snow density are difficult to obtain. Therefore, a method is required to estimate snow density using data that can be automatically observed by other instruments.

In practical applications, newly fallen snow density has often been prescribed as a constant value (e.g., 100 kg m^{-3}) or estimated empirically from meteorological variables such as air temperature[4]. However, temperature primarily reflects differences associated with the presence or absence of melting and thus cannot represent variations in snow density arising from differences in precipitation particle types - such as aggregate and graupel - which dominate under subfreezing conditions. Figure 1 illustrates the relationship between newly fallen snow density and air temperature observed at the Shiozawa Snow Disaster Prevention Experiment Station of the Railway Technical Research Institute. The red dashed line represents the empirical curve given by Eq. (1) [4], where ρ_N denotes the newly fallen snow density [kg m^{-3}] and T_a denotes the air temperature [°C].

$$\rho_N = 53 . \exp\left(0.488\, T_a\right) + 37.0 \quad (1)$$

For all temperature ranges, the correlation coefficient between air temperature and newly fallen snow density was 0.48; however, when considering only subfreezing conditions, the correlation decreased to 0.26, indicating virtually no relationship. This is because air temperature primarily reflects the presence or absence of melting, rather than differences in hydrometeor types such as aggregate and graupel. As shown in Fig. 2, even under temperatures below 0 °C, the density can vary by a factor of two to three depending on the dominant particle type, demonstrating that the influence of hydrometeor type on snow density cannot be neglected. Therefore, in this study, we developed a method to estimate newly fallen snow density by incorporating information on hydrometeor characteristics, using data obtained from a disdrometer capable of automatically observing particle size and fall-speed distributions [11].

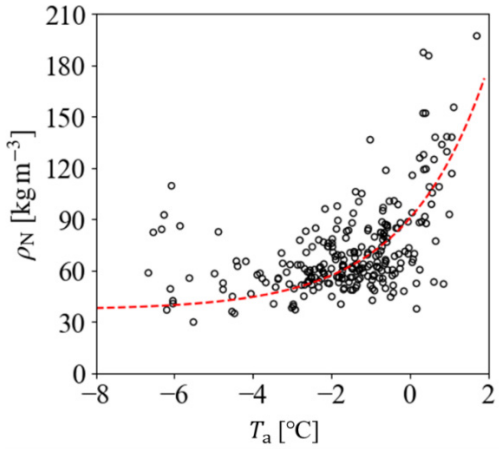


Fig. 1 Relationship between air temperature and newly fallen snow density

2. Methods

2.1 Data

Direct observations of newly fallen snow density and ground-based measurements of snow particles were conducted at the Shiozawa Snow Disaster Prevention Experiment Station of the Railway Technical Research Institute (37.04°N, 138.85°E), located in Minami Uonuma City, Niigata Prefecture, during the winter seasons of 2020-2021 to 2023-2024. Newly fallen snow density was measured using a snow board made of expanded polystyrene insulation. The board was placed outdoors during snowfall events. After one hour of exposure, the depth of snow accumulated on the board was measured with a ruler, and its weight measured using an electronic balance.

Snow particles were observed using disdrometers capable of measuring particle size, fall speed, and shape. An optical disdrometer (Parsivel^2; OTT Hydromet GmbH) and a video disdrometer (G-PIMMS: Ground based Particle Image and Mass Measurement System [12]) were installed within a wind shielding net to conduct ground-based observations of snowfall particles (Fig. 3). The Parsivel^2 is a visible light optical disdrometer that determines particle size and fall speed from the attenuation and duration of a laser beam as hydrometeors pass through the measurement area. For each output interval (set to 1 min in this study) the instrument reports the number of particles in a 32 × 32 matrix of size-fall speed bins (particle size: 0-26 mm; fall speed: 0-22.4 m s^{-1}). In this study, the central values of each bin were used as the representative particle size and fall speed.

The G-PIMMS instrument is designed to directly capture images of precipitation particles, particularly snow. When a particle passes through the infrared sensor mounted on the instrument's lid, a strobe light is triggered, and a CCD camera records an image covering an area of 30 mm in width and 20 mm in height[12]. Advantages of the G-PIMMS include the ability to derive shape-related parameters (e.g., aspect ratio, circularity) directly from the images and to visually identify hydrometeor types. However, unlike the Parsivel^2, it cannot measure fall speed. In addition, the instrument occasionally fails to record data when the opening becomes covered by snow, resulting in reduced observational stability.

Therefore, in this study, Parsivel^2 observations were used for quantitative analyses, while the G-PIMMS was used to visually classify hydrometeor types, as shown in Fig. 2. All observational

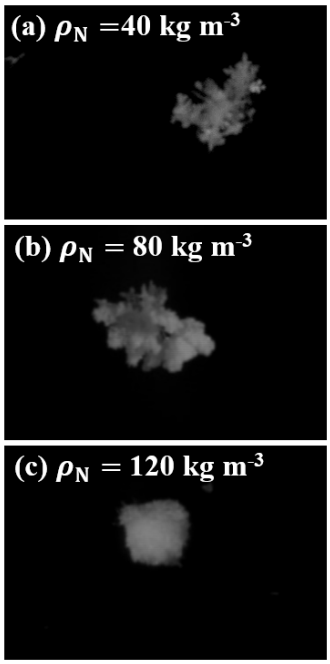


Fig. 2 Example images of snow particles observed at ρ_N=40, 80 and 120 kg m^{-3}



Fig. 3 G-PIMMS and Parsivel^2 installed inside the snow-protection net [11].

data were aggregated at one-hour intervals to match the evaluation period of newly fallen snow density.

2.2 Degree of riming

In the absence of melting (i.e., when no liquid water is present within the snowpack), the snow density is determined by the amount of pore space within a unit volume, since the density of ice itself is fixed. Consequently, newly fallen snow becomes denser as riming progresses, because small ice particles and supercooled droplets fill the voids within the aggregate, corresponding to the transition from particles shown in Fig. 2(a) to the more heavily rimed particles in Fig. 2(b) and Fig.2(c). Focusing on this riming process, we proposed a method to quantify the average degree of riming using the particle size and fall speed distributions obtained from Parsivel^2 [11]. Following previous studies [13, 14], we adopted the assumption expressed in Eq. (2) that, for a given particle size D[mm], the density of a snow particle ρ_s [g cm^{-3}] depends on the degree of riming f_{rim}.

$$\rho_s(D) = \alpha_1 f_{rim}(D) D^{\beta_1} \quad (2)$$

A degree of riming of f_{rim}=1 corresponds to an unrimed, lightly rimed aggregate, whereas f_{rim}=5 represents heavily rimed particles

such as graupel. Following previous studies [15], we assumed that the degree of riming can be expressed as a function of particle size, as shown in Eq. (3), where f_{rim} is defined as the square of the ratio between the observed fall speed V [m s^{-1}] and the empirical fall speed of an unrimed aggregate V_0 [m s^{-1}].

$$f_{\mathrm{rim}}(\mathrm{D})=\left(\frac{V(D)}{V_0(D)}\right)^2 \tag{3}$$

Assuming that snow particles have reached their terminal velocity, where gravitational and drag forces are balanced, the square of the fall speed represents the ratio of the particle density ρ_s to the drag coefficient C_d for a given particle size, as expressed in Eq. (4). Thus, heavily rimed particles, which are denser and more spherical (and therefore have smaller drag coefficients), exhibit larger values of f_{rim}(D) in Eq. (3).

$$V^2=\frac{4g}{3\rho_a}\frac{\rho_s}{C_d}D \tag{4}$$

Here, ρ_a enotes the air density [kg m^{-3}], and g is the gravitational acceleration. For the empirical fall speed of unrimed aggregate V_0, we employed the empirical relationship given in Eq. (5) [16].

$$V_0(\mathrm{D})=0.8\,D^{0.16} \tag{5}$$

Because f_{rim} in Eq. (3) is defined as a function of particle size, a representative value must be derived for each observation case in order to compare it with the newly fallen snow density. In this study, we adopted a precipitation weighted averaging approach and calculated F_{rim}, as defined in Eq. (6), to represent the average degree of riming. We then examined the relationship between F_{rim} and the observed newly fallen snow density.

$$F_{\mathrm{rim}}=\frac{\sum_i f_{\mathrm{rim}}(D_i)M_0(D_i)N(D_i)\Delta D_i}{\sum_i M_0(D_i)N(D_i)\Delta D_i} \tag{6}$$

Here, D_i and ΔD_i represent the particle size and bin width [mm] of the i-th Parsivel^2 size bin, respectively; $N(D_i)$ denotes the number of particles in that bin. The mass of an unrimed snowflake M_0 [mg] was calculated using the empirical relationship given in Eq. (7) [16].

$$M_0=0.073\,D^{1.4} \tag{7}$$

3. Results

3.1 Relationship between degree of riming and newly fallen snow density

The relationship between F_{rim} and the directly measured newly fallen snow density: ρ_N obtained using a snow board was examined. The analysis included 157 cases observed over the 2020-2021 and 2021-2022 winter seasons, during which the air temperature was below 0 °C and the snowfall depth exceeded 10 mm. Cases with snowfall depths less than 10 mm were excluded because measurement uncertainties become relatively large. Figure 4 shows the relationship between the average degree of riming F_{rim} and ρ_N. The correlation coefficient was 0.78, which is substantially higher than the value of 0.26 obtained using the air temperature-based estimation method (Fig. 1). A logarithmic regression was applied to the distribution shown in Fig. 4, yielding the empirical relationship expressed in Eq. (8) [11].

$$\rho_N=42.3\ln(F_{\mathrm{rim}})+53.2 \tag{8}$$

Figure 5 shows example images of snow particles from four cases with different dominant hydrometeor types. As shown in Fig. 4 and 5, the estimation method developed in this study (Eq. 8) can be applied not only to cases dominated by clearly identifiable particle types such as graupel or aggregate, but also to intermediate forms

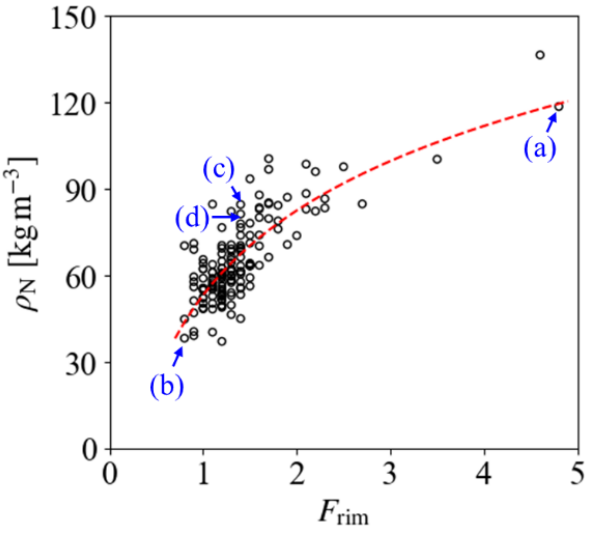


Fig. 4 Relationship between the average degree of riming F_{rim} and the newly fallen snow density ρ_N for the 2020-2021 and 2021-2022 winter seasons. Panels (a)-(d) correspond to the cases illustrated in Fig. 5. The red dashed line represents the empirical curve given by Eq. (8) [11].

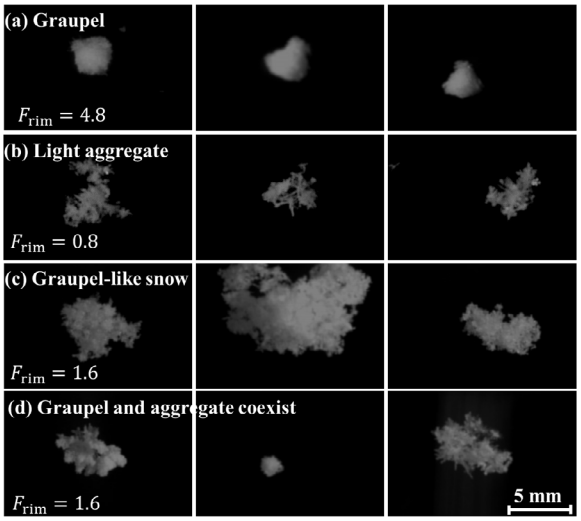


Fig. 5 Example images of snow particles observed in the four cases indicated by arrows in Fig. 4 [11].

such as graupel-like snow (Fig. 5c) and mixed cases in which graupel and aggregate coexist (Fig. 5d). This demonstrates that the method provides a seamless application across a continuum of particle types.

3.2 Comparison with temperature-based estimation methods

Figure 6 compares the proposed method based on the empirical relationship in Eq. (8) [11] with the conventional temperature-based estimation method given by Eq. (1) [4]. Figure 6 shows the observed newly fallen snow densities and the corresponding estimates from both methods for 67 cases observed at Shiozawa during the 2022-2023 and 2023-2024 winter seasons. The root mean square error (RMSE) was 10.5 kg m^{-3} for the developed method in this

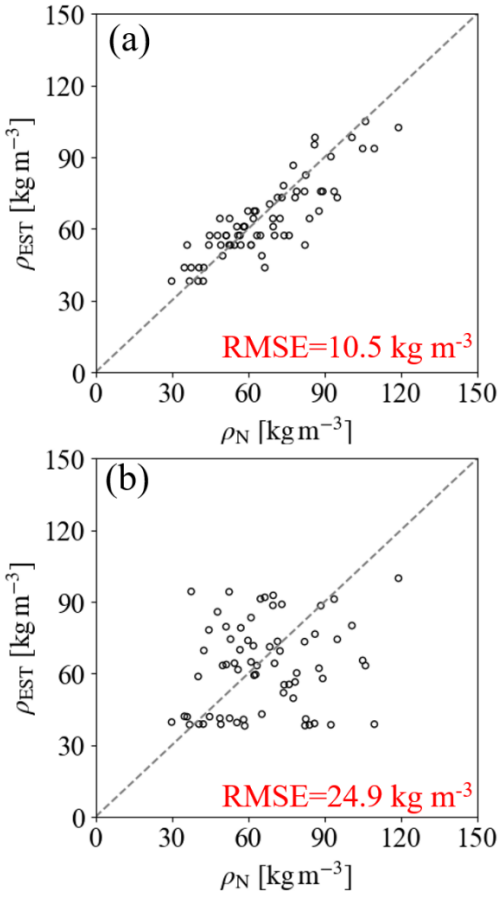


Fig. 6 Observed newly fallen snow densities (ρ_N) at Shiozawa during 2022-2023 and 2023-2024 winter seasons; and estimated newly fallen snow density (ρ_{EST}) using (a) the empirical formula based on disdrometer observations [11]: Eq. (8) and (b) the conventional temperature-based formula [4]: Eq.(1).

study, on the other hand, that of the temperature-based method was 24.9 kg m^{-3}. These results indicate that Eq. (8) provides substantially more accurate estimates of newly fallen snow density than the conventional temperature-based approach.

4. Conclusion

In this study, we developed a method for estimating newly fallen snow density, which is essential for converting precipitation amounts into snowfall depth. Conventional approaches on simple temperature based empirical formulas exhibit poor accuracy for particularly dry snow under subfreezing conditions. To address this limitation, we utilized data from an optical disdrometer capable of automatically measuring particle size and fall speed distributions.

In the absence of melting (i.e., when no liquid water is present within the snowpack), the density of newly fallen snow depends on the amount of pore space within a unit volume, as the density of ice itself is fixed. Snow density therefore increases as riming progresses, with small ice particles and supercooled droplets filling the internal voids of aggregate. Focusing on this process, we proposed a method to quantify the average degree of riming using particle size and fall speed distributions obtained from the optical disdrometer. Analysis of 157 cases observed at the Shiozawa Snow Disaster Prevention Experiment Station during the 2020-2021 and 2021-2022 winter seasons revealed a correlation coefficient of 0.78 between the average degree of riming and newly fallen snow density, substantially higher than the correlation of 0.26 obtained from conventional air temperature-based method. From these results, we derived an empirical relationship between the average degree of riming and snow density [11].

We compared the performance of the proposed method with that of the conventional temperature based empirical formula[4] using observations from the 2022-2023 and 2023-2024 winter seasons. The RMSE values were 10.5 kg m^{-3} for the developed method in this study and 24.9 kg m^{-3} for the temperature-based method, demonstrating a substantial improvement in estimation accuracy.

In the future, we plan to develop a method for estimating and forecasting snow weight and accumulation at arbitrary locations along railway lines.

Acknowledgment

Observational data used in this study were obtained through a collaborative research project conducted by the Railway Technical Research Institute, Yamaguchi University, and the Japan Weather Association.

References

[1] Kamada, M., Takahashi, D., Kurihara, Y., Yokokura, A., and Iikura, S., "Method of Predicting the Snow Accretion Amount on Train Bogies in Consideration of Snow Properties on the Track," *RTRI Report*, Vol. 29, No. 1, pp. 11-16, 2015 (in Japanese).

[2] Judson, A. and Doesken, N., "Density of freshly fallen snow in the Central Rocky Mountains," Bulletin of the American Meteorological Society, Vol. 81, pp. 1577-1587, 2000.

[3] Kajikawa, M., Goto, H., Saruwatari, T., Kanaya, K., Hashimoto, M., and Kikuchi, K., "Relationship between new fallen snow density and shape of snow crystals," *Seppyo,* Vol. 67, pp. 213-219, 2005 (in Japanese).

[4] Noguchi, M., Kobayashi, Iikura, S., Kawashima, K., Iyobe, T., and Fujii, T., "Snow control measures on the Tohoku Shinkansen line between Hachinohe and Shin-Aomori: Part 6-Air temperature dependence of freshly fallen snow density measured during snowfall-," *Proceedings of the 2010 Snow and Ice Research Conference (Sendai),* p. 174, 2010 (in Japanese).

[5] Wayand, N.E., Massmann, A., Butler, C., Keenan, E., Stimberis, J., and Lundquist, J.D., "A meteorological and snow observational data set from Snoqualmie Pass (921 m), Washington Cascades, USA," *Water Resources Research*, Vol. 51, pp. 10092-10103, 2015.

[6] Ishizaka, M., Motoyoshi, H., Yamaguchi, S., Nakai, S., Shiina, T., and Muramoto, K.I., "Relationships between snowfall density and solid hydrometeors, based on measured size and fall speed, for snowpack modeling applications," *The Cryosphere*, Vol. 10, pp. 2831-2845, 2016.

[7] Helfricht, K., Hartl, L., Koch, R., Marty, C., and Olefs, M., "Obtaining sub daily new snow density from automated measurements in high mountain regions," *Hydrology and Earth System Sciences*, Vol. 22, pp. 2655-2668, 2018.

[8] Endo, Y., Kominami, Y., Yamanoi, K., and Niwano, S., "Depth

and density of hourly new snow computed with viscous compression model," *Seppyo*, Vol. 64, pp. 3-13, 2002 (in Japanese).
[9] Lehning, M., Bartelt, P., Brown, B., and Fierz, C., "A physical SNOWPACK model for the Swiss avalanche warning: Part III: Meteorological forcing, thin layer formation and evaluation," *Cold Regions Science and Technology,* Vol. 35, pp. 169-184, 2002.
[10] Yokoyama, K., Ohno, H., Kominami, Y., Inoue, S., and Kawakata, T., "Performance of Japanese precipitation gauges in winter," Seppyo, Vol. 65, pp. 303-316, 2003 (in Japanese).
[11] Takami, K., Kamamoto, R., Suzuki, K., Yamaguchi, K., and Nakakita, E., "Relationship between newly fallen snow density and degree of riming estimated by particles' fall speed in Niigata Prefecture, Japan," *Hydrological Research Letters,* Vol. 16, No. 4, pp. 87-92, 2022.
[12] Suzuki, K., Nakagawa, K., Kaneko, Y., Oki, R., and Nakamura, K., "Microphysical features of solid/melting particles by ground based direct observations for the GPM/DPR algorithm development," *Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 3941-3944, 2016.
[13] Brandes, E.A., Ikeda, K., Zhang, G., Schonhuber, M., and Rasmussen, R.M., "A statistical and physical description of hydrometeor distributions in Colorado snowstorms using a video disdrometer," *Journal of Applied Meteorology and Climatology*, Vol. 46, pp. 634-650, 2007.
[14] Zawadzki, I., Jung, E., and Lee, G., "Snow studies. Part I: A study of natural variability of snow terminal velocity," *Journal of the Atmospheric Sciences*, Vol. 67, pp. 1591-1604, 2010.
[15] Bukovčić, P., Ryzhkov, A., Zrnić, D., and Zhang, G., "Polarimetric radar relations for quantification of snow based on disdrometer data," *Journal of Atmospheric and Oceanic Technology*, Vol. 57, pp. 103-120, 2018.
[16] Locatelli, J.D. and Hobbs, P.V., "Fall speeds and masses of solid precipitation particles," *Journal of Geophysical Research*, Vol. 79, pp. 2185-2197, 1974.

Author



Kazuya TAKAMI, Ph.D.
Senior Researcher, Meteorological Disaster Prevention Laboratory, Disaster Prevention Technology Division
Research Areas: Snowfall, Radar, Wind hazard, Severe weather

Detectability of Residual Stress Anomalies in Various Types of Railway Wheels Subjected to Thermal Overload Based on Back Gauge Change

Kazuyuki HANDA
Materials Technology Division

Wheel rim temperature increases due to tread braking can lead to maintenance challenges such as thermal cracking. While heating within the anticipated design range should not adversely affect the structural safety of the wheels, excessive temperature rises resulting from brake system failures can induce tensile residual stresses in the rim, potentially leading to wheel fractures. Detecting wheels with abnormal residual stress is therefore crucial for railway safety, and monitoring residual displacements in the rim is an implementable approach. However, the shape of wheel webs varies, and the displacement response to thermal input differs depending on the wheel geometry. This study aims to explore general deformation behaviors related to wheel shape and elucidate the relationship between shape and heat resistance. The elastic-plastic finite element analysis revealed that residual deformation after excessive thermal loading consistently results in back gauge expansion, although the deformation and stress responses differ based on the web geometry. Differences in heat resistance are indicated by the variation in the temperature at which compressive yielding onsets. The residual stress detectability of each wheel is determined by the principle that a greater residual displacement at a lower stress indicates higher anomaly detectability.

Key words**:** *wheel, tread brake, residual stress, detectability, back to back distance*

This article is an excerpted and rewritten version of the previously published paper, [Kazuyuki HANDA, Detectability of residual stress anomalies in railway wheels with different geometries subjected to thermal overload based on back gauge change, Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit 239.1 (2025) pp. 58-69.] Copyright © 2025 by IMechE. Reprinted by Permission of SAGE Publications

1. Background and Objective

Railway wheels are safety-critical components designed and manufactured to endure operational loads and prevent failure. The wheel rim, initially under compressive residual stress from heat treatment, is designed to withstand usage without fracturing even with tread surface defects [1] [2]. However, overheating from tread brake malfunctions can cause thermal stress which leads to compressive yielding and plastic deformation. Subsequent cooling can result in the development of tensile residual stresses, diminishing the safety margin [3]. Wheels with excessive tensile residual stresses can potentially fracture, leading to catastrophic consequences.

The surface of tread-braked railway wheels is subjected to stresses and strains due to rolling contact and frictional heat. In particular, the heat input caused by tread brakes is unique to tread-braked railway wheels. This heat input causes changes in material properties, such as hardness and yield stress, as the temperature rises [4] [5]. These changes accelerate fatigue and wear responses, resulting in surface damages such as tread thermal cracks [6-10]. Importantly, it is the macroscopic residual stresses in the rim that determine whether surface cracks will propagate internally beyond the thermomechanical ratcheting zone [11] [12]. Thus, understanding the residual stress state from thermal histories is crucial for preventing wheel failures.

Corresponding to the stress state in the wheel rim, the axial position of the rim section changes due to thermal expansion when the temperature of the wheel rim rises. Under normal conditions, the rim section reverts to its original position upon cooling. However, if the wheel rim overheats and undergoes plastic deformation, the rim position may not fully recover, signalling a change in residual stress. Therefore, measuring the back to back distance between wheels, shown in Fig. 1 (a), can potentially be used to detect anomalies in the residual stress at the wheel rim. In fact, measurement of the back to back distance is recommended for assessing the severity of thermal overload [13].

Although many studies have focused on stress changes due to thermal load, few studies have investigated the extent to which the back to back distance changes with increases in temperature, and how these stress changes correlate with changes in the back to back distance [8] [14] [15]. Moreover, wheel geometry varies among wheels, leading to different responses to changes in stress state. Therefore, to obtain accurate information on the stress state of a wheel from changes in back to back distance, an evaluation specific to each wheel geometry is required. However, the variety of wheel geometries currently in use is huge.

To examine the relationship between stress state and back gauge distance changes for wheels with various web geometries, one possible approach is to conduct a discussion using the geometry itself as a parameter. However, the shape of the wheel web is designed to meet requirements, making it difficult to represent with simple geometric parameters. Therefore, this study aims to obtain insights by comparing multiple existing wheel geometries. It is important to note that the objective is not merely to compare the results of calculations for various geometries, but to use these actual wheel geometries as model cases to understand the trends derived from wheel geometries.

Although various studies have examined the stress state of wheels during braking, they typically focus on the stress changes in wheels with specific geometries [16-18]. There are very few cases where the responses of multiple wheel geometries to identical conditions have been investigated. For certain wheel geometries, compliance with standard requirements confirms that the response to a

given input is within the specified limits [19]. Some studies have investigated the response of wheels with specific geometries through dynamometer tests. However, the responses of many wheels remain uncharacterized. If insights that apply across different wheel geometries can be obtained, and a universally applicable evaluation method can be established, it would contribute to improving the safety of tread-braked wheels.

This study aims to investigate changes in back gauge distances of wheels with various geometries during overheating and subsequent cooling. It seeks to extract insights related to wheel geometry. The results obtained from elastic-plastic finite element analysis, which considers non-linear material properties where thermal-stress-induced deformation surpasses the elastic range leading to plastic deformation, will be validated against experimental data. This paper will also discuss residual stresses to establish a critical temperature threshold for tread-braked wheels.

2. Methodology

2.1 Object of analysis

This study focuses on seven types of wheels currently in use in Japan, characterized by different web geometries. According to the classification by Okagata [20], wheels with an outward sloping web are categorized as Type A, while those with an inward sloping web are classified as Type B, as shown in Fig. 1 (b). Type B wheels are predominantly used on driving bogies for narrow gauge tracks where space between wheels is limited but must accommodate the driving motor and gears. Both Type A and Type B wheels are available in two different diameters. These wheels have initial diameters of 860 mm and 810 mm, respectively, with each of the four wheel designs featuring different web geometries. Notably, the 810 mm wheels have more inclined and offset web geometries [20]. Based on the classification, four types of wheels were modelled, designated as 860A, 860B, 810A, and 810B.

In contrast, the S-shaped wheel incorporates a different design concept from Type A and Type B wheels, characterized by its S-shaped web, which meets EN standard requirements for thermal stability [21]. Only one variant was analyzed since both share identical web shapes. This wheel model is designated as 860S.

The remaining categories include corrugated wheels, which are lighter due to their corrugated web geometries that achieve stiffness with reduced plate thickness [20]. Two versions of corrugated wheels exist, both with an initial diameter of 860 mm, and are classified as A corrugated and B corrugated based on the direction of the web sloping. The Type A corrugated wheel is engineered to mitigate stresses induced by thermal loads and is denoted as the N-A type corrugated wheel [20] [22]. The corrugated wheel models are designated as N-Ac and Bc.

In total, there are seven distinct wheel types: five with an initial diameter of 860 mm designated as 860A, N-Ac, 860B, Bc and 860S, and two with an initial diameter of 810 mm, classified as 810A and 810B. These wheels have a reprofiling allowance of 40 mm in radius at the outer rim. Numerical analysis was conducted using finite element (FE) models for both the initial and use limit diameters of these seven wheel types. In this study, the new diameter (860 mm) of the 860A wheel is referred to as 860A860, and the use limit diameter (780 mm) as 860A780. This nomenclature was applied across all models, encompassing seven wheel types and 14 model types. The FE models are presented in Fig. 1. The FE model of the corrugated wheels was developed in 3D, whereas the other wheels were modeled in 2D.

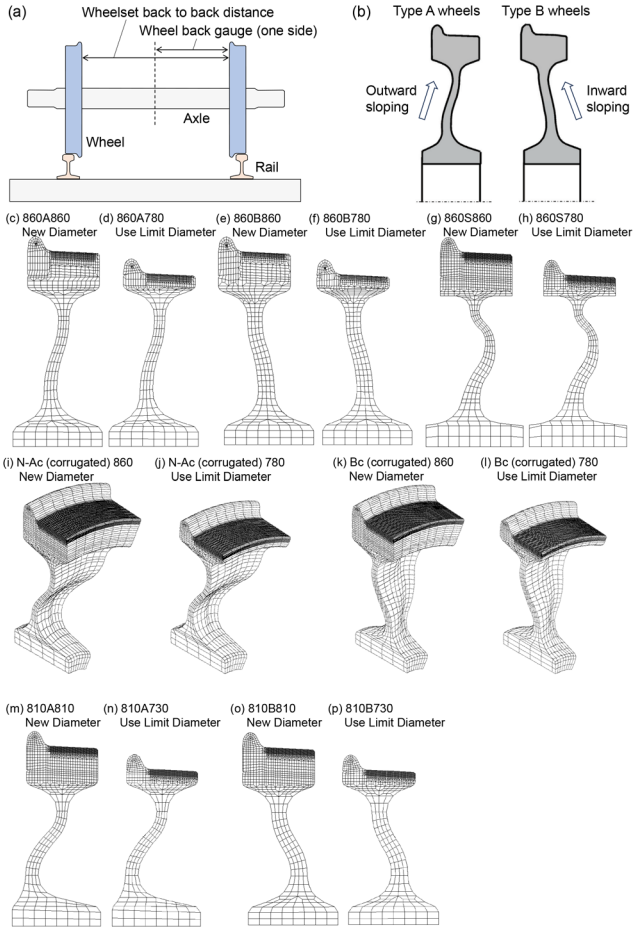


Fig. 1 Schematic showing (a) location of wheel back gauge in a railway wheelset and (b) direction of web slope. Outline of FEA models is presented for 860A wheels with (c) new diameter and (d) use limit diameter; similarly for 860B wheels (e) and (f); for 860S wheels (g) and (h); for N-A corrugated wheels (i) and (j); for B corrugated wheels (k) and (l); for 810A wheels (m) and (n); and for 810B wheels (o) and (p).

2.2 Analysis method

In this study, the effect of differences in wheel geometry on residual deformation and residual stress caused by excessive heat input is investigated. Applying consistent thermal loads to various wheel types in dynamo tests with actual wheels is challenging due to reproducibility, time, and cost. The primary challenge is the variation in the friction coefficient of the tread brake block. Consequently, this study employs finite element analysis to compare the effects of different wheel geometries.

The analysis models were constructed using extrapolated data from drawings legally accessible or from published documents. The detailed geometry of wheel webs is generally highly confidential. Therefore, the models may not perfectly replicate the actual product shapes. However, the focus of this study is to understand the behavior of temperature, stress and displacement at the rim. The authors contend that the approximations of web shape adequately simulate the degree of restraint against the rim's deformation.

The analytical models were two-dimensional (2D) for the

non-corrugated wheels and three-dimensional (3D) for the corrugated wheels. In both types, the element size on the longitudinal cross-section was set to approximately 1 mm square at the tread surface, where braking heat is applied, with a coarser mesh towards the center.

The loading conditions aimed to apply a thermal load to induce residual deformation of the wheel. Applied conditions were selected that would induce residual deformation and residual stress without reaching the austenite transformation temperature of 730°C. The thermal loading conditions were standardized at 33.3 kW for 40 minutes across all wheel diameters, resembling the drag braking condition as outlined in EN 13979/UIC 510-5 [19] [23]. The analysis target time was set to 4 hours, which was sufficient for the temperatures of all elements to return to their initial states. For verification, experimental data involving heating the 860A860 and 860B860 wheels were used.

The elastic-plastic analysis proceeded in two stages. Initially, a transient thermal analysis was performed where the brake block contact area was exposed to a constant heat flux for a set duration. Convective heat transfer and radiation heat dissipation were applied as heat dissipation conditions. Subsequently, thermal-structural one-way coupled analyses were conducted using the obtained temperature distributions. Circumferential displacements of all nodes in both end cross-sections were constrained, and all nodes in the boss inner diameter were fully constrained.

A nonlinear finite element analysis accounting for elastic-plastic deformation was conducted. Temperature-dependent material properties were used. The material properties employed in the analysis were determined from measurements of specimens extracted from a wheel compliant with ISO 1005-6 C64GW-T.

3. Results

3.1 Temperature

The temperature history of the 860A860, 860B860 and N-Ac860 wheels at the tread center surface is depicted in Fig. 2. For wheels of the same diameter, the temperature history was nearly identical, irrespective of the wheel geometry. Fig. 3 (a) illustrates the maximum temperatures at the tread surfaces for each wheel type at 2,400 s, and Fig. 3 (b) displays the corresponding temperature distribution for the 860A860, 860B860, 860S860, N-Ac860 and Bc860 wheels. Under the load conditions of 33.3 kW for 2400 seconds, the wheel surface did not reach 730°C.

3.2 Back gauge changes at elevated temperatures and after cooling

The time history of the back gauge shift of the 860A860 and

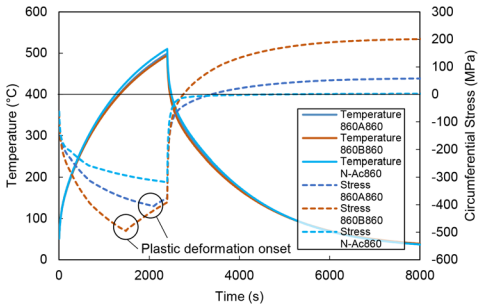


Fig. 2 Temperature and stress histories at tread surface of 860A860, 860B860 and N-Ac860 wheels.

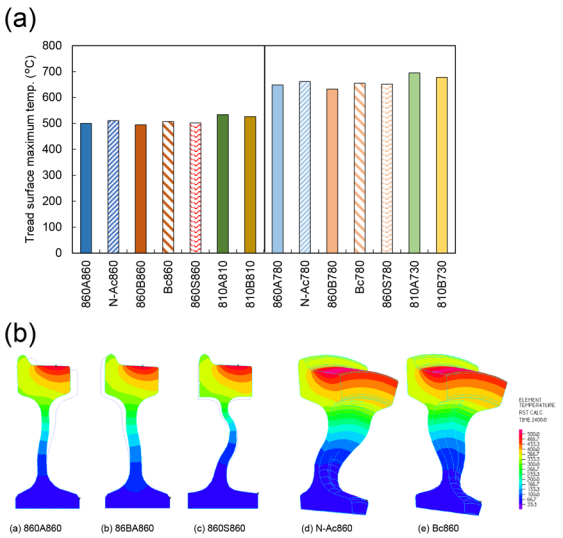


Fig. 3 Wheel temperatures (a) at tread surface of each wheel; and (b) distribution in 860A860, 860B860, 860S860, N-Ac860 and Bc860 wheels at end of braking. Displacements in (b) are displayed at a magnification of five times actual displacement.

860B860 wheels is shown in Fig. 4 (a). The back to back distance of the wheelset is twice the value shown. For the 860A wheel, displacement during temperature rise occurred inward, whereas for the 860B wheel, it occurred outward. The displacement of the 860A wheel was greater than that of the 860B wheel. This deformation is shown in Fig. 3 (b) with a magnification factor of five. Corresponding to the web's slope direction, a back gauge change occurred in the direction permitting thermal expansion of the rim. The 860S wheel allowed expansion primarily in the radial direction, resulting in only a slight back gauge change during temperature rise. The back gauge displacement for each wheel type at 2,400 s is depicted in Fig. 4 (b), showing variations in both the direction and magnitude of displacement depending on the web geometry.

Figure 4 (c) displays the back gauge change at the end of the analysis (14,400 s), with the back to back distance change being twice this value. This displacement was taken at the position of maximum displacement on the rim. As shown in Fig. 3 (a), the maximum temperature of the tread surface was approximately 500°C for the wheels with the new diameter, while it reached 600°C-700°C for the wheels at the use limit diameter, resulting in residual deformation. In particular, the 860A780 wheel exhibited residual deformation. A key observation across all geometries is that residual deformation after thermal loading occurs on the side of back gauge expansion.

3.3 Residual stress after cooling

This analysis focuses on the stress changes resulting from a single heat input. Simplified conditions were adopted to understand the response differences across wheel geometries.

Figure 2 presents the temperature and stress transition at the tread surface of the 860A860 and 860B860 wheels. The stresses are depicted as circumferential normal stress. As the temperature rises, compressive stress develops. Upon reaching the yield stress, the material yields, leading to plastic deformation. As compressive plastic deformation progresses, stress values decrease with temperature rise. Once braking concludes and the system begins to air-cool, the

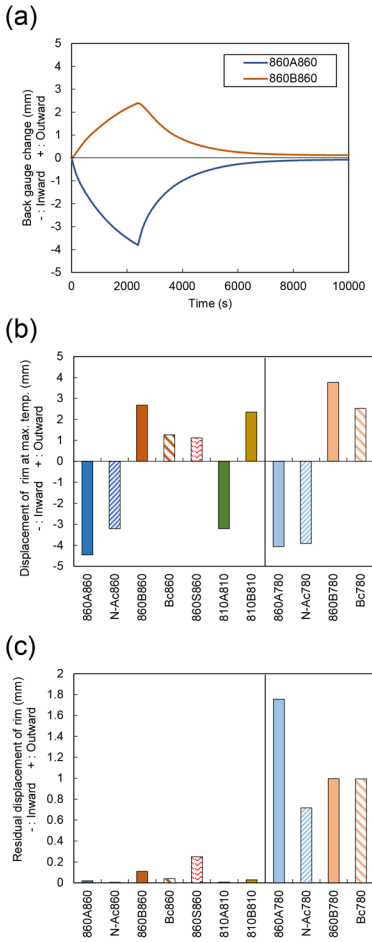


Fig. 4 Wheel back gauge displacement (a) of 860A860 and 860B860 wheels over time; (b) of each wheel at maximum temperature time point (2,400 s); and (c) after cooling (14,400 s).

stresses relax due to thermal contraction. Subsequently, the stresses reverse, generating tensile residual stresses corresponding to the compressive plastic deformation strain.

Figure 5 illustrates the maximum values of residual stress in the rim section at the end of the analysis (14,400 s) for each wheel. Tensile residual stresses were observed in all wheel models after cooling. However, the levels of residual stresses varied significantly depending on wheel geometry.

4. Discussion

4.1 Allowable temperatures

Areas of the tread surface that reach the highest temperatures are most susceptible to yielding and plastic deformation due to the maximum temperature gradient. Therefore, the temperature at which yielding begins at the highest temperature elements was determined. This onset of yielding was identified by the peak circumferential compressive stress during thermal loading, corresponding to the stress peak shown in Fig. 2. For all models, the equivalent stresses were also highest when the compressive stresses peaked. The extracted temperature is considered the minimum at which yielding could occur under the given thermal input conditions and is

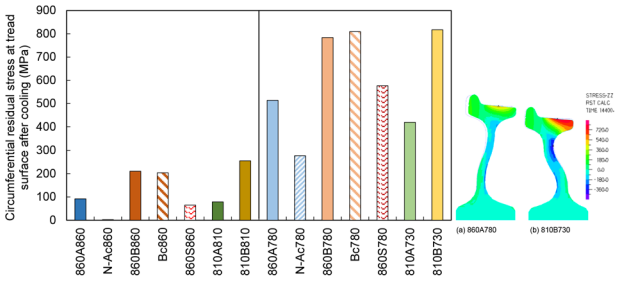


Fig. 5 Circumferential residual stress after cooling of each wheel, and residual stress distribution in 860A780 and 810B730 wheels. Displacements in stress distribution map are displayed at a magnification of five times actual displacement.

referred to as the plastic deformation onset temperature. This temperature can be seen as the maximum allowable temperature for the wheel, as below this threshold, no tensile residual stresses are expected when the initial stress state is neutral.

Figure 6 illustrates the temperatures at which plastic deformation onset occurs for each wheel type. For the N-A corrugated wheel with the new diameter (N-Ac860), no yielding was observed; thus, the values represent the maximum temperature at 2,400 s. Fig. 6 also demonstrates that the plastic deformation onset temperatures are similar for wheels with the same web geometry. This suggests that the ease of deformation at elevated temperatures is influenced by the web shape, which is crucial in determining the plastic deformation onset temperature. This finding indicates that the thermal resistance of a wheel can be assessed using a parameter: the critical temperature tailored to its design geometry.

It is important to note that the plastic deformation onset temperature varied by as much as 150°C among the seven wheel types. This variation ranged from approximately 400°C to 550°C, highlighting the influence of wheel geometry on the thermal response of the wheels. The N-A corrugated wheels exhibited higher allowable temperatures, followed by the 810A wheels and then the 860A and 860S wheels. Conversely, the B corrugated wheels showed allowable temperatures comparable to those of the 860B and 810B wheels. The 860B and 810B wheels displayed relatively low plastic deformation onset temperatures, likely due to their increased stiffness against deformation.

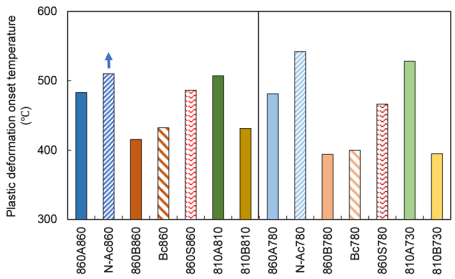


Fig. 6 Plastic deformation onset temperatures at tread surface of wheels investigated.

4.2 Degree of displacement constraint

In analyzing the differences in stress levels during temperature rise across specific wheel geometries, this study highlights the rim displacements of the 860A, 810A, 860S and N-A corrugated wheels

in the plane containing the axle, from 0 s to 2,000 s, as depicted in Fig. 7, alongside the maximum temperature at 2,000 s. Notably, none of the wheels exhibited yielding at this point. The axial displacement towards the inner direction of the wheelset was more pronounced in the 860A wheel, while the radial displacement outward was more significant in the N-A corrugated and 810A wheels. This variation likely contributed to the lower stresses observed in the N-A corrugated wheel. This suggests that the N-A corrugated wheel allows for larger circumferential displacements at lower temperatures, resulting in a higher onset temperature for plastic deformation.

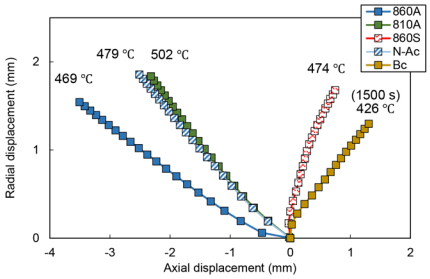


Fig. 7 Change in position of wheel rim on a plane normal to axle during braking.

4.3 Residual stress detectability

The relationship between the maximum circumferential residual stress at the rim of each wheel and the residual deformation (back gauge change) after cooling is illustrated in Fig. 8 (a). From a safety margin perspective, it is beneficial to have small absolute values of residual stress for a given thermal input. Conversely, for the detection of excessive tensile residual stresses, it is preferable for the shape to respond sensitively to changes in the residual stress state. Specifically, the back gauge is suitable for detecting changes using non-contact methods such as laser displacement measurement. If the change is large relative to the generated stress value, the permissible dimensional measurement error can be large.

Noting that greater residual displacement at a smaller generated stress enhances the detectability of anomalies, the residual stress detectability of the wheel can be defined as follows:

$$Drs = \frac{BC}{RS} \quad (1)$$

where
Drs: Residual Stress Detectability factor
BC: Back gauge Change (residual displacement)
RS: Residual Stress

The residual stress detectability for each wheel at the use limit diameter is depicted in Fig. 8 (b). The residual stress detectability factor varied by approximately sixfold depending on the wheel geometry. Wheels exhibiting high residual stress detectability and a high plastic deformation onset temperature are considered to have superior heat resistance and detectability. In this study, these wheels include the 860A, N-Ac, 860S and 810A wheels. Conversely, the 860B, Bc and 810B wheels are perceived to be more vulnerable to thermal loads and less detectable. Among these, only the 860S wheel takes into account the changes in back gauge during and after temperature rise, in compliance with EN requirements [19].

Residual stress detectability can also be applied to other stress values. The integration of residual stress detectability and plastic deformation onset temperature could provide a unified evaluation of heat resistance properties derived from wheel web geometry, offering a practical approach to the prevention of potential wheel rim failures.

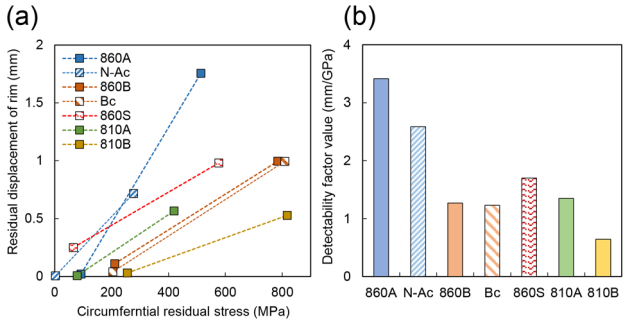


Fig. 8 Residual displacement of wheel rim: (a) its relationship with circumferential residual stress; and (b) proportional coefficient expressed per unit residual stress (mm/GPa), referred to as detectability factor.

5. Conclusions

This study investigated the changes in back gauge distance of wheels with various geometries under thermal loading and subsequent cooling, aiming to derive insights from the wheel geometry. Stresses and deformations were analyzed using the elastic-plastic FE method, assuming stress neutrality in the initial state. The key findings are summarised as follows:

1. The changes in back gauge distance during temperature rise were influenced by the direction of the slope of the wheel web, resulting in either expansion or contraction of the back gauges. For wheels that conformed to the requirements of the EN standard, these changes were minimal.
2. In areas subjected to compressive yielding and plastic deformation due to high temperatures at the wheel tread, the maximum circumferential tensile residual stresses that developed after cooling ranged from approximately 280 MPa to 820 MPa depending on the web geometry.
3. The extent of back gauge change after cooling in wheels subjected to tensile residual stress varied from about 0.5 mm to 1.8 mm depending on the web geometry. However, residual displacement after thermal loading consistently resulted in back gauge expansion.
4. The analyses of back gauge changes during temperature rise for two representative wheel geometries matched experimental measurements conducted under equivalent conditions.
5. The allowable temperature of the wheel, defined as the upper limit temperature where plastic deformation does not occur under a certain thermal input, was determined by the temperature at which compressive yielding occurs at the point of highest temperature on the tread, namely the plastic deformation onset temperature. This onset temperature varied from approximately 400°C to 550°C among the seven wheel types. However, it was consistent for wheels with the same web geometry.
6. The back gauge displacement per unit of tensile residual stress after cooling has been organized as the residual stress detectability factor. This factor indicates the extent of back gauge change per unit of residual stress. In this study, the residual stress detect-

ability factor varied by approximately sixfold depending on the wheel geometry among the seven types of wheels.

Acknowledgment

The author would like to thank F. Morimoto of the Railway Technical Research Institute for technical assistance with numerical simulation models, and R. Imata of the East Japan Railway Company for advice and discussion.

References

[1] Gordon J, Perlman A. B., "Estimation of Residual Stresses in Railroad Commuter Car Wheels Following Manufacture. Final Report," *Federal Railroad Administration, U.S. Department of Transportation*, USA, June 2003.
[2] Molyneux-Berry P, Bevan A, "Residual Stresses in Railway Wheels and Their Effect on Damage Rates through the Life of a Wheel," In: *Computing and Engineering Researchers' Conference*, University of Huddersfield, UK, December 2013: 31-36.
[3] Nishihara S, Tokimasa K., "Study on the Residual Stresses in Railroad Solid Wheels and Their Effect on Wheel Fracture," *Bulletin of the JSME* 1976; 19: 459-468.
[4] Handa K, Kimura Y, Mishima Y. Ferrite and Spheroidized, "Cementite Ultrafine Microstructure Formation in an Fe-0.67 Pct C Steel for Railway Wheels under Simulated Service Conditions," *Met. Mat. Trans. A* 2009; 40: 2901-2908.
[5] Handa K, Kimura Y, Yasumoto Y, et al., "Effect of deformation and annealing temperatures on ultrafine microstructure development and yield strength of pearlitic steel through continuous recrystallization," *Mat. Sci. Eng. A* 2010; 527: 1926-1932.
[6] Ekberg A, Kabo E., "Fatigue of railway wheels and rails under rolling contact and thermal loading—an overview." *Wear* 2005; 258: 1288-1300.
[7] Handa K, Kimura Y, Mishima Y. "Surface cracks initiation on carbon steel railway wheels under concurrent load of continuous rolling contact and cyclic frictional heat." *Wear* 2010; 268: 50-58.
[8] Handa K. Morimoto F. "Influence of wheel/rail tangential traction force on thermal cracking of railway wheels." *Wear* 2012; 289: 112-118.
[9] Caprioli S, Vernersson T, Handa K. "Thermal cracking of railway wheels: Towards experimental validation." *Tribology Int.* 2016; 94: 409-420.
[10] Landström E, Steyn E, Ahlström J, et al. "Thermomechanical testing and modelling of railway wheel steel," *Int. J. Fatigue* 2023; 168: 107373.
[11] Salehi I, Cookson J, Mutton P, et al. "Assessment of Overheated wheels for Increased Serviceability Using the Ultrasonic Birefringence Residual Stress Measurement Technique." In: *IIHA 2015 Conference*, 21-24 June 2015, Perth, Australia.
[12] Strážovec P, Suchánek A, Šťastniak P, et al. "Detection of residual stress in a railway wheel." *Trpn. Res. Proc*. 2019; 40: 898-905.
[13] Joint Network Secretariat, "Normal Procedure Task Force. Broken wheels - Final output," *Report for the EU Agency for railways*, Brussels, 28. November 2019.
[14] Donzella G, Scepi M, Solazzi L, et al. "The Effect of Block Braking on the Residual Stress State of a Solid Railway Wheel." *Proc. IMechE, Part F: J Rail and Rapid Transit* 1998; *212*: 145-158.
[15] Teimourimanesh S, Vernersson T, Lundén R. "Thermal capacity of tread-braked railway wheels. Part 1: Modelling." *Proc. IMechE, Part F: J Rail and Rapid Transit* 2016; 230: 784-797.
[16] Landström E, Vernersson T, Lundén R. "Analysis and testing of tread braked railway wheel—Effects of hot spots on wheel performance." *Int. J. Fatigue* 2024; 180: 108116.
[17] Lundén R. "Contact region fatigue of railway wheels under combined mechanical rolling pressure and thermal brake loading." *Wear* 1991; 144: 57-70.
[18] Walia M, Esmaeili A, Vernersson T, et al. "Thermomechanical capacity of wheel treads at stop braking: A parametric study." *Int. J. Fatigue* 2018; 113: 407-415.
[19] EN13979-1:2003+A2:2011. "Railway applications - Wheelsets and bogies - Monobloc wheels - Technical approval procedure - Part 1: Forged and rolled wheels."
[20] Okagata Y. "Design Technologies for Railway Wheels and Future Prospects." *Nippon Steel & Sumitomo Metal Technical Report* 2013; No. 105: 26-31.
[21] Ronchi A, Tanabe N. "Application of EN-based design for non-European scenarios: the case of Lucchini RS and JR-West." In: *XIX International Wheelset Congress*, Venice, Italy, 16-20 June 2019, paper no. 6A-3.
[22] Yamamura Y, Nakata M, Anjiki M. "Development of Brakeheat-proof Corrugated Wheel." *Sumitomo Metals* 1994; 46: 77-84.
[23] UIC Code 510-5:2007. "Technical approval of monobloc wheels - Application document for standard EN13979-1." *International Union of Railways*.

Author



Kazuyuki HANDA, Ph.D.
Director, Head of Materials Technology Division
Research Areas: Wheel and Brake, Railway Materials

Evaluation of Yield Management with a Randomized Controlled Trial (A/B Test)

Ryosuke MATSUMOTO
Data Analytics Laboratory, Information and Communication Technology Division

This study proposes a method to evaluate yield management through sales slot control for discount railway tickets using a randomized controlled trial (A/B test). Implemented on an operating railway line, this method empirically shows how sales slot adjustments influence revenue. The results demonstrate that the appropriate strategy is time-dependent: in certain periods, increasing slots stimulates demand and boosts revenue, while in others, decreasing slots suppresses down-selling for a similar positive effect. As the first published case of an A/B test for yield management in the Japanese railway industry, this paper provides a practical framework for implementation and analysis.

Key words**:** *yield management, sales slot control, randomized controlled trial, A/B test*

1. Introduction

For which day types and time-of-day periods should a railway operator increase or decrease the number of sales slots for fixed-price, limited-seat tickets? This question is important because it directly affects railway revenue. Yield management involves using sales slot control or price control to enhance revenue for non-storable services, such as transportation services. The opening question is particularly important in contexts where railway pricing is generally fixed and yield management mainly involves sales slot control, as in the Japanese railway market. Moreover, in railway markets such as Japan, where a declining population is reducing both passenger demand and the available workforce, yield management is an effective approach for maintaining or increasing railway revenue without large-scale capital investment and with minimal operational resources.

To explain how sales slot control can increase railway revenue, consider a situation with only two types of online tickets: standard tickets and discount tickets. Under this setting, only discount tickets are sold with a limited number of seats and are subject to sales slot control. In the following discussion, increasing or decreasing the number of sales slots is referred to as expanding or reducing sales slots. First, expanding the sales slots for discount tickets may induce additional demand among passengers with a relatively low willingness to pay. At the same time, however, it may also lead to down-selling; that is, passengers with a relatively high willingness to pay may shift from standard tickets to discount tickets, thereby reducing the average fare per passenger. In this context, passengers with a relatively low willingness to pay may include students and non-regular workers, whereas those with a relatively high willingness to pay may include business travelers. Conversely, reducing the number of sales slots for discount tickets may increase the average fare per passenger by suppressing down-selling. At the same time, however, it may also cause demand loss among passengers who can no longer purchase discount tickets. Therefore, the revenue effect of sales slot control depends on the trade-off between changes in demand and changes in the average fare per passenger. However, it is not easy for railway operators to determine the optimal number of sales slots that balance this trade-off based solely on existing rules of thumb.

To address this issue, this study proposes a method for evaluating the effects of yield management through a field experiment. The field experiment adopted in this study is a randomized controlled trial (A/B test), a method known for its high internal validity in causal inference. In addition, with the cooperation of a railway operator, this study applies the A/B test to a line in revenue service. The results demonstrate that the proposed method works in practice. They also show how increasing or reducing sales slots for discount tickets affects railway revenue.

Among railway operators in Japan, yield management through sales slot control appears to have been practiced traditionally. In academic research, however, studies on yield management have mainly focused on price control [1, 2]. To the author's knowledge, studies focusing on yield management through sales slot control in the railway sector are limited to a series of studies by Nakagawa et al. [3, 4]. Nakagawa et al. developed an algorithm for calculating optimal sales slots by solving a mathematical optimization problem with the objective of maximizing railway revenue. However, their model of passenger behavior in the algorithm was not based on actual demand data. Instead, it relied on a web-based stated-choice survey that asked passengers how they would behave if discount tickets were unavailable because the sales slots had already been filled. They also examined the validity of the algorithm through a case study on a line in revenue service. In the case study, the sales slots calculated by the algorithm were applied in practice. However, the study covered only three days, corresponding to a three-day weekend. For these reasons, some issues remained regarding the internal validity of Nakagawa et al.'s analytical results.

As described above, the randomized controlled trial (A/B test) adopted in this study is considered to offer particularly high internal validity in causal inference. For example, Cabinet Office, Government of Japan [5] classifies the internal validity of results obtained from causal inference methods into six levels: 1a, 1b, 2a, 2b, 3, and 4. In this classification, meta-analyses comparing the results of multiple randomized controlled trials (A/B tests) are assigned to the highest level, 1a. Randomized controlled trials (A/B tests) are assigned to the second highest level, 1b. Although "randomized controlled trial" and "A/B test" essentially refer to the same method, this paper uses the term "A/B test" hereafter because it better fits the following discussion.

An A/B test is a type of social experiment in which two patterns, A and B, are prepared. These patterns differ only in the factor to be tested. Subjects, such as individuals or days, are then randomly assigned to one of the two groups, and each pattern is applied to the corresponding group. The results are then compared. In a typical A/B test, the current or baseline measure is applied to the control group, while a new measure is applied to the treatment group. The effect of the new measure is then statistically evaluated by compar-

ing outcome variables, such as sales or other key performance indicators (KPIs), between the two groups. Random assignment of subjects helps ensure that the groups are comparable. This makes it possible to statistically determine whether the difference in results is attributable to the measure itself.

However, to the author's knowledge, no cases of A/B testing for railway yield management have been published. In addition, many studies have reported the implementation and analysis of A/B tests, mainly in the context of online experiments. However, previous studies have also identified potential pitfalls in A/B testing. For example, assigning participants to A/B groups inappropriately or using unsuitable statistical measures for a given case can lead to false-negative or false-positive results [6]. Accordingly, when applying an A/B test to railway yield management, it is essential to establish an appropriate experimental design and analytical method in advance to avoid these pitfalls.

The remainder of this paper is organized as follows. Section 2 proposes the implementation and analytical methods for the A/B test. Section 3 applies the A/B test to an operating railway line and analyzes the results. Section 4 concludes the paper.

2. Implementation and analytical methods for A/B test

This section first identifies the appropriate unit of assignment to A/B groups. It then provides a mathematical demonstration showing that effect evaluation can be conducted using this assignment method. Finally, it presents the specific procedure for evaluating yield management effects using the A/B test.

2.1 Assignment unit for A/B groups

The key requirement in determining the appropriate unit of assignment to A/B groups is to satisfy the no-interference assumption. Under this assumption, the outcome of a unit is not affected by the treatment assignments of other units.

For the route segment examined in this study, the minimum unit for setting sales slots for discount tickets is the train. For example, the number of sales slots can be set for each individual train, such as 10 slots for a train on a specific day. Table 1 illustrates how random assignment to A/B groups would be conducted at the train level and at the daily level. In Table 1, the numbers of sales slots in patterns A and B are set at 10 and 20, respectively.

Using Table 1, this subsection examines whether A/B groups should be assigned at the train level or at the daily level. First, suppose that A/B groups are assigned at the train level. Consider a passenger who wishes to purchase a discount ticket for Train 3, but cannot do so because all the sales slots have been filled. In this case, the passenger may switch to an adjacent train, such as Train 1 or Train 5, and purchase a discount ticket for one of those instead. However, this creates interference in demand across A/B groups, so train-level assignment does not satisfy the no-interference assumption. Next, suppose that A/B groups are assigned at the daily level. Consider a passenger traveling on October 13 who wants to purchase a discount ticket for Train 3. In this case, the passenger may still switch to another train on the same day. At the same time, as discussed below, this study focuses on normal days, when passengers are assumed to be unlikely to change their travel date to the previous or following day. For this reason, assignment on a per-day basis is considered to satisfy the no-interference assumption.

From the above discussion, this study concludes that the daily level is the appropriate unit for assigning A/B groups and testing the effects of yield management.

Table 1 Example of A/B group assignment

Panel A: Train-level assignment

Train	A/B group (random assignment)	Sales slots for discount tickets
Train 1	B	20
Train 3	A	10
Train 5	B	20
Train 7	A	10
...	...	...

Panel B: Daily-level assignment

Date	A/B group (random assignment)	Sales slots for discount tickets
October 12	B	20 slots for all trains
October 13	A	10 slots for all trains
October 14	B	20 slots for all trains
October 15	A	10 slots for all trains
...	...	...

2.2 Mathematical demonstration of A/B test validity

For a given date d, the treatment effect of Group B relative to Group A, denoted by τ_d, is defined as in (1). The average treatment effect across multiple sample dates, denoted by τ^{ATE}, is defined as in (2). Although this study uses the labels Group A and Group B rather than treatment group and control group, the term "treatment effect" is widely used in the literature and is therefore also used in this paper.

$$\tau_d \equiv Y_d^B - Y_d^A \tag{1}$$

$$\tau^{\mathrm{ATE}} \equiv \mathrm{E}[Y_d^B - Y_d^A] \tag{2}$$

Here, Y_d^G denotes a potential outcome, such as demand or farebox revenue, that would be observed on date d if date d were assigned to group $G \in \{A, B\}$. However, because a given date d belongs to either Group A or Group B, but not both, Y_d^A and Y_d^B cannot both be observed in the actual data. This is known as the "fundamental problem of causal inference". Therefore, the date-specific treatment effect τ_d cannot be directly measured in real-world data. The average treatment effect τ^{ATE}, defined as an expectation, corresponds to the effect of yield management estimated in this study. However, without randomization, τ^{ATE} is not necessarily identified if observations are assigned to Group A and Group B at the analyst's discretion. Under such discretionary assignment by the analyst, the dates assigned to Group A and Group B are denoted by G_d= A and G_d = B, respectively. Then, the difference between the expected outcome for Group B and the expected outcome for Group A, denoted by τ^G, can be expressed as in (3).

$$\begin{aligned}\tau^G &\equiv \mathrm{E}[Y_d^B \mid G_d=\mathrm{B}] - \mathrm{E}[Y_d^A \mid G_d=\mathrm{A}] \\ &= \mathrm{E}[Y_d^B \mid G_d=\mathrm{B}] - \mathrm{E}[Y_d^A \mid G_d=\mathrm{A}] \\ &\quad + \mathrm{E}[Y_d^A \mid G_d=\mathrm{B}] - \mathrm{E}[Y_d^A \mid G_d=\mathrm{B}] \\ &= \mathrm{E}[Y_d^B - Y_d^A \mid G_d=\mathrm{B}] \\ &\quad + \mathbf{E}[\boldsymbol{Y_d^A \mid G_d=\mathrm{B}}] - \mathbf{E}[\boldsymbol{Y_d^A \mid G_d=\mathrm{A}}]\end{aligned} \tag{3}$$

The second and third terms in the last line of (3), shown in bold, are referred to as self-selection bias. Specifically, self-selection bias captures the difference between the expected KPI under Group A for dates assigned to Group B and the expected KPI under Group A for dates assigned to Group A. Because this value does not necessarily equal zero, τ^G may therefore be biased. It should also be noted that this value does not converge to zero even as the number of observations increases.

The first term in (3), $\mathrm{E}[Y_d^B - Y_d^A \mid G_d = \mathrm{B}]$, represents the average treatment effect for dates assigned to Group B, which is the average treatment effect on the treated (ATT). This value is denoted by τ^{ATT}. The condition under which τ^G equals τ^{ATT} is that the self-selection bias is zero; that is, $\mathrm{E}[Y_d^A \mid G_d = \mathrm{B}] = \mathrm{E}[Y_d^A \mid G_d = \mathrm{A}]$. When group assignment is random, conditioning on G_d provides no additional information for calculating expectations. Therefore, the conditional expectations given G_d are equal to the corresponding unconditional expectations. Thus, the relationship in (4) holds. By combining (3) and (4), (5) is obtained. This shows that randomization ensures that τ^G equals τ^{ATE}.

$$\begin{aligned}&\mathrm{E}[Y_d^B \mid G_d=\mathrm{B}]=\mathrm{E}[Y_d^B \mid G_d=\mathrm{A}]=\mathrm{E}[Y_d^B]\\&\mathrm{E}[Y_d^A \mid G_d=\mathrm{B}]=\mathrm{E}[Y_d^A \mid G_d=\mathrm{A}]=\mathrm{E}[Y_d^A]\end{aligned} \tag{4}$$

$$\begin{aligned}\tau^G &\equiv \mathrm{E}[Y_d^B \mid G_d=\mathrm{B}]-\mathrm{E}[Y_d^A \mid G_d=\mathrm{A}]\\&=\mathrm{E}[Y_d^B-Y_d^A \mid G_d=\mathrm{B}]\\&\quad+\mathbf{E}[\boldsymbol{Y_d^A \mid G_d=\mathrm{B}]-\mathrm{E}[Y_d^A \mid G_d=\mathrm{A}]}\\&=\mathrm{E}[Y_d^B-Y_d^A \mid G_d=\mathrm{B}]\\&=\mathrm{E}[Y_d^B-Yd^A]\\&\equiv\tau^{\mathrm{ATE}}\end{aligned} \tag{5}$$

2.3 Procedure for evaluating yield management effects using A/B test

Figure 1 illustrates a framework for evaluating yield management effects using an A/B test. The example uses the same setting as in Section 1, with only two types of tickets: standard and discount tickets. It also follows the A/B patterns described in Section 2.1, where the sales slots are set uniformly at 10 and 20 slots, respectively, for all trains.

In Fig. 1, the daily average demand is used as the outcome variable. For standard tickets, the difference obtained by subtracting the value for Group A from that for Group B, hereafter referred to as the B-A difference, is −45. This indicates that expanding the sales slots from pattern A to pattern B results in 45 passengers switching from standard tickets to discount tickets. For discount tickets, by contrast, the B-A difference is 60. After accounting for these 45 passengers, the remaining difference is 15, indicating that expanding the sales slots induces demand from 15 additional passengers. The resulting change in farebox revenue depends on the balance between the revenue loss from down-selling by 45 passengers and the revenue gain from the induced demand of 15 passengers.

The procedure for evaluating yield management effects using the A/B test is outlined below. First, A/B patterns for the sales slots are defined based on the characteristics of the operating railway line, and the test period is randomly assigned to A/B groups at the daily level (Step 1). After the A/B test is conducted on a revenue service line, verification is made to check whether the A/B group assignment was conducted randomly as intended (Step 2). After confirming that randomization was successfully performed, the effects of yield management are evaluated at the daily level, as illustrated in Fig. 1 (Step 3). Finally, the effects of yield management are evaluated at the time-of-day level (Step 4). The details of each step are described in Section 3.

3. Application and analysis of A/B test on an operating line

This section first describes the conditions and procedure for conducting the A/B test. It then presents the analysis results of the A/B test conducted on a line in revenue service. The test followed the four-step procedure for evaluating yield management effects described in Section 2.3. As described in Section 1, this analysis verifies the practical applicability of the proposed method and demonstrates that the method can quantitatively identify how changes in the sales slots for discount tickets affect farebox revenue.

3.1 Conditions and procedure for implementing A/B test

The A/B test was conducted on the origin-destination (OD) pair of X Station and Y Station in both directions. X Station and Y Station are the central stations of a designated city and a core city, respectively, both of which are defined under the Local Autonomy Act. The two stations are connected by direct intercity train services with a travel time of about two hours. These intercity train services operate according to a timetable determined during the annual timetable revision. Therefore, unless disruptions such as delays or rolling stock failures occur, the number of regular trains, departure and arrival times, train formation, the composition of reserved-seat and

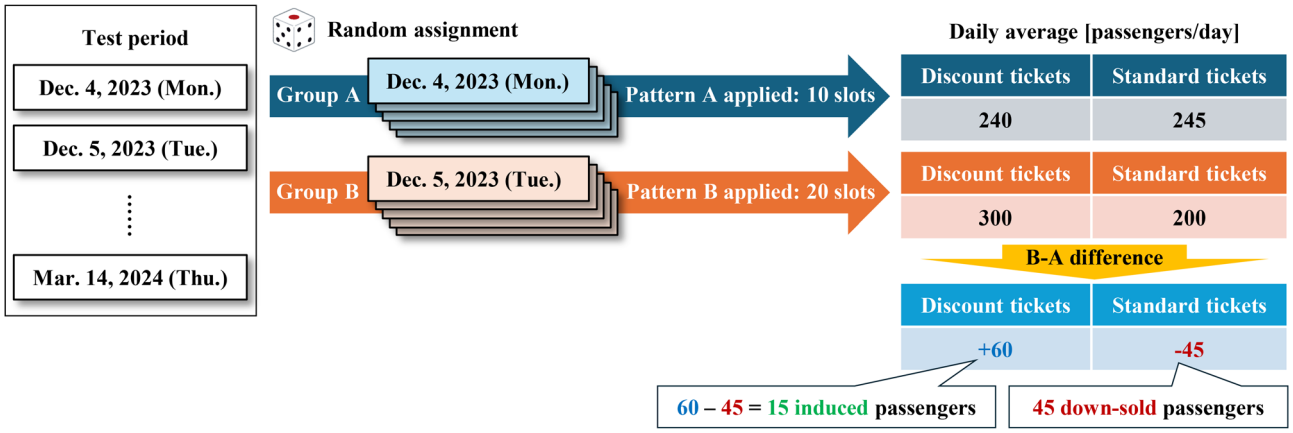


Fig. 1 Framework for evaluating yield management effects using A/B test

non-reserved-seat cars, and seating capacity are the same each day. Although some trains have different formations depending on the day type, such as weekdays or holidays, these conditions are the same within each analysis unit because the analysis is conducted separately by day type. Consequently, the operating conditions are identical at the daily level, even though they may differ to some extent across trains within a day. Therefore, the no-interference assumption is considered to hold. In addition to these intercity train services, the main alternative modes of travel for this OD pair include express buses and private cars. In this paper, the flow from X Station to Y Station is referred to as outbound, and the flow from Y Station to X Station is referred to as inbound.

Tickets for this OD pair can broadly be divided into two categories: full-fare tickets purchased at ticket counters or ticket vending machines, and tickets purchased online. Online tickets are further divided into standard and discount tickets. In the following discussion, standard online tickets are referred to as standard tickets. Sales slots are set only for discount tickets, and the A/B test controls the number of sales slots for these tickets. Discount tickets are priced about 20% lower than standard tickets, whereas full-fare tickets, on a reserved-seat basis, are priced at about twice the level of standard tickets. Thus, full-fare tickets differ from the other two ticket types in that they cannot be purchased online and are substantially more expensive. For this reason, the demand affected by this A/B test, which controls the sales slots for discount tickets, is assumed to be limited to standard tickets and discount tickets; demand for full-fare tickets is therefore not examined. Discount tickets are available only for reserved seats, whereas standard tickets are available for both reserved and non-reserved seats. However, since standard tickets have the same price for reserved and non-reserved seats, passengers generally use reserved seats. Therefore, demand for non-reserved seats is expected to be affected only marginally by the A/B test. It should be noted that, when using intercity train services in Japan, passengers pay a base fare and an express surcharge. In this paper, all prices refer to the sum of the base fare and the express surcharge.

The A/B test was conducted from Friday, December 1, 2023, to Sunday, June 30, 2024. Despite the annual timetable revision in March 2024, the number of regular trains remained unchanged. Departure and arrival times were also generally similar before and after the revision.

Although the detailed procedure for implementing the A/B test is described later, the number of sales slots is set relatively low in pattern A and relatively high in pattern B. The expected results are described below, interpreting pattern B as an expansion of sales slots relative to pattern A. First, when the sales slots for discount tickets are expanded, some demand for reserved-seat standard tickets is expected to shift to discount tickets. By contrast, demand for non-reserved-seat standard tickets is expected to remain largely unchanged because such passengers choose non-reserved seats even though reserved and non-reserved seats are priced the same. In addition, if expanding the sales slots for discount tickets stimulates demand for these tickets, total demand is expected to increase. If no additional demand is generated, total demand is expected to remain unchanged. The revenue effect depends on the balance between two factors: the loss from down-selling and the gain from induced demand for discount tickets. Higher revenue under pattern B indicates that expanded sales slots generated sufficient additional demand. Higher revenue under pattern A indicates that reduced sales slots were more effective in preventing excessive discounting. The expected signs of the A/B test results derived from the above discussion are summarized in Table 2.

Before implementing the A/B test, the railway operator determined the number of sales slots by classifying each date into several day types according to demand characteristics: weekdays, pre-holiday days, Saturdays, Sundays, medium-length holiday periods, and super-peak periods. Weekdays refer to non-holiday Mondays to Thursdays, while pre-holiday days mainly refer to Fridays. Medium-length holiday periods mainly refer to three-day weekends. The super-peak periods considered here are the year-end and New Year holidays, the Golden Week holiday period from late April to early May, and the Obon holiday period in mid-August. This study also evaluates the A/B test results based on this day-type classification. However, medium-length holiday periods and super-peak periods were excluded because demand was expected to be unusually high during these periods. Therefore, the A/B test was evaluated for the following four day types: weekdays, pre-holiday days, Saturdays, and Sundays. In addition, some Saturdays and Sundays were affected by large-scale events held near X Station, such as concerts by popular artists. On these dates, the railway operator applied pattern A regardless of the A/B group assignment. These dates were therefore removed from the evaluation sample for the A/B test. Table 3

Table 2 Expected signs of A/B test results

Sales slots for discount tickets		Difference in demand (B - A)				Difference in farebox revenue (B - A)
A	B	Discount tickets	Reserved-seat standard tickets	Non-reserved-seat standard tickets	Total	Total
Low	High	+	−	0 or −	+ or 0	+ or 0 or −

Table 3 Conditions and number of evaluation days for A/B test

Day type	Number of sales slots per train		Number of evaluation days	
	A	B	A	B
Weekdays	10*1	20	55	51
Pre-holiday days	10*1	20	15	16
Saturdays	20*2	30	10	8
Sundays	20*2	30	11	8

*1: 15 slots for some extra trains only
*2: 30 slots for some extra trains only

shows the number of sales slots and the number of evaluation days by A/B pattern.

This study uses ticketing records held by the railway operator. Table 4 presents descriptive statistics organized by day type, direction, and ticket type. Table 4 shows that the largest number of evaluation days was available for weekdays, with 106 days. By contrast, the numbers of evaluation days for Saturdays and Sundays were limited to 18 and 19 days, respectively, partly because large-scale events were more likely to be held on these days. Nevertheless, these sample sizes were larger than the three days examined in the previous study, which corresponded to a three-day weekend.

3.2 Verification of randomization

It is necessary to verify whether the A/B group assignment was random in practice. For example, at X Station, which is located in a large city, special demand may arise from events or other factors that are difficult for the analyst to fully identify. If such exceptional dates are concentrated in either Group A or Group B, the effect of yield management through sales slot control cannot be properly isolated. Therefore, to verify that randomization was properly conducted, this study examines covariates that are not affected by sales slot control for discount tickets between X Station and Y Station during the A/B test period. The analysis shows that the mean values of these covariates do not differ significantly between the A/B groups.

Important covariates include operating conditions, such as travel time and the number of trains. As described in Section 3.1, these operating conditions were the same each day. Other possible covariates include calendar arrangements, weather conditions such as temperature and precipitation, and local events held along the line. However, these factors are extensive and diverse, making it impractical to examine all of them comprehensively. In addition, for local events, it is impossible to fully identify all individual cases that may affect demand, such as large private group trips. To address these limitations, this study uses OD demand on the same railway line for which sales slots for discount tickets were not controlled as a covariate. This is because, if demand between X Station and Y Station is affected by observable or unobservable factors, similar effects are also expected to appear in the demand for other OD pairs on the same railway line.

From the above discussion, this study used demand for other OD pairs on the same railway line as the X Station–Y Station pair as covariates to verify randomization. Specifically, the covariates were the demand for the three OD pairs with the highest demand among those for which no discount tickets were available. These three OD pairs correspond to six directional OD flows in total. For these covariates, two-sample t-tests were conducted. The null hypothesis was that the B-A difference, defined as the mean value for Group B minus the mean value for Group A, was zero, and the tests were conducted as two-sided tests. Table 5 shows the results. Overall, the results indicate that statistically significant differences between the A/B groups were rarely observed. Although significant differences were observed at the 10% significance level for some OD flows on Saturdays and Sundays for non-reserved-seat standard tickets, no significant differences were observed even at the 10% significance level for all other tickets and OD flows. At the 5% significance level, no significant differences were observed for any ticket or OD flow. These results suggest that the A/B group assignment was reasonably well randomized.

3.3 Daily-level analysis of A/B test

Table 6 shows the results of the daily-level A/B test for the X Station–Y Station pair (Step 3). As shown in Table 2, the B-A difference in farebox revenue may be either positive or negative, whereas the B-A difference in demand is expected to be either positive or negative depending on the ticket category. Therefore, the two-sample t-tests in Table 6 are reported as two-sided tests for farebox revenue and as one-sided tests for demand.

First, the overall trends are interpreted regardless of day type. For discount tickets, the B-A difference was estimated to be significantly positive at the 1% significance level in almost all categories. By contrast, for reserved-seat standard tickets, the difference between the A/B groups was estimated to be significantly negative at the 1% significance level in all categories. These results provide statistically significant evidence supporting the hypothesis that expanding sales slots led to down-selling from reserved-seat standard tickets to discount tickets. For non-reserved-seat standard tickets, no significant B-A differences were observed even at the 10% significance level, except in a few cases. This result is consistent with the hypothesis that sales slot control for discount tickets does not sub-

Table 4 Descriptive statistics

Day type	Direction	Demand [passengers/day]								Number of evaluation days		Number of excluded days*	
		Discount tickets		Reserved-seat standard tickets		Non-reserved-seat standard tickets		Total					
		Mean	SD	Mean	SD	Mean	SD	Mean	SD				
Weekdays	Outbound	289.4	72.4	219.9	68.7	61.1	20.1	570.4	116.9	106	174	1	16
	Inbound	281.1	59.8	203.9	56.5	64.3	18.6	549.3	96.8				
Pre-holiday days	Outbound	349.7	98.8	258.7	75.2	96.2	30.9	704.6	111.8	31		1	
	Inbound	380.9	95.2	306.2	77.0	113.1	27.3	800.2	94.4				
Saturdays	Outbound	516.3	55.2	251.1	58.7	76.6	18.6	844.1	96.1	18		7	
	Inbound	510.7	47.4	259.9	60.0	128.7	41.2	899.3	103.5				
Sundays	Outbound	525.3	72.2	304.4	70.5	124.6	41.8	954.3	102.3	19		7	
	Inbound	556.4	61.9	244.4	49.1	99.5	27.2	900.3	97.2				

*: Excluded days due to large-scale events are shown. In addition, 22 days were excluded because they were medium-length holiday periods or super-peak periods.

Table 5 Results of randomization verification

Day type	OD	Direction	Demand [passengers/day]					
			Reserved-seat standard tickets			Non-reserved-seat standard tickets		
			A	B	B-A difference	A	B	B-A difference
Weekdays	X Station–I Station	Outbound	115.9	110.0	−5.9	37.8	38.4	0.6
		Inbound	100.3	96.5	−3.8	42.2	44.0	1.8
	X Station–J Station	Outbound	70.3	69.9	−0.5	15.2	16.0	0.8
		Inbound	63.0	60.8	−2.3	17.3	18.2	0.8
	X Station–K Station	Outbound	64.0	66.4	2.4	18.6	17.5	−1.1
		Inbound	59.3	58.2	−1.1	18.2	18.6	0.5
Pre-holiday days	X Station–I Station	Outbound	136.1	150.8	14.7	48.9	51.1	2.1
		Inbound	119.5	125.4	5.9	59.6	65.3	5.7
	X Station–J Station	Outbound	84.3	84.1	−0.3	18.4	21.7	3.3
		Inbound	82.0	77.3	−4.8	25.3	29.1	3.9
	X Station–K Station	Outbound	83.1	83.1	−0.1	22.6	21.8	−0.8
		Inbound	71.4	70.7	−0.7	26.4	26.2	−0.2
Saturdays	X Station–I Station	Outbound	200.7	195.9	−4.8	65.5	57.8	−7.8
		Inbound	149.4	145.0	−4.4	84.3	78.0	−6.3
	X Station–J Station	Outbound	90.3	79.0	−11.3	18.3	15.9	−2.4
		Inbound	78.4	77.8	−0.7	36.3	28.1	−8.2*
	X Station–K Station	Outbound	92.5	93.9	1.4	19.2	17.3	−2.0
		Inbound	81.4	81.1	−0.3	33.6	27.0	−6.6
Sundays	X Station–I Station	Outbound	169.3	153.4	−15.9	52.8	42.9	−9.9*
		Inbound	139.5	131.0	−8.5	78.2	64.3	−13.9*
	X Station–J Station	Outbound	78.5	76.0	−2.5	14.5	14.8	0.3
		Inbound	64.7	66.1	1.4	25.9	20.1	−5.8*
	X Station–K Station	Outbound	79.6	76.9	−2.8	16.4	16.6	0.3
		Inbound	73.4	77.1	3.8	24.2	21.6	−2.6

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.
Numbers in red indicate negative values.

Table 6 Results of daily-level A/B test

Day type	Direction	Demand [passengers/day]									Farebox revenue [JPY/day]	
		Discount tickets			Reserved-seat standard tickets			Non-reserved-seat standard tickets			Total	Total
		A	B	B-A difference	A	B	B-A difference	A	B	B-A difference	B-A difference	B-A difference
Weekdays	Outbound	249.6	332.3	82.7***	250.6	186.8	−63.9***	62.3	59.8	−2.5	16.3	+
	Inbound	243.7	321.5	77.8***	229.7	176.1	−53.6***	66.7	61.7	−4.9*	19.3	+
Pre-holiday days	Outbound	256.7	437.0	180.3***	310.9	209.8	−101.1***	90.4	101.6	11.2	90.4**	+
	Inbound	287.7	468.3	180.6***	362.2	253.8	−108.5***	107.9	118.0	10.1	82.3***	+
Saturdays	Outbound	495.9	541.9	46.0**	288.2	204.8	−83.5***	81.9	70.0	−11.9	−49.4	−
	Inbound	488.2	538.9	50.7***	302.5	206.6	−95.9***	140.4	114.1	−26.3	−71.5	−*
Sundays	Outbound	478.1	590.3	112.2***	347.7	244.8	−103.0***	127.5	120.6	−6.8	2.4	−
	Inbound	522.0	603.6	81.6***	270.5	208.5	−62.0***	97.1	102.9	5.8	25.4	+

***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.
Numbers in red indicate negative values.
For confidentiality reasons, farebox revenue is shown only by sign.

stantially affect demand for non-reserved seats.

The results are then interpreted by day type. First, for weekdays and Sundays, the B-A difference in total demand was positive in both cases. The positive sign of the B-A difference alone suggests that expanding the sales slots stimulated demand. However, the difference was not statistically significant in either case, not even at the 10% significance level. Secondly, for pre-holiday days, the B-A difference in total demand was positive and statistically significant at the 1% and 5% significance levels. This indicates that expanding the sales slots for discount tickets stimulated demand. By contrast, although the B-A difference in farebox revenue was also positive, it was not statistically significant even at the 10% significance level. Finally, on Saturdays, the B-A differences in total demand and farebox revenue were both negative. In theory, expanding the sales slots for discount tickets is not expected to reduce total demand. Therefore, the negative B-A difference in total demand on Saturdays is considered to be a result that is implausible as an effect of sales-slot expansion and is likely attributable to factors other than sales slots, such as the calendar arrangement or events. As shown in Table 4, although demand on pre-holiday days, Saturdays, and Sundays was higher than on weekdays, the number of evaluation days was lower. Therefore, the results for these day types may be less stable than those for weekdays. Demand may also vary by season and by month. Even within the same month, demand patterns may differ between the beginning and end. However, for day types other than weekdays, there are only about four days of each type per month. Therefore, if days with particularly high demand happen to be concentrated in Group A, such a result may arise.

Daily-level analysis may average out demand patterns that differ by time of day. For example, if down-selling is large in the morning whereas induced demand exceeds down-selling around midday, effects that actually occur may not be detected at the daily level. Therefore, the next section examines changes in farebox revenue throughout the day using a more detailed time-of-day-level analysis.

3.4 Time-of-day level analysis of A/B test

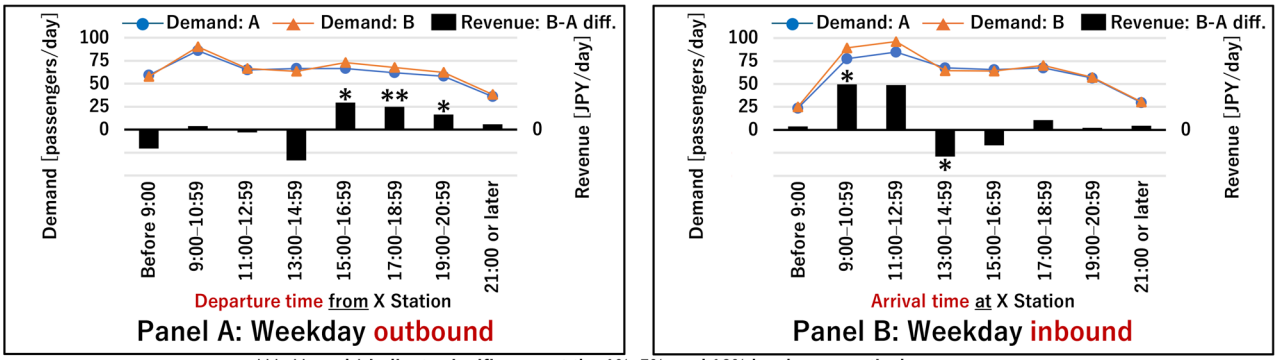
This section examines the A/B test in greater detail by time of day, focusing on weekdays because they provide the largest number of evaluation days and relatively stable daily-level results (Step 4). When the unit of assignment differs from the unit of analysis, statistical testing should use cluster-robust standard errors rather than a two-sample t-test [6, 7]. In this study, statistical significance is tested using cluster-robust standard errors based on parameter estimates from the regression model shown in (6).

$$y_{i,d} = \beta_1 + \Sigma_{t=2}^{T} \beta_t S_t(i) + \tau_1 B(d) + \Sigma_{t=2}^{T} \tau_t S_t(i) \cdot B(d) + \varepsilon_{i,d} \quad (6)$$

Here, subscript i denotes the train, d denotes the date, and t denotes the time-of-day ID, which is a serial number from 1 to T. The variable $y_{i,d}$ represents the KPI of train i on date d. $B(d)$ is a dummy variable equal to 1 when date d belongs to Group B, and $S_t(i)$ is a dummy variable equal to 1 when train i belongs to time-of-day period t. The term $\varepsilon_{i,d}$ is the error term, and β and τ are parameters. Therefore, the significance test for the difference in KPI between the A and B groups in time-of-day period t is conducted by testing the null hypothesis that τ_t equals zero.

Figure 2 shows the demand values for the A and B groups, the B-A differences in farebox revenue, and the results of two-sided statistical significance tests when time-of-day periods are divided into eight categories: before 9:00, 9:00–10:59, ..., 19:00–20:59, and 21:00 or later. For weekday outbound trains, Panel A of Fig. 2 shows that the B-A differences in farebox revenue are significantly positive at the 5% or 10% significance levels in three consecutive time-of-day periods: 15:00–16:59, 17:00–18:59, and 19:00–20:59. This indicates that pattern B is recommended for increasing farebox revenue during these periods. In all three periods, demand is also higher under pattern B than under pattern A. Therefore, the increase in farebox revenue during these periods can be attributed to pattern B, under which the revenue gain from induced demand exceeded the revenue loss caused by down-selling. For weekday inbound trains, Panel B of Fig. 2 shows that the B-A difference in farebox revenue is significantly positive at the 10% significance level in the 9:00–10:59 period, whereas it is significantly negative at the 10% significance level in the 13:00–14:59 period. This indicates that patterns B and A are recommended for the 9:00–10:59 and 13:00–14:59 periods, respectively. The demand in these periods is higher under pattern B in the 9:00–10:59 period, whereas it is almost the same under patterns A and B in the 13:00–14:59 period. Therefore, expanding the sales slots in the 9:00–10:59 period appears to have induced enough demand to offset the decrease in revenue per passenger, as observed for weekday outbound trains. By contrast, reducing the sales slots in the 13:00–14:59 period appears to have suppressed down-selling.

These results show that statistically significant differences based on cluster-robust standard errors were observed at the time-of-day level. However, neither total demand nor farebox revenue differed significantly at the daily level for weekdays. This result suggests that time-of-day-specific or train-specific sales slot settings



***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.
For confidentiality reasons, only 0 is shown on the vertical axis for farebox revenue.

Fig. 2 Results of time-of-day level A/B test

can be more effective than a uniform setting throughout the day. Specifically, the analysis shows two distinct mechanisms: in some periods, expanding sales slots increases farebox revenue because the revenue gain from induced demand exceeds the revenue loss caused by down-selling, whereas in other periods, reducing sales slots increases farebox revenue by suppressing down-selling. These findings demonstrate that the proposed A/B test method can quantitatively identify how changes in sales slots for discount tickets affect farebox revenue.

4. Conclusion

This study proposed a method for evaluating the effects of yield management through an A/B test, focusing on sales slot control for discount tickets. The method was then applied to an A/B test on a line in revenue service. The results demonstrated that the method can quantitatively evaluate how increases or decreases in these sales slots affect farebox revenue. The key finding was that the appropriate sales slot setting strategy differs by time of day. The results showed that, in some periods, expanding sales slots increased farebox revenue by inducing demand that outweighed the decrease in revenue per passenger. In other periods, reducing sales slots increased farebox revenue by suppressing down-selling. These findings can provide railway operators with important information to support their decisions when formulating time-of-day-specific yield management strategies.

Future work includes developing adjustment methods that can produce more robust results for Saturdays and Sundays, for which it is difficult to obtain as many evaluation days as for weekdays. One limitation of this study is that, although the A/B test ensured high internal validity, the external validity of the results could not be verified for periods or locations where the experiment was not conducted. This issue should be addressed by continuing the A/B test, extending the test to other railway lines, and updating the analysis accordingly.

Despite these remaining issues, this study makes an academic contribution to yield management research in transport economics. It empirically demonstrates how sales slot control for discount tickets affects demand and farebox revenue, an issue that has not been fully clarified in previous research. This study also makes a practical contribution as the first published case of an A/B test conducted on a Japanese railway. It shows the appropriate assignment of A/B groups by date and an analytical method for cases where A/B groups are assigned by day but analyzed by train. This paper is expected to serve as a practical guide for future A/B tests in the railway sector and other transport fields.

This paper is an English-language version of the *RTRI Report* article [8], which provided a detailed discussion of practical implications based on the content presented in *The Japanese Journal of Transportation Economics* [9].

Acknowledgment

The author is deeply grateful to the railway operator for its generous cooperation, making it possible to conduct the A/B test and complete this paper.

References

[1] McGill, J. I., and Van Ryzin, G. J., "Revenue management: Research overview and prospects," *Transportation Science*, Vol. 33, No. 2, pp. 233-256, 1999.

[2] Sato, K. and Sawaki, K., *Revenue Management*. Kyoritsu Shuppan Co., Ltd., Tokyo, Japan, 2020 (in Japanese).

[3] Nakagawa, S., Suzuki, T., Matsumoto, R., Fukasawa, N., and Ozaki, N., "Yield Management Method Focused on Discount Ticket Sales Quota," *Quarterly Report of RTRI*, Vol. 61, No. 4, pp. 261-266, 2020.

[4] Nakagawa, S., Suzuki, T., Matsumoto, R., Fukasawa, N., and Ozaki, N., "Yield Management Method Focused on Discount Ticket Sales Quota," *RTRI Report*, Vol. 34, No. 2, pp. 11-16, 2020 (in Japanese).

[5] Cabinet Office, Government of Japan, "海外におけるEBPMの事例とエビデンス集作成について," 2020 (in Japanese).

[6] Kohavi, R., Tang, D., and Xu, Y., *Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing*. Cambridge University Press, Cambridge, UK, 2020.

[7] Kawaguchi, K. and Sawada, M., *The Econometrician's Guide to Causal Inference*. Nippon Hyoron Sha, Tokyo, Japan, 2024 (in Japanese).

[8] Matsumoto, R., "Evaluation of Yield Management with Randomized Controlled Trial (A/B Testing)," *RTRI Report*, Vol. 40, No. 2, pp. 52-61, 2026 (in Japanese).

[9] Matsumoto, R., "Evaluation of Yield Management with Railway Discount Ticket Slot Control: Based on A/B testing," *The Japanese Journal of Transportation Economics*, No. 68, pp. 33-40, 2025 (in Japanese).

Author



Ryosuke MATSUMOTO
Senior Researcher, Data Analytics Laboratory, Information and Communication Technology Division
Research Areas: Transportation Economics, Applied Econometrics

Characterization of Instrumented Wheelset Considering Contact Condition between Wheel and Rail

Takayuki TANAKA **Shoya KUNIYUKI**
Vehicle Mechanics Laboratory, Railway Dynamics Division

Instrumented wheelsets are widely used for measuring wheel loads and lateral forces acting between wheel and rail. These forces are fundamental quantities for running safety assessments of railway vehicles. The authors proposed a testing apparatus to clarify the characteristics of strain output of an instrumented wheelset which depend on the contact condition between wheel and rail. This paper describes the results of characteristic tests conducted using this apparatus. The aim of the tests was to confirm two aspects: the accuracy of the estimated force by the instrumented wheelset, and the validity of an assumption generally utilized in the safety assessment against flange climbing.

Key words: *Railway running safety assessment, Instrumented wheelset, Wheel-rail contact, Wheel load and lateral force*

1. Introduction

Evaluating dynamic performance of railway vehicles regarding running safety is crucial for preventing derailments. In running safety assessments, the forces acting between the wheel and rail are commonly evaluated. In particular, the wheel load (P) and the lateral force (Q) are crucial parameters, representing the vertical force acting on the wheel from the rail and the transverse force acting in the sleeper direction, respectively. Among these, the ratio of lateral force to wheel load, known as the derailment coefficient (Q/P), is widely used as an indicator for evaluating running safety.

In running tests conducted to assess running safety and curving performance of railway vehicles in Japan, the wheel load and lateral force are typically measured using an instrumented wheelset. In the instrumented wheelset, the wheel load and lateral force are estimated from strain generated in the wheel web, which is detected by strain gauges attached to the wheel. Figure 1 shows an example of the arrangement of strain gauges on an instrumented wheelset and the bridge circuit for outputting strain signals for the new continuous method [1, 2].

Accurate measurement of wheel load and lateral force using an instrumented wheelset requires prior calibration to determine the relationship between the strain and the corresponding forces. Based on the calibration results from a dedicated testing apparatus, the strain measured during vehicle operation is converted into the wheel load and lateral force. Figure 2 illustrates the wheel load and lateral force calibration apparatus in the Railway Technical Research Institute (RTRI) [3]. The calibration procedure is as follows:

(1) Wheel load: A contact probe is pressed against the reference position on the wheel tread (for example, at 65 mm from the wheel back face in the case of a modified arc tread), and a vertical load corresponding to the wheel load is applied (Fig. 2(a)).

(2) Lateral force: A hydraulic cylinder is inserted between the back

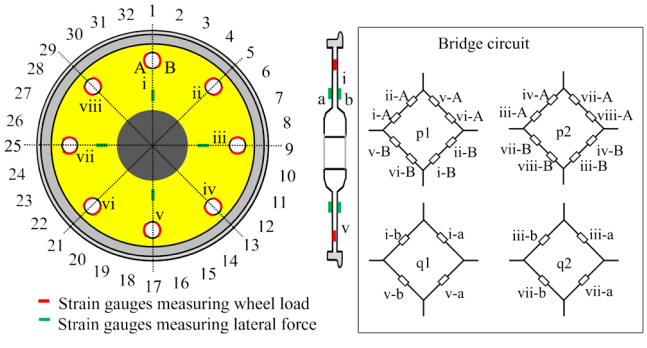
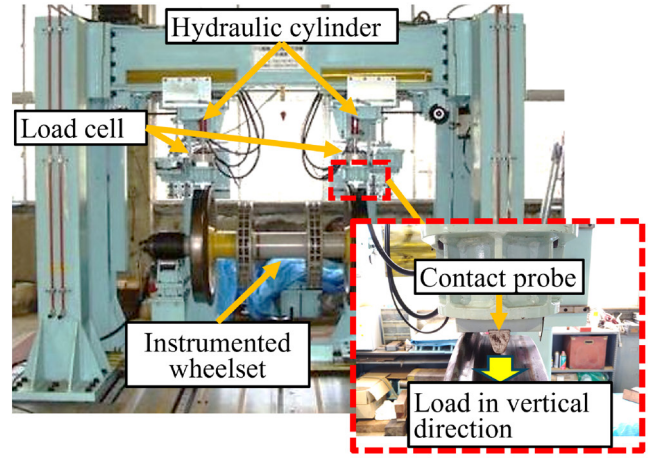
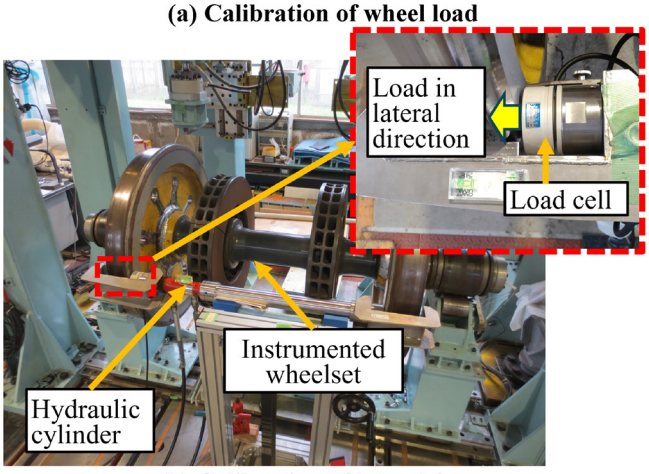


Fig. 1 Example of instrumented wheelset and its bridge circuit



(a) Calibration of wheel load



(b) Calibration of lateral force

Fig. 2 Calibration apparatus of instrumented wheelset

faces of the left and right wheels and extended while pressing load cells against the wheel back faces at the height of the tread reference position, thereby applying a lateral load corresponding to the lateral force (Fig. 2(b)).

Since the strain gauges on the instrumented wheelset are arranged as shown in Fig. 1, the measured strain output varies depending on the circumferential position of the applied load. Therefore, during calibration, measurement points are defined by equally dividing the wheel circumference (a total of 32 points labeled 1 through 32 as shown in Fig. 1), and the strain output is measured at each point according to the applied load. In actual operation, however, the contact condition between the wheel and rail varies significantly depending on factors such as track geometry and the surface profiles of the wheel and rail and cannot be uniquely defined as in the calibration process. In particular, worn wheels and rails lead to two-point contact or complex contact conditions (so-called conformal contact) that cannot be modeled by a Hertzian contact ellipse [4]. However, the differences between the strain outputs under such conditions and those obtained during calibration have not been sufficiently discussed. Furthermore, during flange climbing, which can occur when negotiating sharp curves, the outer wheel comes into contact at the flange face. To analyze this phenomenon, the resultant force is commonly defined as the combination of the tangential force acting along the contact surface and the normal force acting perpendicular to it, as shown in Fig. 3. This resultant force is decomposed into vertical and horizontal components to represent the wheel load and lateral force, respectively [5]. In this approach, the contact surface is defined based on the wheel flange angle, and the relationships among the force components are derived accordingly; however, no experimental studies have verified this relationship.

To address this, the authors have proposed a test method, specifically instrumented wheelset characteristic test, that statically simulates contact conditions occurring during operation and measures the corresponding strain output [3]. This paper outlines the instrumented wheelset characteristic test and presents the evaluation results of the instrumented wheelset output under simulated contact conditions between worn rail and wheel. Furthermore, it discusses the equivalent contact angle inferred from instrumented wheelset outputs when loads corresponding to normal force are applied to the flange face [6].

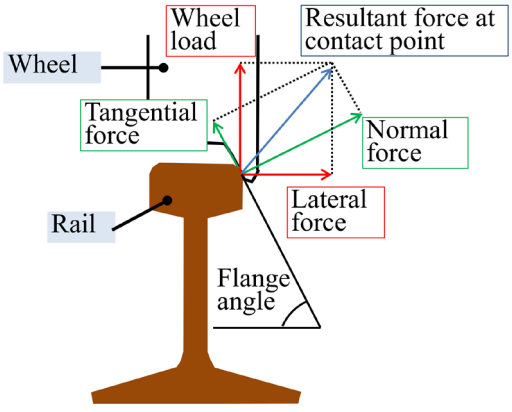


Fig. 3 Forces acting between wheel and rail during flange contact

2. Wheel-Rail Contact Conditions During Vehicle Operation

This chapter summarizes the differences in wheel-rail contact conditions between actual running conditions and calibration conditions. Five key factors essential for describing the contact conditions are selected and outlined in the following sections.

2.1 Lateral Contact Position on Wheel

During operation, the position of the wheelset relative to the track centerline varies depending on the track geometry. For example, during straight-line running, if the lateral position of the wheelset center aligns with the track centerline, the contact position of both wheels will be near the tread center, which is close to the contact position used in calibration. In contrast, when negotiating a curve, the wheelset shifts outward relative to the track centerline. As a result, the contact position of the contact position of the wheel on the outer rail moves toward the flange, while that of the inner rail wheel shifts toward the anti-flange side.

2.2 Number of Contact Points

For wheel and rail profiles commonly used in Japan, the contact typically occurs at a single point when both components are in a new condition. During calibration, loads are applied to the wheel using a single wedge-shaped contact probe, resulting in a single contact point. However, when wheels and rails are worn, or when passing through turnouts, multiple contact points may occur depending on the lateral displacement of the wheelset [7].

2.3 Contact Shape (Contact Area)

For new wheel and rail profiles, the contact area is generally approximated as an elliptical shape. During calibration, the contact area is also approximately elliptical, although with a relatively smaller minor axis. On the other hand, in the case of worn wheel and rail profiles - such as when the throat of the wheel flange comes into contact with the rail shoulder - the contact area may deviate significantly from an elliptical shape, exhibiting a more complex shape [8].

2.4 Load Application Conditions

During actual operation, both the wheel load and lateral force act on the wheel simultaneously. In contrast, during the calibration, loads corresponding to the wheel load and lateral force are applied independently.

2.5 Direction of Load Application

As described in Section 2.4, since the wheel load and lateral force act simultaneously during operation, the resultant force vector varies depending on their ratio and the position where the load is applied. In particular, during flange climbing, the force balance is as shown in Fig. 3. In expressing this relationship of the force balance, the flange angle, which defines the contact surface, becomes an important parameter. However, during calibration, the load corresponding to the wheel load is applied at a tread location where the flange angle can be assumed as approximately 0°. Moreover, the load corresponding to the lateral force is applied to the back face of the wheel. Therefore, during operation, the wheel is subjected to an oblique resultant force, whereas during calibration, the loads are

applied only in pure vertical or horizontal directions.

2.6 Scope of Instrumented Wheelset Characteristic Test

Among the factors described in Section 2.1-2.5, the effect of changes in the wheel-rail contact position, described in Section 2.1, on the strain output of the instrumented wheelset has already been investigated in previous studies [9, 10], and its characteristics are well understood. In particular, when contact occurs near the flange, it has been shown that the influence of bending moments caused by the wheel load can lead to an overestimation of the lateral force. This paper investigates the influence of the differences described in Section 2.2 to 2.5, which have not been sufficiently examined in previous studies on the measurement of wheel load and lateral force.

3. Characteristic Test of Instrumented Wheelset

The instrumented wheelset characteristic test is conducted by adding auxiliary fixtures to the wheel load and lateral force calibration apparatus in RTRI. This setup allows static loads to be applied to the wheel while simulating various contact conditions. Depending on the type of contact interface and loading fixtures, the test is classified into the following three patterns:

Pattern 1: Two-point contact simulation test

This test simulates a two-point contact condition using two contact probes that can simultaneously apply loads to the flange and the tread. The load applied by each contact probe to the wheel can be measured independently.

Pattern 2: Normal load application to the wheel flange face

In this test, a contact probe arranged perpendicular to the straight portion of the wheel flange is used to directly apply and measure the normal force acting on the flange surface.

Pattern 3: Test using a rail segment as a contact body

This test simulates actual wheel-rail contact by using a rail segment as the contact interface. The rail segments can be replaced according to the purpose of the test, including new rail profiles and worn rail profiles.

Figure 4 shows a schematic diagram of the instrumented wheelset characteristic test. Figures 5(a)-(c) illustrate the loading conditions corresponding to Patterns 1 through 3, respectively, showing how the contact probes apply loads to the wheel.

Based on the differences in contact conditions between actual operation and calibration described in Chapter 2, the instrumented wheelset characteristic test was conducted from the following perspectives. The test wheelset used in this study was a conventional instrumented wheelset with a modified arc tread profile. In addition, to investigate how the output varies with the circumferential loading position, loads were applied at five points (Points 1 to 5) on the wheel circumference as shown in Fig. 1. It should be noted that Point 1 was defined as 0°, with subsequent points spaced at 45-degree intervals in the clockwise direction.

(1) Investigation of instrumented wheelset output under two-point and conformal contact conditions

In Pattern 1, the two-point contact condition between the wheel and rail is statically simulated. The proportion of load shared by each contact point (load distribution ratio) is adjusted, defining two different load distribution conditions. Additionally, Pattern 3 is used to simulate conformal contact conditions by employing contact probes that reproduce the shape of worn rail. The tests using Pattern 1 correspond to the differences described in Section 2.2, while those using Pattern 3 correspond to Section 2.3; together, they aim to verify the differences described in Section 2.4. The strain outputs obtained from the instrumented wheelset are converted into equivalent wheel load and lateral force using calibration coefficients determined in advance. Hereafter, these are variables referred to as the output wheel load and output lateral force, respectively, and collectively as the output loads. The applied loads corresponding to wheel load and lateral force in each test are measured using load cells installed in the apparatus (hereafter referred to as input loads). If the output loads are equal to the input loads, it can be concluded that the strain output of the instrumented wheelset is independent of the contact condition.

(2) Investigation of the relationship between normal force and

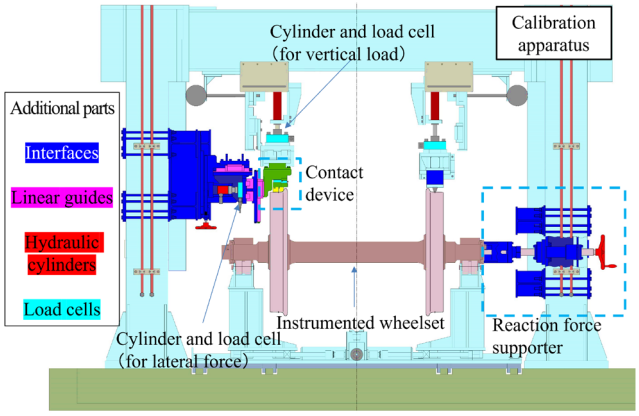
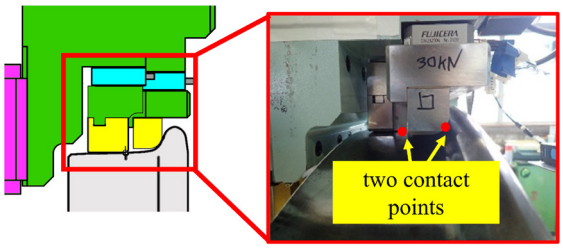
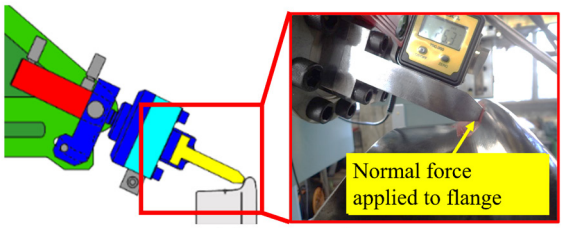


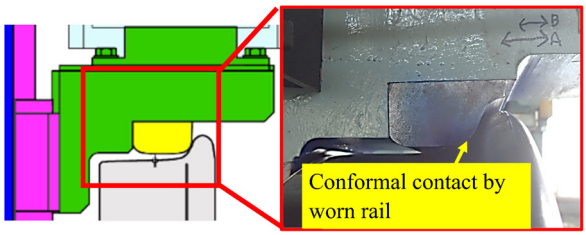
Fig. 4 A schematic diagram of the instrumented wheelset characteristic test



(a) Pattern 1



(b) Pattern 2



(c) Pattern 3

Fig. 5 Loading conditions in the test

wheel load/lateral force outputs

In Pattern 2, a load perpendicular to the straight portion of the wheel flange (namely, a normal force acting on the flange surface) is applied. This test aims to verify the difference described in Section 2.5. Under static loading conditions without wheel rotation, the tangential force component shown in Fig. 3 does not occur. Therefore, in this characteristic test, the vertical and horizontal components of the applied normal force N correspond to the wheel load and lateral force, respectively. That is, denoting the contact angle as α, the following relationships are obtained:

$$P = N\cos\alpha \quad (1)$$

$$Q = N\sin\alpha \quad (2)$$

When the wheel contacts the rail at the flange, the contact angle is generally assumed to be equal to the flange angle. Under this assumption, letting the flange angle be α' (65° for the wheel used in this study), the following relationship holds:

$$Q/P = \tan\alpha' \quad (3)$$

In this paper, the validity of (3) is examined based on the results of the characteristic tests of the instrumented wheelset.

4. Test Results

4.1 Investigation of Instrumented Wheelset Output under Different Contact Conditions

(1) Two-point Contact Condition

In Pattern 1, a vertical load of 30 kN and a lateral load of 20 kN were applied to the instrumented wheelset. It should be noted that the input loads in each direction were controlled within ±0.1 kN of the target values. The two contact points were visualized by applying paint to the wheel surface and pressing the contact probes against it, as shown in Fig. 6. The two contact points were adjusted to be located at the center of the tread and at the straight portion of the flange at distances of 65 mm (tread side) and 29 mm (flange side) from the wheel back face, respectively. For each contact point, two load-sharing conditions for wheel load and lateral force (Conditions I and II) were set, as shown in Table 1. The load center of the wheel load applied at the two contact points is defined as the point where the moment due to wheel load becomes zero. Based on the load distribution ratio and the loading positions, this center was determined to be 50.3 mm from the wheel back face for Condition I and 57.2 mm for Condition II.

The test results are shown in Fig. 7. In this figure, the input loads are compared with the output loads which were obtained by converting the strain outputs from the instrumented wheelset into equivalent wheel load and lateral force using calibration coefficients. For the lateral force, Fig. 7 also shows the corrected values obtained by applying the sensitivity variation formula with respect to changes in the lateral position of the wheel load, as reported in Ref. [10]. As a result, for the wheel load, the output wheel load tended to be smaller than the input load under both conditions. The average difference was approximately 0.7 kN (about 2.3%), although one measurement point showed a difference roughly twice as large (33.75° in Condition II). In contrast, the output lateral force tended to be larger than the input load, with differences of approximately 0.8 kN (about 4.0%) and 0.4 kN (about 2.0%) for Conditions I and II, respectively. One measurement point also showed a relatively large deviation (45° in Condition II, with a difference of 1.2 kN). When correction was applied to the output lateral force, the difference for all points under both conditions remained within approximately 0.5 kN (about 2.5%).

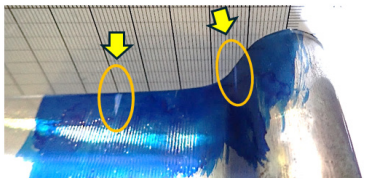
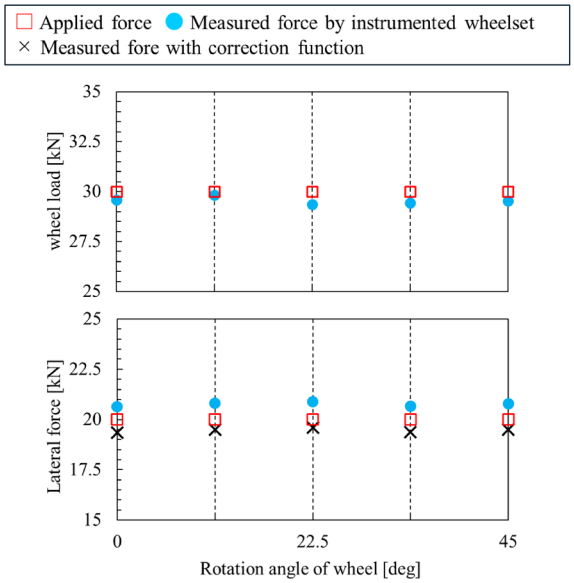


Fig. 6 An example of two-point contact

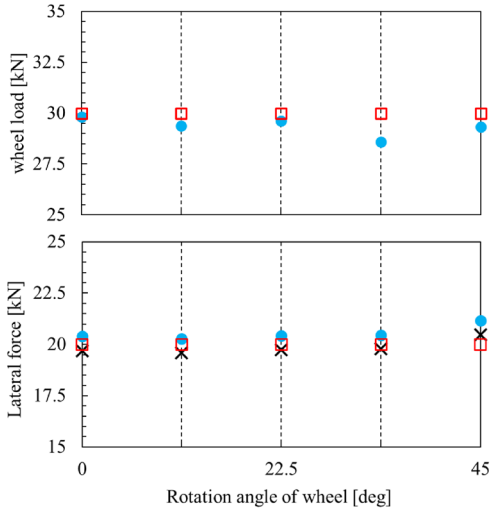
Table 1 Two load conditions in pattern 1

Condition	Wheel load		Lateral force	
	Tread	Flange	Tread	Flange
I	18	12	2.5	17.5
II	24	6	6	14

(Unit: kN)



(a) Condition I



(b) Condition II

Fig. 7 Testing result of Pattern 1

(2) Conformal Contact Condition

In Pattern 3, vertical and lateral loads of 60 kN and 10 kN respectively, were applied to the instrumented wheelset. In this test, two types of contact conditions were simulated: a "full-contact shape" and an "actual worn shape". The contact probes were fabricated by machining JIS 60 kg rail sections: one type was designed to produce wide-area contact with the wheel, while the other replicated worn rail shapes commonly observed in the field. These shapes are shown in Fig. 8. The contact probes were pressed against the throat region of the wheel flange. The contact areas were measured by applying red lead paste to the contact probes and transferring it to the wheel surface; examples are shown in Fig. 9. In the full-contact case, the contact area extended from the flange to the throat region, whereas in the actual worn case, two contact areas were observed: an elliptical contact on the flange and a non-elliptical contact on the tread. The measurement results for wheel load and lateral force are shown in Fig. 10. In comparison with Pattern 1, the difference between input and output wheel load was within approximately 0.6 kN (about 1%), and the difference for lateral force was within approximately 0.5 kN (about 5%). However, for lateral force, since the contact region had spatial extent and the load center of the equivalent wheel load could not be defined, no correction based on lateral load position was applied in this analysis. The measured lateral force tended to overestimate the input load; however, applying such a correction is expected to further reduce the observed discrepancy.

4.2 Relationship between Normal Force and Output Wheel Load/Lateral Force

In Pattern 2, a 50 kN load corresponding to the normal force acting on the flange was applied near the center of the straight portion of the wheel flange (27 mm from the wheel back face). The measurement results of wheel load and lateral force are shown in Fig. 11. In this figure, the vertical (wheel load equivalent) and hori-

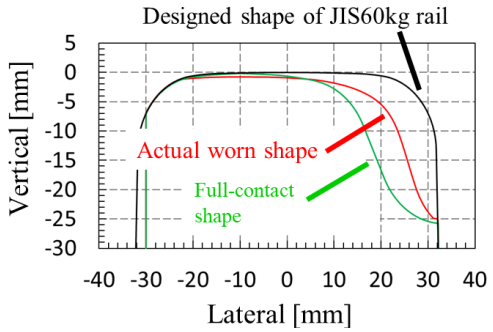


Fig. 8 Shapes of contact probes in Pattern 3

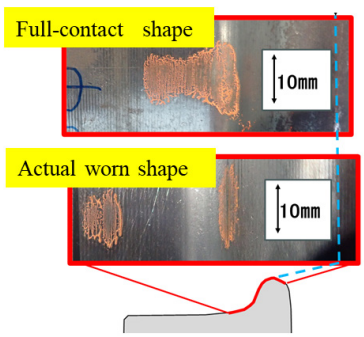


Fig. 9 Examples of contact area measured in Pattern 3

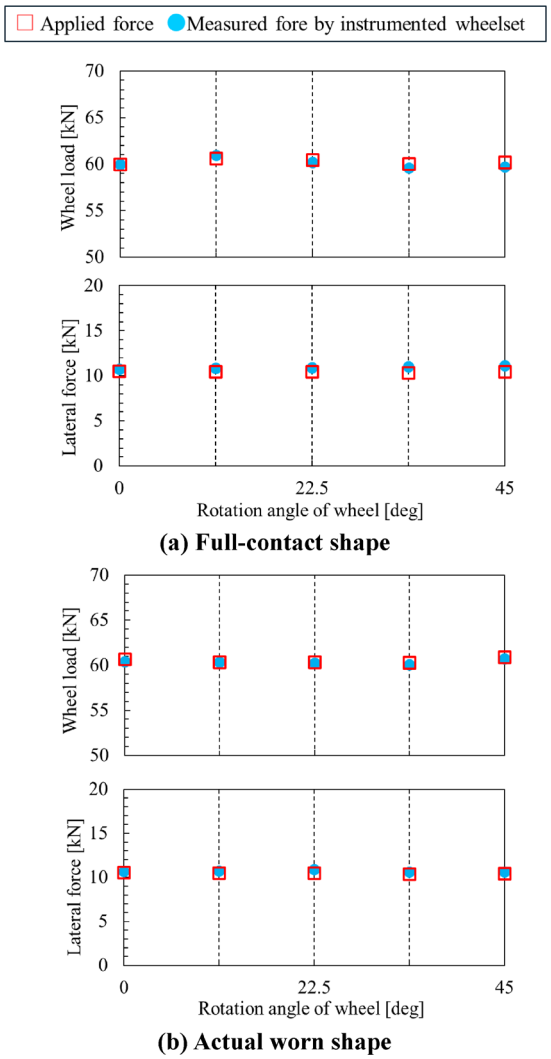


Fig. 10 Testing result of Pattern 3

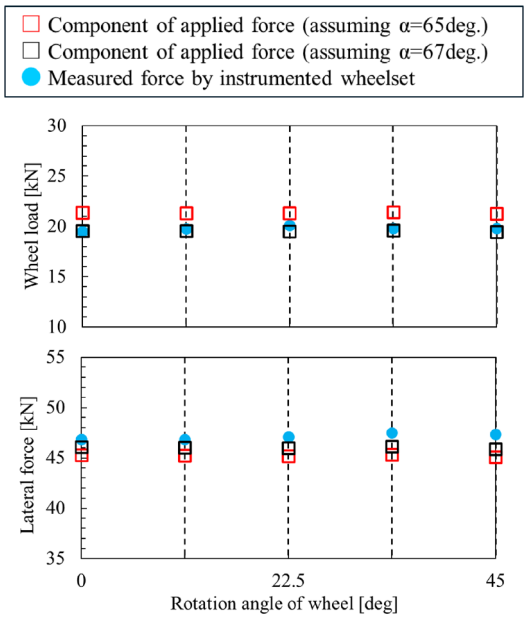


Fig. 11 Testing result of Pattern 2 and estimation of components of applied force

zontal (lateral force equivalent) components corresponding to the input load are shown for assumed contact angles of 65° (equal to the design flange angle of the modified arc tread) and 67°. For the lateral force measurements, corrections were applied as in Pattern 1, taking into account the sensitivity variation caused by the lateral position of the wheel load. By calculating the force components while varying the assumed contact angle around 65°, it was found that assuming a contact angle of 67° yielded values closest to the wheel load and lateral force estimated from the instrumented wheelset measurements.

5. Discussion

Based on the experimental results described in the previous chapter, the following observations can be made.

(1) Under the simulated two-point contact and conformal contact conditions investigated in this study, the loads estimated from the instrumented wheelset output tended to be slightly smaller than the input load for the wheel load, but slightly larger for the lateral force. The differences were within a few percent in each case. These results indicate that the effect of contact conditions on the instrumented wheelset output is relatively small. Furthermore, considering the relationship between the input and output loads, when the derailment coefficient (defined as the ratio of lateral force to wheel load) is used as an index for running safety evaluation, this tendency results in a larger estimated derailment coefficient than the actual value. Therefore, running safety evaluation based on instrumented wheelset measurements can be regarded as conservative, providing a safety margin. In addition, for the lateral force, applying a correction for the effect of the lateral position of the wheel load was confirmed to further reduce the discrepancy. This result demonstrates that such corrections effectively improve the accuracy of lateral force measurements.

(2) When investigating the relationship between the normal input load and the output forces, assuming a contact angle of 67° yielded results closer to the instrumented wheelset output. This is more accurate than the results obtained using the design flange angle of 65°. During the test, no slippage was observed at the contact surface, and the posture of the contact probe was measured to be $65 \pm 0.1°$. Based on these observations, this discrepancy can be attributed to two reasons. First, the deformation of the wheel and axle under static loading alters the effective contact angle. Second, small surface irregularities at the contact interface may also exert an influence. Such phenomena may also occur during actual vehicle operation, potentially affecting how the flange angle is treated in running safety evaluations for flange climbing. Further investigation will be conducted in future work to clarify this issue.

6. Conclusions

This study investigated the output characteristics of an instrumented wheelset, which is widely used for measuring wheel-rail interaction forces, under the various contact conditions that may occur during actual operation. The differences in contact conditions between calibration and running were summarized, and an instrumented wheelset characteristic test was proposed. The outline of the test and the results are as follows.

(1) Characteristic tests of the instrumented wheelset were conducted using a dedicated test apparatus, reproducing two-point contact and conformal contact conditions that may occur under worn wheel-rail interfaces. Under the conditions examined in this study, the loads estimated by the instrumented wheelset were slightly smaller than the input loads for the wheel load, but slightly larger for the lateral force. In each case, differences were within a few percent. These results indicate that the influence of such variations in contact conditions on the output of the instrumented wheelset is relatively small. Furthermore, considering this tendency in the relationship between input and output loads, the derailment coefficient (defined as the ratio of lateral force to wheel load) obtained from instrumented wheelset measurements becomes larger than the actual value when contact conditions vary. Therefore, running safety evaluations based on the derailment coefficient can be considered conservative, providing a safety margin.

(2) An investigation was also conducted to examine whether the wheel load and lateral force, serving as components of the normal force, are resolved based on the flange angle, as generally assumed in analyses of flange climbing. The results showed that assuming a contact angle of 67° - approximately 2° larger (67°) than the design flange angle of the tested wheelset (65°) - provides better agreement with the measured output loads. Further studies will be performed based on the findings of this study by incorporating structural analysis of the instrumented wheelset to validate the results and examine their generality. These efforts aim to further improve the measurement accuracy of wheel load and lateral force using instrumented wheelsets.

References

[1] Ishida, H., Matsuo, M., Tezuka, K. and Ueki, K., "Method of Measuring Wheel and Rail Contact Forces and Derailment Quotients Continuously. Development of a Measuring Device," *Transactions of the Japan Society of Mechanical Engineers. C*, Vol. 63, No. 614, pp. 3417-3423, 1997 (in Japanese).

[2] Suzuki, M., Totake, T., Hondo, T. and Kuniyuki, S., "Development of Measuring System for Wheel and Rail Contact Forces to be Realized Labor Saving of Operation, Using Commercial Measurement Instrument," *RTRI Report*, Vol. 33, No. 8, pp. 17-22, 2019 (in Japanese).

[3] Tanaka, T., Kuniyuki, S., Doi, H. and Nishiyama Y., "Characteristic test of instrumented wheelset considering contact condition," *The Proceedings of 31st Transportation and Logistics Conference*, 2022 (in Japanese).

[4] Tanaka, T. and Sugiyama, H., "Numerical procedure for non-Hertzian wheel-rail contact model integrated in quasi-steady railway motion solver," *Proceedings of the ASME 2020 International Design Engineering Technical Conferences and Computers and Information in Engineering Conferences*, August 17-19, 2020.

[5] Tanaka, T., Doi, H., Miyamoto, T., and Iijima, H., "A study of lateral-force estimation equation focusing on moment of a train bogie," *Transactions of the Japan Society of Mechanical Engineers*, Vol. 82, No. 842, pp. 16-00054, 2016 (in Japanese).

[6] Tanaka, T. and Kuniyuki, S., "Characteristic test of instrumented wheelset considering contact condition-Report of characteristic test results-," *The Proceedings of 33rd Transportation and Logistics Conference*, 2024 (in Japanese).

[7] Terao, A. and Sugiyama, H., "Development of Numerical Procedure for Multiple Wheel/Rail Contact for Curve and turnout Negotiations of Railroad Vehicles," *The Proceedings of Dynamics and Design conference 2012*, 2012 (in Japanese).

[8] Sugiyama, H., Sekiguchi, T., Matsumura, R., Yamashita, S. and

Suda, Y., "Wheel/rail contact dynamics in turnout negotiations with combined nodal and non-conformal contact approach," *Multibody System Dynamics*, Vol. 27, pp. 55-74, 2012.
[9] Hondo, T., "Estimation method of longitudinal displacement of wheel/rail contact point by using PQ measurement wheelset," *The Proceedings of J-RAIL 2018*, 2018 (in Japanese).
[10] Noguchi, Y., "Method for Measuring Wheel/Rail Contact Positions Using Strain Gauges," *IEEJ Transactions on Industry Applications D*, Vol. 141, No. 3, pp. 241-248, 2021 (in Japanese).

Authors



Takayuki TANAKA, Ph.D.
Senior Researcher, Vehicle Mechanics Laboratory, Railway Dynamics Division
Research Areas: Vehicle Dynamics, Running Safety



Shoya KUNIYUKI, Dr.Eng
Senior Researcher, Vehicle Mechanics Laboratory, Railway Dynamics Division
Research Areas: Vehicle Dynamics, Running Safety

Effects of Operating Conditions on Traction Motor Bearing Life

Daisuke SUZUKI **Ken TAKAHASHI**
Lubricating Materials Laboratory, Material Technology Division

Currently, there is no practical method for determining the life of bearings used in traction motors for railway vehicles. Therefore, as a first step toward establishing such a method, the effects of operating conditions on bearing life were investigated. In this study, the effects on bearing life were investigated using a test rig capable of rotating traction motor bearings under conditions close to those experienced in actual operating conditions by varying bearing temperature, vibration acceleration, and grease mass. The results demonstrated that varying any of these conditions affects bearing life, and that these effects can be quantified within the range of conditions examined in this study.

***Key words*:** *bearing, traction motor, wear, tribology, cage, bearing life*

1. Introduction

Figure 1 shows the primary bearings used in the bogies of railway vehicles. As shown in Fig. 2, axle journal bearings and gear unit bearings fail due to rolling fatigue of the outer ring raceway, while traction motor bearings fail due to cage wear caused by sliding contact between the rolling elements and the cage. These failure mechanisms are fundamentally different. To ensure safe and reliable railway operation, it is necessary to evaluate bearing life and replace bearings before failure occurs. For axle journal bearings and gear unit bearings, bearing life can be calculated using methods standardized in ISO 281:2007 [1] or JIS B 1518:2013 [2], which are based on the rolling fatigue life formula developed by Lundberg and Palmgren [3]. On the other hand, because the failure mode of traction motor bearings differs from rolling fatigue, these standardized methods cannot be applied to calculate an appropriate bearing life. Furthermore, since many factors in addition to applied load and total number of rotations influence bearing life, no practical method for evaluating traction motor bearing life has been established.

To establish a method for evaluating traction motor bearing life, it is necessary to investigate how variations in operating conditions affect bearing life. This study presents the results of an investigation into the effects of bearing emperature, vibration acceleration, and grease mass (initial grease fill amount) on bearing life. These investigations were conducted using a test rig capable of rotating traction motor bearings under conditions close to actual operating conditions [4,5]. The test conditions were selected based on the following considerations. Bearing temperature was selected because bearings reach relatively high temperatures during operation and their temperatures fluctuate significantly due to factors such as heat generated by the traction motor and ambient temperature. Vibration acceleration was selected because vibrations caused by train operation are transmitted to the bearings via the bogie frame and the traction motor. Furthermore, these vibrations vary significantly with wheel and rail conditions and train speed. Grease mass was selected because the traction motor bearings are lubricated with grease despite operating at relatively high rotational speeds, and both excessive and insufficient grease significantly affect bearing life. Furthermore, since the rotational speed and load of the bearings are largely determined by the traction motor specifications, and the variations in specifications between traction motor types are minimal, these parameters were not varied as test conditions in this study.

2. Test method

2.1 Test Bearing

As shown in Fig. 1, deep groove ball bearings and NU-type cylindrical roller bearings are commonly used in traction motors. By using an NU-type cylindrical roller bearing at one end, axial displacement caused by thermal expansion of the shaft is allowed, thereby preventing excessive axial loads on the bearing. This study focuses on an NU214 cylindrical roller bearing that is widely used in traction motors. Figure 3 shows the components of the test bearing, and Table 1 lists its specifications. This bearing consists of an

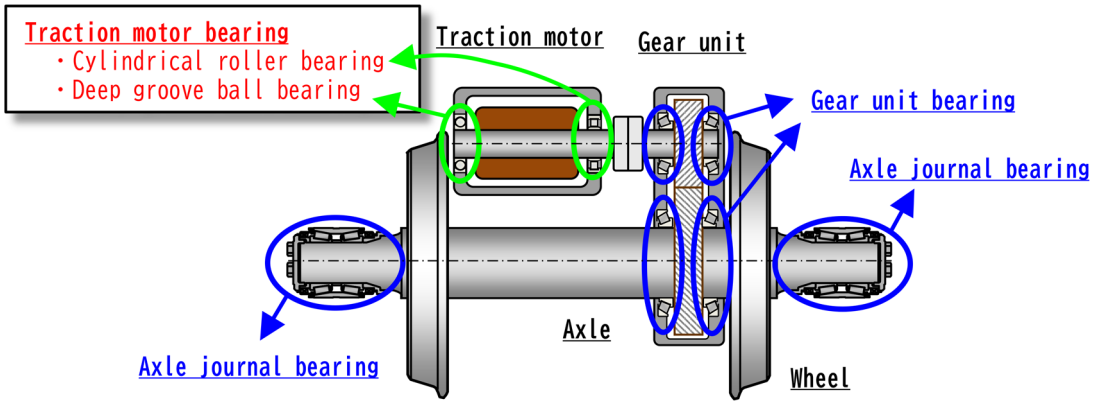


Fig. 1 Primary bearings used in bogie

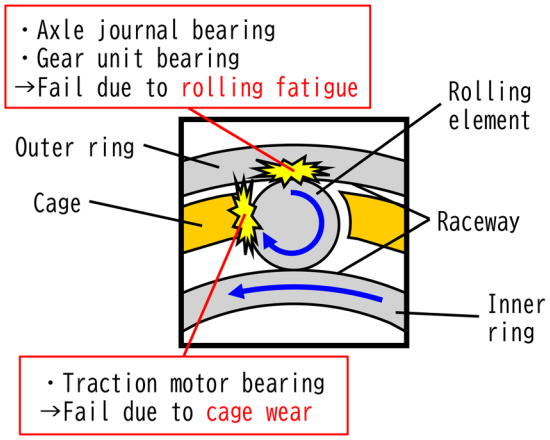


Fig. 2 Causes of bearing failure

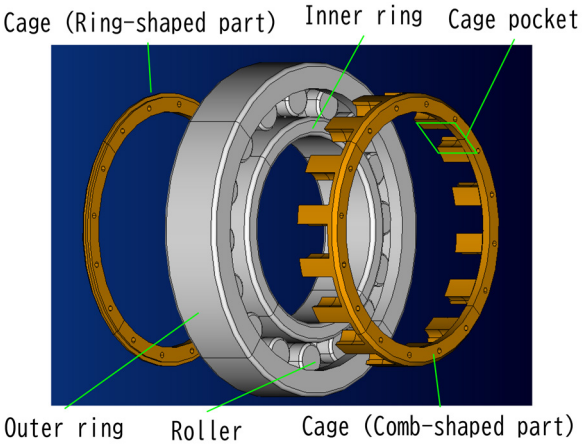


Fig. 3 Components of test bearing

Table 1 Specifications of test bearing

Bearing type		Cylindrical roller bearing
Inner diameter		70 mm
Outer diameter		125 mm
Width		24 mm
Radial clearance		0.090-0.125 mm
Pitch circle diameter of roller		97.5 mm
Number of rollers		16
Roller diameter		13 mm
Length of roller		13 mm
Cage guide		By rolling elements
Material	Race rings	JIS SUJ2
	Rollers	JIS SUJ2
	Cage	JIS CAC301
Basic dynamic load rating		83500 N

inner ring, an outer ring, rollers, and a cage. The cage is assembled by riveting a ring-shaped part to a comb-shaped part. In traction motor bearings, roller-riding cages are primarily used, and this study focuses on this type. Consequently, the cage only comes into contact with the rollers, and sliding contact occurs between the rollers and the cage pockets. Since the inner ring, outer ring, and rollers are made of bearing steel, and the cage is made of high-strength cast brass, most wear at the sliding contact between the rollers and the cage occurs in the softer cage material.

2.2 Test rig

Figure 4 shows a schematic of the test rig used in this study. This test rig enables rotational testing of bearings under conditions that simulate the operating conditions of traction motors, including load, rotational speed, vibration acceleration, and bearing temperature. The test rig is designed to mount the test bearings on both ends of the main shaft, which is driven by a motor via a shaft coupling. To reproduce a load equivalent to the mass of the traction motor rotor, the test load is applied to the test bearing housing via a load-adjusting spring connected to a support bearing housing. Since the test load acts equally on both test bearings, the radial load on each test bearing is half the total test load. Furthermore, to simulate vibrations transmitted from the bogie frame to the traction motor, as well as vibrations generated by the traction motor itself, the entire test section (excluding the motor) can be vibrated vertically using a vibration exciter. Additionally, to simulate the heat generated by the traction motor, the test bearing on the shaft-end side (opposite the motor) can be heated using a heater. The maximum radial load is 1250 N per test bearing, and the maximum rotational speed is 7060 min^{-1} (equivalent to the maximum rotational speed of the traction motor). The maximum vibration acceleration is 100 m/s^2 (amplitude). The test load was measured using a load cell, the rotational speed was measured using an optical tachometer installed on the motor shaft, the test bearing acceleration (RMS value from 0 to 1000 Hz) was measured using a piezoelectric accelerometer installed in the test bearing housing, and the test bearing temperature was measured using a thermocouple installed on the outer diameter surface of the outer ring (hereinafter referred to as the bearing temperature).

2.3 Test conditions

After defining standard conditions, the bearing temperature, vibration acceleration, and grease mass were individually varied to serve as verification conditions, and their effects on bearing life were investigated. Table 2 shows the standard and verification con-

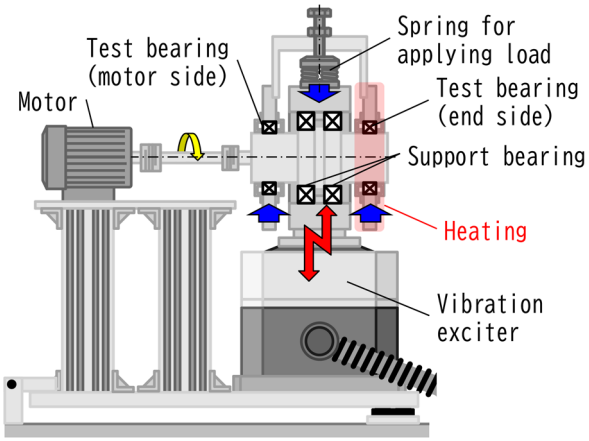


Fig. 4 Schematic of test rig

Table 2 Test conditions (standard and verification conditions)

	Standard conditions	Verification conditions
Maximum rotational speed	6100 min^{-1}	←
Rotation pattern	Forward 110 min → Pause 10 min → Reverse 110 min → Pause 10 min	←
Load	922 N/Brg.	←
Bearing temperature	Approx. 30 ℃	Approx. 35-55 ℃ (in increments of approx. 5 ℃)
Vibration acceleration	50 m/s^2	0, 100 m/s^2
Vibration frequency	40 Hz (sine wave)	←
Lubrication	Lithium complex soap grease	←
Grease mass	1 g	0.5, 2, 4, 8 g

ditions determined on the basis of the following considerations: Higher rotational speeds are expected to increase cage wear per rotation; therefore, the rotational speed was set to 6100 min^{-1}, based on the maximum rotational speed of a traction motor under actual operating conditions. The rate of change in rotational speed during acceleration and deceleration was set to 100 min^{-1}/s. Based on the mass of a standard traction motor rotor, the load was set to 1844 N (922 N per test bearing, 1.1% of the basic dynamic load rating). Since the bearing temperature rise in actual railway vehicles was approximately 20 ℃ above the ambient temperature, the standard condition was set to approximately 30 ℃ (assuming an ambient temperature of 10 ℃). The verification conditions were set within the range of approximately 35-55 ℃ in 5 ℃ increments, assuming a maximum ambient temperature of 35 ℃. Referring to the general rules for the design of bogie frame strength, which specify a vertical motor acceleration of 3-10 G (29.4-98.0 m/s^2) [6], the verification points were set at 0 and 100 m/s^2 (amplitude). The test utilized a 40 Hz sine wave, which is the dominant frequency observed in actual vehicle measurements.

The grease used was a lithium complex grease (base oil: mineral oil), which is widely used in traction motor bearings. Since the specified grease mass of 10 g would require an extended period for bearing seizure to occur, 1 g (1/10 of the specified amount) was established as the standard condition. Additionally, 0.5, 2, 4, and 8 g were set as the verification conditions. Prior to grease filling, the bearings were thoroughly cleaned and degreased, and grease was applied evenly to the inner surface of the cage. Each test was continued until bearing seizure occurred, and the total number of rotations at that point was defined as the bearing life. Bearing seizure was defined as the point at which the temperature rise per minute (hereafter referred to as the bearing temperature rise rate) reached 1 ℃/min or higher. However, due to the configuration of the test rig, a threshold for the bearing temperature rise rate could not be set automatically to terminate the test; therefore, the test was continued until the bearing temperature reached approximately 90 ℃.

3. Results

Table 3 shows the test results obtained under the conditions listed in Table 2 for a total of 14 bearings, numbered sequentially from 1 to 14 (Bearings 1-14). The test results include the specific conditions under which each test was conducted, along with the total number of rotations until bearing seizure occurred. It also indicates whether the results for each bearing were used as verification data for bearing temperature, vibration acceleration, or grease mass. Note that, since the bearing temperature fluctuated during each test, the average bearing temperature for each test is presented.

3.1 Bearing temperature

To evaluate the effect of bearing temperature on bearing life, the bearing temperature verification data from Table 3 are plotted in Fig. 5, with the bearing temperature on the horizontal axis and bearing life on the vertical axis. The bearing numbers corresponding to Table 3 are also indicated in the figure (this notation is adopted throughout the study). Since bearing life decreases exponentially with increasing bearing temperature, a logarithmic scale is employed for the vertical axis. Bearing 4 (bearing temperature: 34.7 ℃), which exhibited the longest bearing life, lasted approximately four times as long as Bearing 8 (bearing temperature: 56.2 ℃), which exhibited the shortest bearing life. Prior studies investigating the relationship between bearing temperature and lubrication life under grease lubrication have reported that lubrication life has an exponential relationship with bearing temperature, and that lubrication life is halved when the bearing temperature rises by approximately 10 ℃ [7]. Although that study focused on deep-groove ball bearings and considered various failure modes other than cage wear, it also addressed the relationship between bearing temperature and bearing life. Therefore, the relationship between bearing temperature and bearing life obtained in this study was approximated using an exponential function, similar to the previous study. The approximation equation relating bearing temperature to bearing life, obtained by the least-squares method, is shown below.

$$L = 10^{a_T T + b_T}$$
$$a_T = -0.0175,\ b_T = 8.89,\ R^2 = 0.884 \quad (1)$$

Here, L represents bearing life, T represents bearing temperature [℃], a_T and b_T are constants, and R^2 is the coefficient of determination. The line shown in Fig. 5 represents the curve given by Equation (1). Based on the approximation equation, the temperature required to halve bearing life is calculated to be 17.2 ℃. Although this is less pronounced than in the previously reported example (approximately 10 ℃), bearing life decreases significantly as the bearing temperature rises. It is likely that when the bearing temperature rises, the oxidation and evaporation of the grease base oil are accelerated, thereby shortening bearing life. Furthermore, it is also likely that as the bearing temperature rises, the oil film thickness at the

Table 3 Test results of bearing life

	Bearing temperature	Vibration acceleration	Grease mass	Bearing life (Total number of rotations)	To verify*
Bearing 1	29.9 ℃	50 m/s^2	1 g	2.0×10^8	T, V, G
Bearing 2	31.8 ℃	50 m/s^2	1 g	2.2×10^8	T, V, G
Bearing 3	32.6 ℃	50 m/s^2	1 g	2.2×10^8	T, V, G
Bearing 4	34.7 ℃	50 m/s^2	1 g	2.2×10^8	T
Bearing 5	42.3 ℃	50 m/s^2	1 g	1.2×10^8	T
Bearing 6	46.8 ℃	50 m/s^2	1 g	1.1×10^8	T
Bearing 7	50.9 ℃	50 m/s^2	1 g	1.2×10^8	T
Bearing 8	56.2 ℃	50 m/s^2	1 g	0.6×10^8	T
Bearing 9	33.0 ℃	0 m/s^2	1 g	5.0×10^8	V
Bearing 10	40.5 ℃	100 m/s^2	1 g	0.7×10^8	V
Bearing 11	39.3 ℃	50 m/s^2	0.5 g	0.7×10^8	G
Bearing 12	31.8 ℃	50 m/s^2	2 g	4.3×10^8	G
Bearing 13	32.1 ℃	50 m/s^2	4 g	5.2×10^8	G
Bearing 14	38.7 ℃	50 m/s^2	8 g	5.6×10^8	G

*T: Bearing temperature, V: Vibration acceleration, G: Grease mass

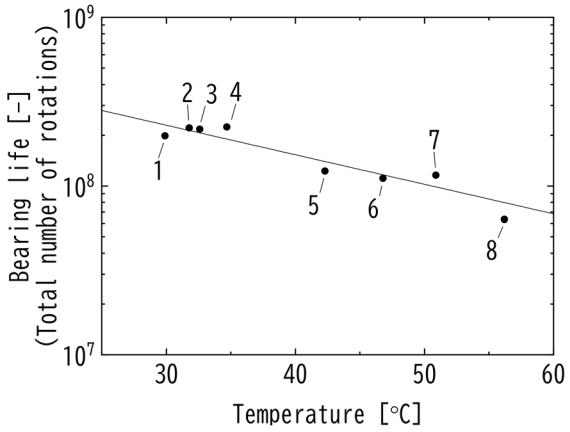


Fig. 5 Test result (effect of bearing temperature on bearing life)

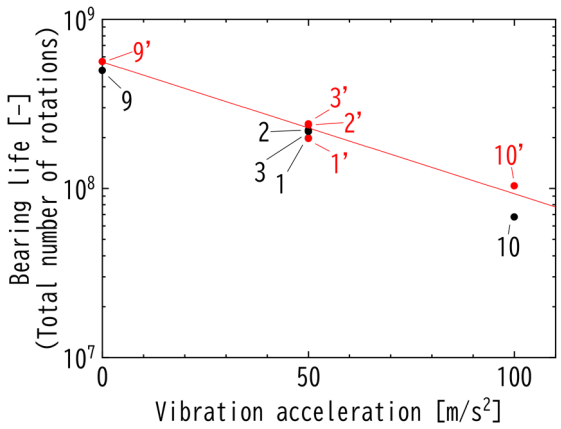


Fig. 6 Test result (effect of vibration acceleration on bearing life)

sliding contact areas decreases, making the oil film more susceptible to rupture caused by the ingress of wear particles.

3.2 Vibration acceleration

To evaluate the effect of vibration acceleration on bearing life, the vibration acceleration verification data shown in Table 3 are plotted as black dots in Fig. 6, with vibration acceleration on the horizontal axis and bearing life on the vertical axis. Since bearing life decreases exponentially with increasing vibration acceleration, a logarithmic scale is employed for the vertical axis. Bearing 9 (vibration acceleration: 0 m/s^2), which exhibited the longest bearing life, lasted approximately seven times as long as Bearing 10 (vibration acceleration: 100 m/s^2), which exhibited the shortest bearing life. However, the bearing temperature of Bearing 10 was about 10 ℃ higher than the target temperature of approximately 30 ℃. This is because, under conditions of high vibration acceleration, the vibration exciter generates heat, which is transferred to the bearing. As shown in Fig. 5, bearing life decreases as bearing temperature rises; therefore, the bearing life of Bearing 10 would be expected to be longer if its temperature were maintained at 30 ℃. Accordingly, using Equation (1), the life of each bearing was converted to an equivalent value at a temperature of 30 ℃. Using Equation (1), the following conversion formula was obtained.

$$L_{30\,℃}=\frac{10^{a_T\times 30+b_T}}{10^{a_T T+b_T}}L=10^{a_T(30-T)}L \qquad (2)$$

Here, $L_{30\,℃}$ represents the bearing life converted to a temperature of 30 ℃. The red points in Fig. 6 represent bearing life converted to a bearing temperature of 30 ℃, and the corresponding bearing numbers are marked with an apostrophe (hereafter, the same notation is applied). The temperature correction results in a more linear

distribution of the data points. Based on this, an exponential function was used to approximate the relationship between vibration acceleration and bearing life (converted to a bearing temperature of 30 ℃). The approximation equation obtained using the least squares method is shown below.

$$L_{30℃}=10^{a_V V+b_V}$$
$$a_V=-0.00780,\ b_V=8.75,\ R^2=0.989 \quad (3)$$

Here, V represents the vibration acceleration [m/s^2], and a_V and b_V are constants. The red line in Fig. 6 represents the curve given by Equation (3). Furthermore, using the approximation equation, the increase in vibration acceleration required to halve bearing life is calculated to be 38.6 m/s^2. The decrease in bearing life as vibration acceleration increases is considered to be due to an increase in the contact force between the rolling elements and the cage, which accelerates cage wear.

3.3 Grease mass

To evaluate the effect of grease mass on bearing life, the grease mass verification data from Table 3 are plotted as black dots in Fig. 7, with grease mass on the horizontal axis and bearing life on the vertical axis. Bearing 14 (grease mass: 8 g), which exhibited the longest bearing life, lasted approximately eight times as long as Bearing 11 (grease mass: 0.5 g), which exhibited the shortest bearing life. However, the temperatures of Bearings 11 and 14 were approximately 10 ℃ higher than the target of approximately 30 ℃. This is because Bearing 11 generated heat due to poor lubrication from the insufficient grease mass, while Bearing 14 generated heat due to agitation loss caused by excess grease mass. The bearing life of these bearings would be expected to be longer if the bearing temperature were 30 ℃. Therefore, using Equation (2), the life of each bearing was converted to an equivalent value at a bearing temperature of 30 ℃. The red points in Fig.7 represent bearing life converted to a bearing temperature of 30 ℃. Based on these results, a logarithmic function was used to approximate the relationship between grease mass and bearing life (converted to a bearing temperature of 30 ℃). The approximation equation obtained using the least squares method is shown below.

$$L_{30℃}=a_G \log_{10} G+b_G$$
$$a_G=5.85\times10^8,\ b_G=2.41\times10^8,\ R^2=0.987 \quad (4)$$

Here, G represents the grease mass [g], and a_G and b_G are constants. The red line in Fig. 7 represents the curve given by Equation (4). Furthermore, using the approximation equation, the increase in bearing life when doubling the grease mass is calculated to be 1.76 × 10^8. Based on this, given an initial grease mass of 8 g, doubling it to 16 g is expected to yield only a modest 23% increase in bearing life. On the other hand, because excessive grease filling raises concerns regarding agitation loss and grease leakage, manufacturers recommend a grease filling ratio (the ratio of grease volume to the bearing's internal volume) of approximately 30% [8,9]. Since the specified grease mass of 10 g for the test bearing corresponds to a filling ratio of approximately 30%, concerns regarding excessive grease are minimal within the examined range. Therefore, the increase in grease mass is considered to have extended bearing life. However, since increasing the grease mass beyond the current level could shorten bearing life due to temperature rises from agitation loss, the current filling mass is considered reasonable.

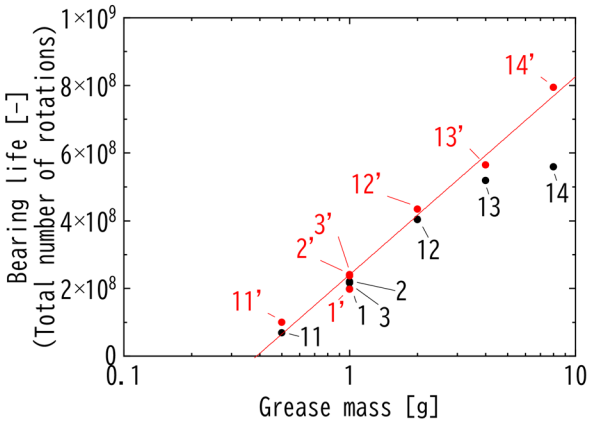


Fig. 7 Test result (effect of grease mass on bearing life)

4. Conclusions

Since there is currently no practical method for determining the life of traction motor bearings, this study investigates, as a first step toward establishing such a method, the effects of varying operating conditions on bearing life. Using a test rig that enables rotational testing of bearings under conditions that simulate the operating conditions of traction motors, the effects on bearing life were investigated by varying specific operating parameters, namely bearing temperature, vibration acceleration, and grease mass. The main findings are summarized as follows.

(1) An investigation into the effect of bearing temperature on bearing life revealed that bearing life decreases exponentially with increasing bearing temperature. The temperature increase required to halve the bearing life was determined to be 17.2 ℃. This behavior is attributed to accelerated oxidation and evaporation of the grease base oil at higher temperatures, along with a reduction in oil film thickness, which increases susceptibility to wear particle ingress.

(2) An investigation into the effect of vibration acceleration on bearing life revealed that bearing life decreases exponentially with increasing vibration acceleration. The increase in vibration acceleration required to halve the bearing life was determined to be 38.6 m/s^2. This behavior is attributed to increased impact forces between the rolling elements and the cage, which accelerate cage wear.

(3) An investigation into the effect of grease mass on bearing life revealed that increasing the grease mass resulted in a logarithmic increase in bearing life. The increase in bearing life when the grease mass was doubled was calculated to be 1.76 × 10^8. Because the tested grease mass was within approximately 30% of the bearing internal volume, a threshold below which grease leakage and excessive heat generation are avoided, the increase in grease mass is considered to have contributed to the extension of bearing life.

Based on findings (1)-(3), the effects of each operating condition on bearing life were clarified within the experimental scope defined in this study. Future work will focus on evaluating bearing life under combined variations of these operating conditions.

References

[1] International Organization for Standardization, "Rolling bearings - Dynamic load ratings and rating life," ISO 281: 2007, 2007.

[2] Japanese Standards Association “Rolling bearings - Dynamic load ratings and rating life,” JIS B 1518: 2013, 2013 (in Japanese).
[3] Lundberg, G. and Palmgren, A., “Dynamic Capacity of Rolling Bearings,” *Journal of Applied Mechanics*, Vol. 16, No. 2, pp. 165-172, 1949.
[4] Suzuki D. and Takahashi K., “Lubrication life of cylindrical roller bearings for traction motors of railway vehicles (Effects of bearing temperature, vibration acceleration, and grease mass),” *Transactions of the JSME*, Vol. 90, No. 935, DOI: 10.1299/transjsme.24-00074, 2024 (in Japanese).
[5] Suzuki D. and Takahashi K., “Effect of Operating Conditions on the Life of Traction Motor Bearings,” *RTRI Report*, Vol. 40, No. 3, pp. 11-18, 2026 (in Japanese).
[6] Japanese Standards Association “Rolling stock-Bogie-General rules for design of bogie frame strength,” JIS E 4207: 2019, 2019 (in Japanese).
[7] Research Committee on Temperature Dependency of Grease Life, Japan Society of Lubrication Engineers, “Report of Joint Research on Temperature Dependency of Grease Life (1) -Effect of Temperature on Functional Life of Greases -,” Vol. 30, No. 10, pp. 725-732 1984 (in Japanese).
[8] NSK, “Rolling Bearings for Industrial Machinery,” CAT. No.E1103, p. C130, 2016.
[9] NTN, “Ball and Roller Bearings,” CAT. No. 2203/E, p. A-114, 2020.

Authors



Daisuke SUZUKI, Dr.Eng.
Senior Researcher, Lubricating Materials Laboratory, Material Technology Division
Research Areas: Tribology, Machine Elements, Rolling Bearings



Ken TAKAHASHI, Ph.D.
Senior Researcher, Lubricating Materials Laboratory, Material Technology Division
Research Areas: Tribology, Machine Elements, Rolling Bearings

Summaries of Papers in RTRI REPORT (in Japanese)

Exhaust Pipe Modifications to Reduce Soot Staining on Vehicle Bodies Based on Roof-level Exhaust Flow Simulations of Diesel Railcars

Natsuki HARADA, Yuhei NOGUCHI, Yuto ARAKI, Tokuzo MIYACHI

(Vol.40, No.8, 1-8, 2026.8)

Exhaust gas emitted from diesel railcars can sometimes cause soot staining on surface of car body. Computational fluid dynamics (CFD) simulations were conducted in this study to investigate the exhaust flow around a car body and to analyze both the causes of soot staining and its countermeasures. Based on the findings from CFD simulations, the exhaust pipes of the actual car were modified. The degree of staining with and without the countermeasure was evaluated based on the image brightness of the car body. As a result, the modified exhaust pipe reduced soot staining on the car body surface.

Neutral Temperature Characteristics of Continuous Welded Rail Based on Rail Creeping Surveys and Long-Term Axial Force Measurements

Shingo MATSUTANI, Yuki NISHINOMIYA, Shingo TAMAGAWA, Akito HOSOMI

(Vol.40, No.8, 9-19, 2026.8)

Continuous welded rail management controls rail axial force using its equivalent temperature, taking into account rail creeping and maintenance work. Furthermore, changes in neutral temperature are appropriately updated after resetting. However, track buckling has still occurred, even under such management, suggesting that unexpected rail axial forces may be generated by temporal changes in neutral temperature after resetting. This study analyzed changes in the amount of rail creeping, focusing on the track structures, and measured the axial force and temperature over a long term from the time of long-welded rail installation to understand the variations in neutral temperature characteristics. As a result, the authors confirmed the quantitative impact of the factors contributing to rail creeping and found that the possibility of a large rail axial force occurring due to temporal changes in the neutral temperature after resetting is low.

Long-Term Continuous Day-and-Night Analysis of Rail Temperature Using an All-Sphere Visible Element

Fumihiro URAKAWA, Tsutomu WATANABE

(Vol.40, No.8, 20-33, 2026.8)

This paper proposes a method for long-term, continuous day-and-night analysis of rail temperature. The method is based on a weakly coupled framework that combines detailed thermal evaluation of solar and longwave radiation with heat conduction analysis, using an all-sphere visible element. Atmospheric attenuation of longwave radiation and reflected radiation from surrounding objects is modeled using the Lambert–Beer law, with absorption coefficients adjusted for seasonal water vapor conditions. Validation against field measurements showed that the RMSE of the analyzed rail temperature ranged from 1.4 °C to 2.5 °C in summer and 1.6 °C to 2.1 °C in winter.

Simulation-based Verification of Braking Distance Controllability under Shunting Train Conditions

Yuta KIZAKI, Kohei MIYAHARA, Shin-ichi NAKAZAWA

(Vol.40, No.9, 1-11, 2026.9)

In train shunting at railway facilities, it is difficult to control the stopping distance due to differences from main line operation regarding brake control elements. In this study, we create a simulation environment to reproduce expected shunting train conditions and conduct Train Automatic Stopping Control (TASC) simulations. We will also confirm the controllability of the stopping distance using a previous brake control method under shunting train conditions.

Development and Application of a Numerical Simulation for a Large-Scale Low-Noise Wind Tunnel

Koji NAKADE, Hiroaki NAKANO, Takumi ABE, Atsushi IDO, Tatsuya INOUE

(Vol.40, No.9, 12-21, 2026.9)

This study presents a numerical wind tunnel that replicates the airflow environment of the Maibara Large-Scale Low-Noise Wind Tunnel at the Railway Technical Research Institute. The tool integrates detailed geometric modeling using a 3D laser scanner with high-fidelity flow simulations performed on the RTRI's computational platform. Validation against physical wind tunnel tests demonstrated a strong consistency in airflow characteristics. By reducing the number of required test conditions and measurement points, this numerical tool enhances the efficiency and accuracy of wind tunnel testing. When used alongside conventional experiments, it accelerates research workflows and provides a complementary means for investigating aerodynamic issues.

Investigation of Causes of Breakage of Front Rod L-shaped Fitting and Maintenance Methods to Prevent It

Shunsuke SHIOMI, Sohei SHIGEMORI, Yoshikazu OSHIMI

(Vol.40, No.9, 22-32, 2026.9)

Breakage of front-rods caused by environmental conditions around turnouts and point machines has been an issue in maintaining signalling equipment. This study first investigated the relationship between environmental conditions and strain in front rods through field surveys and measurement tests. Secondly, this study conducted a stress analysis of front rods using a FEM. The results identified the locations of stress concentration in front rods and clarified how the extension adjustment condition affects stress. Using these findings, precautions for maintenance work on turnouts and point machines were proposed to prevent breakage of front-rods.

Stereo Camera-Based Obstacle Detection System in Railway Yards

Hiroki MUKOJIMA, Ryo KAGEYAMA, Nozomi NAGAMINE, Syoji KOMURO

(Vol.40, No.9, 33-42, 2026.9)

Several railway operators are currently advancing toward driverless operations. To achieve end-to-end autonomy, operational conditions must be investigated not only along mainlines but also within depots and station premises. Our prior work addressed a forward obstacle detection system for mainline. Because environments such as yards and stations feature distinct characteristics—including object types, operation speeds, sightlines, and operational rules—this paper redefines the performance requirements and detection targets. Focusing on these environments, this paper defines the performance requirements based on current technology, proposes a train-borne obstacle detection architecture tailored to depot/station operations, and reports evaluation results for representative scenarios.

Startup Characteristics of an Eddy Current Brake Using Capacitor Self-Excitation Phenomenon under Low-Speed and Large-Gap Conditions

Yasuaki SAKAMOTO, Keigo UKITA

(Vol.40, No.10, 1-10, 2026.10)

To explore the potential of electric braking in railway vehicle trailer cars, this study investigates capacitor self-excitation in eddy current brakes (ECBs) with the aim of enabling regenerative braking. While the characteristics under favorable conditions for self-excitation—such as high speeds and small air gaps—have been examined, the behavior under less favorable conditions, including low speeds and large air gaps, remains unclear. This study focuses on analyzing the behavior of the ECB under such less favorable conditions. Anomalous sustained excitation with small currents was observed, deviating from conventional behavior. From experimental and theoretical analyses, this phenomenon is presumed to be caused by the nonlinear magnetic properties of the secondary cores, including initial magnetization and magnetic saturation.

Dynamic Measurement Method for Wheel-Rail Contact Position Using Ultrasonic Waves

Shinya FUKAGAI

(Vol.40, No.10, 11-24, 2026.10)

Information on the contact positions between train wheels and rails when travelling through curves is essential for evaluating new wheel and rail profiles, investigating wear and damage, and analyzing the sources and mechanisms of wheel–rail noise. The objective of this study is to dynamically measure wheel–rail contact positions using surface waves. Measurement tests were conducted using a full-scale test machine, a test track, and a train. The reflection position of the surface waves corresponded to the edge of the wheel–rail contact area. This result indicates the usefulness of surface waves as a method for measuring wheel–rail contact positions. Additionally, the study showed that the characteristic differences in contact positions between the leading and trailing axles of a bogie, as well as the two-point contact near the gauge corner of a high rail, can be evaluated using real measurements.

A Quantitative Method for Predicting Wheel-flat Dimensions Generated during Braking

Daisuke YAMAMOTO, Hirotaka SAKAI

(Vol.40, No.10, 25-35, 2026.10)

Wheel flats generated during braking cause noise, vibration, and damage to track infrastructure. However, due to measurement difficulties under severe thermomechanical conditions, their formation and growth mechanisms remain insufficiently quantified. This study proposes a practical method for predicting wheel-flat dimensions based on a one-way coupled analysis that integrates vehicle dynamics and wheel–rail frictional heat analyses. By accounting for temperature-dependent material properties and contact ellipse growth, the proposed method quantitatively predicts wheel-flat dimensions. Validation against field measurements demonstrates good agreement, confirming the method's applicability for evaluating influential parameters governing wheel-flat development.

Development of Gap Adjustment Device between Switch Rail and Stock Rail, and Proposal of a Turnout Point to Reduce Maintenance Labor

Shingo TAMAGAWA, Masashi UEDA, Shintaro YASUDA, Akito HOSOMI

(Vol.40, No.10, 36-45, 2026.10)

The aim of this study is twofold: to develop a gap adjustment device for use between the switch rail and the stock rail, and to propose a new turnout structure that contributes to reduced maintenance efforts. The authors designed and produced adjustment devices that adjust the gaps between the switch rail and the stock rail expanding and contracting after switching. The developed devices were installed on a turnout, and a series of performance tests were conducted. As a result, we confirmed that the devices enable automate the inspection and adjustment of gaps. A running test involving a vehicle passing over the turnout also confirmed that the devices functioned normally during and after the vehicle's passage.

Study on the Introduction of Digital Technology for Collecting and Utilizing Accident and Near-Miss Reports

Noriko ONOMA, Yasunori OKADA, Munendo FUJIMICHI, Ryo NAKAMURA

(Vol.40, No.10, 46-53, 2026.10)

Railway operators are increasingly using digital methods to collect information on accidents and near-misses, although such data are rarely disclosed. To promote the widespread adoption of digitalization, this study surveyed implementation cases through questionnaires and interviews. The investigation identified the types and prevalence of digital technologies, and operational practices, as well as their benefits. This paper reports on the current status of adoption, practical approaches, and effects revealed through these surveys.

Editorial Board

Chairperson: Masamichi SOGABE
Co-Chairperson: Takamasa HAYASAKA
Editors: Ryohei IKEDA, Hiroyuki KANEMOTO, Tamami KAWASAKI, Yoshitaka KUBOTA, Susumu NAKAJIMA, Taisuke SANAGAWA, Erimitsu SUZUKI, Takamasa SUZUKI, Takashi YONEYAMA

QUARTERLY REPORT of RTRI

第 67 巻　第 4 号
2026 年 11 月 1 日　発行
監修・発行所：公益財団法人鉄道総合技術研究所
〒 185-8540　東京都国分寺市光町 2-8-38
発行人：曽我部正道
問い合わせ：鉄道総研広報

Vol. 67, No. 4
Published date: 1 November 2026
Supervision/Publisher: Railway Technical Research Institute
Address: 2-8-38 Hikari-cho, Kokubunji-shi, Tokyo 185-8540, Japan
Issuer: Dr. Masamichi SOGABE
Contact us: Public Relations, Railway Technical Research Institute
Mail Address: rtripr@rtri.or.jp

Copyright © 2026 Railway Technical Research Institute, Tokyo JAPAN All rights reserved.

QUARTERLY REPORT of RTRI

RAILWAY TECHNICAL RESEARCH INSTITUTE

Printed in Japan