Hierarchical Anomaly Detection and Model Update Framework for Steel Manufacturing

Yohei Harada[1], Masafumi Matsushita[2], and Takehide Hirata[3]

[1]*JFE Steel, Kawasaki, Kanagawa, 210-0855, Japan*
yoh-harada@jfe-steel.co.jp

[2]*JFE Steel, Chiyoda-ku, Tokyo, 100-0011, Japan*
mas-matsushita@jfe-steel.co.jp

[3]*JFE Techno-Research, Kawasaki, Kanagawa, 210-0855, Japan*
This work was conducted while the author was affiliated with JFE Steel.
t-hirata@jfe-tec.co.jp

ABSTRACT

In an integrated steelmaking process, equipment failures can significantly impact overall operations. Therefore, predictive detection and prevention of failures are critical. In this study, A three-layer hierarchical predictive detection system has been developed in order to utilize large-scale, multivariate operational data. This system is designed to identify overall trends by leveraging big data, detect correlation breakdowns through domain knowledge, and detect shifts in single-signal levels. The effectiveness of the proposed system has been confirmed through its application to actual operational data from the steel manufacturing process. In addition, general anomaly detection models, including our system, rely on quantifying deviations from a normal state as an anomaly score. In manufacturing settings, data drift often occurs due to factors such as equipment part replacements or changes in operational conditions. When data drift occurs, it becomes necessary to redefine the normal state. However, in manufacturing environments, temporary runs or experimental operations mean that the data following a drift is not necessarily guaranteed to normal data. Therefore, it is necessary to evaluate whether the data distribution is normal before and after the drift on a case-by-case basis. Current approaches do not provide a quantitative means to make this decision, leading to the issue that model updates depend on the judgment of experts. To address this, we propose a method that utilizes similar equipment conditions to guide the timing and procedure for model updates. By applying Jensen–Shannon divergence to measure differences among four data distributions—derived from two machines and two distinct periods— we provide appropriate guidance for model construction based on a table of potential anomalies. Through validation using real data from two adjacent continuous casters, we confirmed that identifying abnormal equipment and time periods enables us to propose appropriate normal operating windows. These validation results indicate that the proposed system allows for comprehensive predictive maintenance, integrating domain knowledge and thereby contributing to stable operations in steelmaking facilities.

1. INTRODUCTION

Data-driven approaches have become increasingly important in steel manufacturing, enabled by the availability of large-scale process data. Among these, predictive maintenance is a major application. Monitoring techniques for steelmaking have particularly focused on detecting breakout events—failures in which molten steel leaks from the mold—and various approaches have been proposed to address this issue (Zhang & Dudzic, 2006; Ansari et al., 2022). In the hot rolling process, a variety of anomaly detection methods have been proposed, targeting specific equipment such as the rolling drive systems (Naruse, Midorikawa, and Tanaka, 2012; Akechi, Midorikawa, & Kobayashi, 2012), hydraulic servo systems (Kitamura et al., 1991; Nozaki, 2010), and motor current monitoring for table rolls used in steel transport (Naruse et al., 2017; Hirata, Hachiya, & Suzuki, 2021). While these methods are effective for individual equipment units, the large number of units in steel plants makes comprehensive monitoring and management across all equipment increasingly complex. Consequently, recent years have seen the introduction of systems capable of monitoring multiple signals in the steelmaking process using general or integrated approaches (Kato, Hangai, & Fukami, 2023; Inoue, 2024; Sarda et al., 2021; Jakubowski et al., 2021). In this study, we developed and implemented a hierarchical anomaly sign detection system, which leverages large-scale data to monitor multiple equipment units in a layered monitoring framework (Hirata, Matsushita, Iizuka, & Suzuki, 2021). The system integrates multiple complementary monitoring approaches, including an entire monitoring of large data sets, domain-informed analysis and basic statistical methods.

On the other hand, when operating many anomaly detection models, a new problem of model management arises. This is because normal data frequently drifts due to equipment upgrades, component replacements, and changes in operating conditions, rendering the predefined normal range invalid. The most common approach to this data drift issue is to detect distribution changes using statistical measures, followed by model updates (Lai, 1995; Chu, Stinchcombe & White, 1996). More recently, sequential model update methods based on online learning have been proposed (Watte & Heinrichs, 2024). As a specific application, a model update method utilizing a dedicated drift detector has been proposed for collaborative robots (Kermonov, Nabissi, Longhi & Bonci, 2023). However, in the steelmaking process, parts are frequently repaired and reused, and components following routine maintenance are not necessarily in optimal condition. Although drift may be detected, a unique challenge lies in determining what truly constitutes a normal state. As an example of an approach to a similar issue, a method that switches between historical models has been proposed (Agate, Drago, Ferraro, & Re, 2022). However, when numerous models exist, their number may increase further, complicating management. To address this issue, this study proposes a method that compares similar equipment and provides a detailed explanation of the approach. While Harada, Hirata, Matsushita, Eto & Sato (2023) provided only an overview of the proposed method, this paper offers a comprehensive description of its technical details, specific application results, and a discussion.

2. HIERARCHICAL ANOMALY DETECTION SYSTEM

Figure 1 illustrates the structure of the developed hierarchical anomaly sign detection system. The steelmaking process is characterized by a diverse array of machinery and equipment organized in a hierarchical configuration. Accordingly, monitoring is conducted at three levels: the entire process, the equipment, and the instruments. At each level, appropriate methods are applied to ensure effective anomaly detection. This multi-layered approach enables comprehensive and precise monitoring from multiple perspectives. Furthermore, the adoption of flexible and broadly applicable monitoring techniques facilitates rapid and seamless deployment across various processes and plants. Another key feature of the system is its ability to visualize anomaly scores for each monitoring item using a heat map, thereby enabling efficient and comprehensive oversight of many monitored points. Figure 2 presents the system configuration and user interface. The interface displays the monitoring targets along the y-axis and time along the x-axis, with colors indicating computed anomaly levels for each cell. When the user clicks any cell, a pop-up window appears, providing detailed information. An additional function automatically generates a report that summarizes anomaly detection trends. This screen is accessible via the web from control rooms or office environments, contributing to the rapid diagnosis of anomalies.

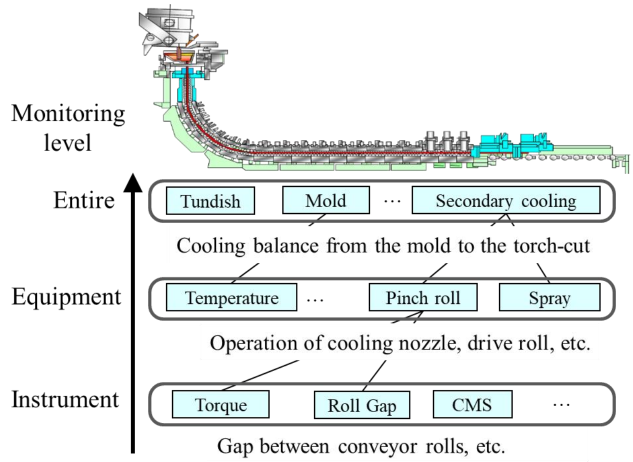


Figure 1. Outline of hierarchical anomaly detection system using a continuous casting machine.
Adapted from: Harada et al. (2023), Fig.1.
Copyright (2025) The Iron and Steel Institute of Japan

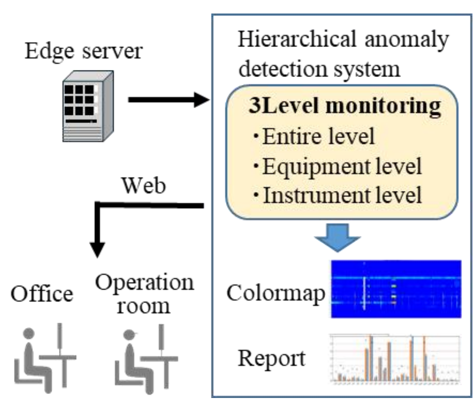
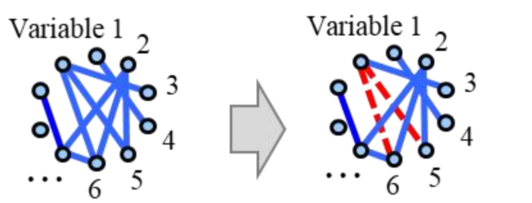


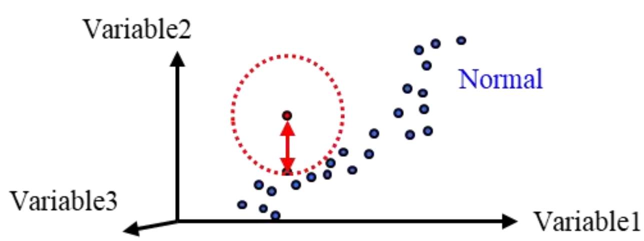
Figure 2. System configuration of hierarchical anomaly detection system.

2.1. Monitoring methods at each level

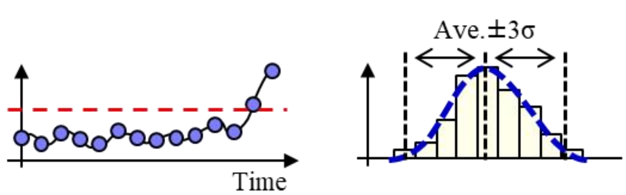
At the highest Entire level, where the number of variables considered exceeds several hundred, Lasso regression (Tibshirani, 1996) is utilized, as shown in Figure 3(a). Lasso regression sets the coefficients of less influential variables to zero, thereby providing a significant reduction in unnecessary explanatory variables from a large set. At the intermediate Equipment level, as shown in Figure 3(b), the primary method introduced is to monitor inter-variable correlations using data-driven models (DBM: Data-Based Models). This method targets systems where multiple data points change simultaneously. When a new measurement is obtained, it computes the distance to the previously registered normal data. This distance is regarded as the degree of deviation from the normal state. If it exceeds a threshold, the system is judged to be anomalous. Regardless of whether the relationships among variables are linear or nonlinear, if equipment exhibits a definable correlation, it becomes a candidate for this monitoring approach. In continuous casting machines, the focus is on flowmeters in the cooling system and on drive rolls. At the lower Instrument level, as shown in Figure 3(c), statistical methods or predefined thresholds are employed for upper and lower limit control. The data handled at this level consists of relatively simple indicators, such as motor vibration values, temperature, and roll gap.



(a) Entire level monitoring: Detecting structural breaks in correlational data using Lasso regression.



(b) Equipment level monitoring: Monitor anomalies based on the distance from the normal distribution.



(c) Instrument level monitoring: Statistical Process Control.

Figure 3. Overview of monitoring methods at each hierarchical level.

2.2. Detection example

As an example of mid-level equipment monitoring, we focus on the secondary cooling spray in a continuous casting machine. In this process, molten steel is continuously cast and cooled to form steel slabs, while the secondary cooling zone utilizes water sprays for slab cooling. It is well established that the relationship between the spray's flow rate and pressure can be modeled using a quadratic function. Figure 4 presents a time-series chart of anomaly scores for the secondary cooling zone, indicating that the scores gradually increased over several days and decreased following pipe cleaning. Figure 5 presents a scatter plot comparing the model data (representing the normal distribution) with the actual data from three days before and after the cleaning. During the period immediately preceding the cleaning, when anomaly scores were elevated, the distribution shifted upward and to the left relative to the model data. This suggests that higher pressure was required to achieve the same flow rate, indicating possible clogging of the spray nozzles. Following the cleaning, the distribution returned to a state closely matching the model, indicating that normal operating conditions had been restored.

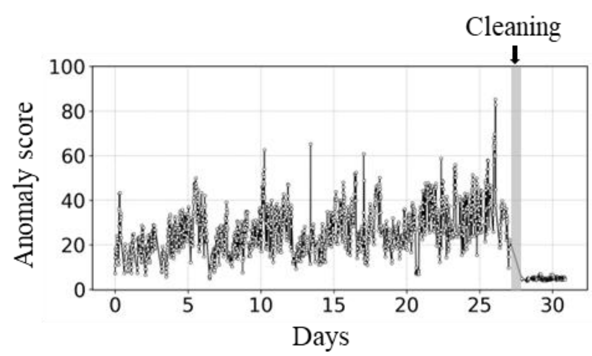


Figure 4. Time-series chart of anomaly scores for the secondary cooling zone spray.

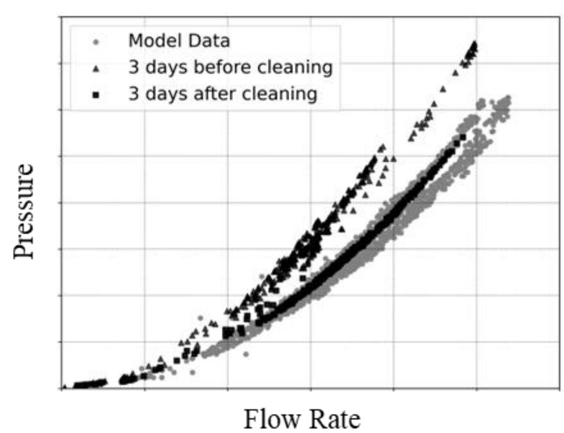


Figure 5. Scatter plot showing the relationship between flow rate and pressure in the secondary cooling zone spray.

3. MODEL UPDATING METHOD

Thus far, detection methods and examples of the hierarchical anomaly detection system have been introduced. However, as the number of monitoring models grows, managing the models themselves becomes increasingly challenging. For instance, in a continuous casting machine, hundreds of models may be employed, each requiring a defined normal operational range. Nevertheless, the more models there are, the harder it becomes to update continuously and appropriately each one. In most cases, model updates follow these steps:

(1) Focus on models with relatively elevated anomaly scores that remain below the detection threshold, as these may still indicate potential anomalies.

(2) Investigate recent trends to hypothesize the reasons behind the elevated scores and examine the values of the relevant variables.

(3) Based on the identified cause, decide whether to update the model by treating the current values as normal, or to regard the state as abnormal and refrain from updating.

Steps (1) and (2) can be partially automated, but for step (3), an engineer with in-depth field knowledge must rely on experience and insights, which presents a significant challenge to full automation. Expert engineers often gather supporting evidence by comparing the current state with similar existing equipment. The proposed method was developed with a focus on facilitating decision-making process.

3.1. Proposal method

Figure 6 presents a schematic diagram of the proposed method, which estimates normal data by comparing similar equipment across different time periods. First, assuming that the time periods correspond to pre- and post-repair or regular maintenance phases, equipment A and B are divided into Time Periods 1 and 2, resulting in distributions labeled A1, A2, B1, and B2. Next, the differences among the four distribution pairs — (1) A_1 vs. A_2; (2) B_1 vs. B_2; (3) A_1 vs. B_1; and (4) A_2 vs. B_2—are evaluated using the Jensen–Shannon divergence, and a threshold is applied to determine statistical significance. Finally, as summarized in Table 1, the appropriate distribution to be used as normal data is selected for each case. However, for Cases 4 and 13, identifying whether Equipment A or B represents the normal condition may be infeasible, making it difficult to narrow down the possibilities. Subsequently, after evaluating each signal, the period determined to represent a "normal distribution" is designated as the new normal period for the model associated with that signal, and the model's normal values are updated accordingly. For instance, in a model utilizing signals from Case 5 (where B_1 and B_2 are normal) and Case 14 (where A_2 and B_2 are normal), B_2—recognized as normal by both signals—is adopted as the updated normal data. This approach enables flexible adaptation even in scenarios where multiple signals exhibit drift.

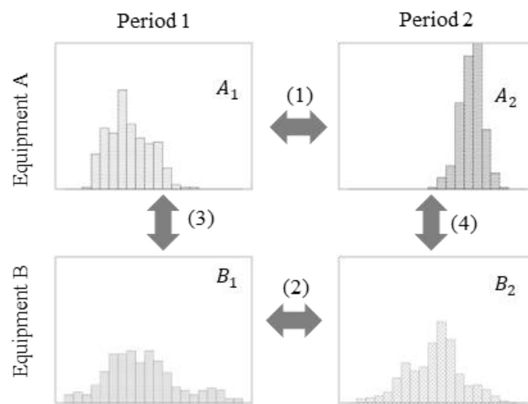


Fig.6 Comparison of data distributions for Equipment A and B across two periods.
Adapted from: Harada et al. (2023), Fig.3.

Table 1. Normal distribution judgment table based on four conditions.

Case	Condition				Normal distribution
	(1)	(2)	(3)	(4)	
1	○	○	○	○	Nothing
2	○	○	○	-	A_2, B_2
3	○	○	-	○	A_1, B_1
4	○	○	-	-	A_1, A_2 or B_1, B_2
5	○	-	○	○	B_1, B_2
6	○	-	○	-	A_2, B_1, B_2
7	○	-	-	○	A_1, B_1, B_2
8	○	-	-	-	B_1, B_2
9	-	○	○	○	A_1, A_2
10	-	○	○	-	A_1, A_2, B_2
11	-	○	-	○	A_1, A_2, B_1
12	-	○	-	-	A_1, A_2
13	-	-	○	○	A_1, A_2 or B_1, B_2
14	-	-	○	-	A_2, B_2
15	-	-	-	○	A_1, B_1

3.2. Evaluation method

The verification was conducted by analyzing the actual current data of the pinch rolls, obtained from a continuous caster during normal operations. The continuous caster distributes molten steel from a single tundish to two molds, producing two slabs simultaneously. Of the two production lines, one is designated as Equipment A, and the other as Equipment B. The transport pinch rolls, which consist of motor-driven rolls spaced at regular intervals, are used to extract the slab, which has solidified externally. Period 1 was defined as approximately one week prior to regular maintenance, while Period 2 corresponded to approximately one week following maintenance. The continuously collected data, converted into per-slab values, were used as representative indicators. All conditions (1) through (4) were tested using a threshold of 0.3 to determine significant differences between distributions. This threshold was set by referencing equipment standards and the distributions observed during past abnormal conditions.

3.3. Evaluation results

Table 2 presents the evaluation results for drive rolls #1 through #4, while the corresponding histograms are shown in Figure 7. Each histogram overlays the four distributions: A_1, A_2, B_1, and B_2. For Roll #1, all distributions were classified as normal, and Figure 7(a) illustrates that the distributions overlap. For Roll #2, all distributions except A_1 were considered normal, and Figure 7(b) confirms that the mean value of A_1 is lower than those of the others. For Roll #3, all distributions except B_2 were judged to be normal, and Figure 7(c) shows that the mean of B_2 is higher than the other distributions. Although B1 also differs in mean and variance compared to A_1 and A_2, its overall distribution remains within a reasonable range, suggesting that it does not significantly deviate from expected behavior. As part of an actual operational improvement initiative, adjustments were made to the operation in response to the observation that the values of B_2 exceeded the expected normal range. For Roll #4, either A or B was determined to represent the normal distribution, and Figure 7(d) reveals a marked difference between the distributions of Equipment A and Equipment B.

Table 2. Results of JS divergence and estimated normal distributions for Drive Rolls #1 to #4.

Roll	JS divergence				Case and estimated normal distribution
	(1)	(2)	(3)	(4)	
#1	0.07	0.02	0.01	0.01	16: All
#2	**0.69**	0.02	**0.40**	0.05	6: A_2, B_1, B_2
#3	0.12	**1.64**	0.01	**0.89**	11: A_1, A_2, B_1
#4	**1.17**	**0.72**	0.10	0.02	4: A_1, A_2 or B_1, B_2

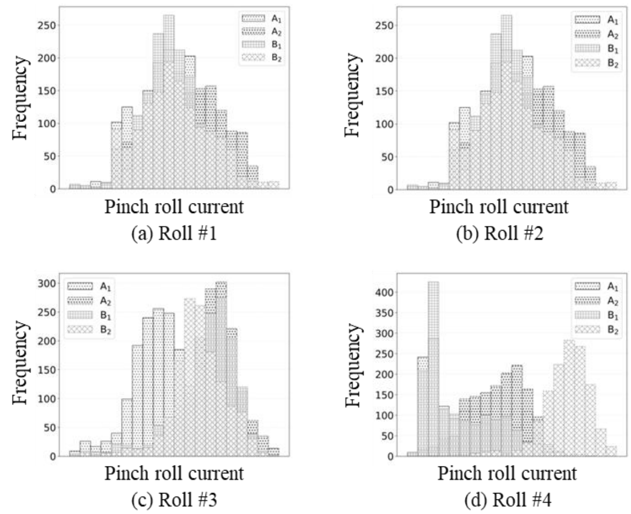


Figure 7. Histograms of distributions for A_1, A_2, B_1, and B_2 across Rolls #1 to #4.

3.4. Discussion

There remains scope for further exploration regarding the method used to compare distributions. In this study, we employed JS divergence, considering its computational efficiency and symmetrical properties, but alternative approaches such as the Wasserstein distance, Kullback-Leibler divergence, or custom drift detection algorithms could also be considered. However, KL divergence was not adopted because it is an asymmetric metric, which could lead to biased conclusions toward either equipment A versus B or period 1 versus 2. Although Wasserstein distance offers symmetry, its computational cost becomes significant for high-dimensional models or non-Gaussian distributions, which could impact existing system performance. If these limitations can be addressed, distribution comparison using alternative methods could be employed to estimate normal distributions by referencing Table 1. Furthermore, it may be beneficial to introduce a custom penalty function that assigns penalties to distributions exceeding the equipment's operational limits, based on historical failure data.

Similarly, determining the appropriate threshold values for drift detection warrants further consideration. In our verification, a uniform threshold of 0.3 was applied to all conditions (1) through (4); however, in practical applications, thresholds should be tailored to the characteristics of each specific piece of equipment. Moreover, by assigning distinct thresholds to conditions (1) through (4), designers can flexibly determine during the design phase whether to prioritize inter-equipment variability or temporal changes.

This method is applicable to configurations in which identical equipment is arranged in parallel, such as multiple

CVD systems in semiconductor manufacturing lines or parallel filling machines in food packaging lines, which are designed to improve throughput. Furthermore, it can also be applied to equipment with symmetrical structures or operations, such as sizing presses that apply pressure from both sides in metal rolling processes, or simultaneous processing systems for left and right doors in automotive manufacturing lines. These examples demonstrate the potential for broad applicability across various industrial domains.

4. CONCLUSION

In this study, a hierarchical anomaly detection framework was implemented in a continuous casting machine, which is an important process in steelmaking, and its detection results were presented. To address data drift issues in anomaly detection systems that rely on numerous models, a method was devised to identify representative normal data by comparing the distributions of similar equipment. To validate the approach, we applied the technique to the drive rolls of the continuous casting machine and confirmed its practical viability. The proposed method enables a reduction in reliance on subjective judgment. In the future, integration of sensor calibration records, repair histories, and system update logs is planned to achieve further automation.

REFERENCES

Agate, V., Drago, S., Ferraro, P., & Re, G. L. (2022). Anomaly detection for reoccurring concept drift in smart environments. *In 2022 18th International Conference on Mobility, Sensing and Networking (MSN)* (pp. 113–120). IEEE. https://doi.org/10.1109/MSN57253.2022.00031

Akechi, Y., Midorikawa, S., & Kobayashi, S. (2012). Nondestructive diagnosis for existing structures using amplitude attenuation of high frequency wave. *JFE Technical Report*, (17), 23–27.

Ansari, M. O., Ghose, J., Chattopadhyaya, S., Ghosh, D., Sharma, S., Sharma, P., Kumar, A., Li, C., Singh, R., & Eldin, S. M. (2022). An intelligent logic-based mold breakout prediction system algorithm for the continuous casting process of steel: A novel study. *Micromachines*, 13(12), 2148. https://doi.org/10.3390/mi13122148

Chu, C.-S. J., Stinchcombe, M., & White, H. (1996). Monitoring structural change. *Econometrica*, 64(5), 1045–1065.

Harada, Y., Hirata, T., Matsushita, M., Eto, K., & Sato, M. (2023). Application of anomaly detection system to steelmaking plant and development of model update guidance[in Japanese]. *Materials and Processes*, No.36, p.2.

Hirata, T., Hachiya, Y., & Suzuki, N. (2021). Anomaly-sign detection techniques for steel manufacturing facilities utilizing data science. *JFE Technical Report*, (26), 15.

Hirata, T., Matsushita, M., Iizuka, Y., & Suzuki, N. (2021). Fault detection technique for hierarchical monitoring of steel making facilities based on data science. *Tetsu-to-Hagane*, 107(11), 897–905. https://doi.org/10.2355/tetsutohagane.TETSU-2021-031

Inoue, M. (2024). Unsupervised anomaly detection in mixed processes using clusters. *Nippon Steel Technical Report*, (131), 85–89. https://www.nipponsteel.com/common/secure/en/tech/report/pdf/131-10.pdf

Jakubowski, J., Stanisz, P., Bobek, S., & Nalepa, G. J. (2021). Explainable anomaly detection for hot-rolling industrial process. In *2021 IEEE 8th International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 281–290). IEEE. https://doi.org/10.1109/DSAA53316.2021.00045

Kato, K., Hangai, S., & Fukami, S. (2023). Integrated platform for facility monitoring for anomaly detection. *Nippon Steel Technical Report*, (131), 65–72. https://www.nipponsteel.com/common/secure/en/tech/report/pdf/131-09.pdf

Kermenov, R., Nabissi, G., Longhi, S., & Bonci, A. (2023). Anomaly detection and concept drift adaptation for dynamic systems: A general method with practical implementation using an industrial collaborative robot. *Sensors*, 23(6), 3260. https://doi.org/10.3390/s23063260

Kitamura, A., Nakayama, M., Konishi, M., & Yabumoto, J. (1991). A study on the optimal control of tension and thickness in tandem cold rolling mills [in Japanese]. *Transactions of the Institute of Systems, Control and Information Engineers*, 4, 235.

Lai, T. L. (1995). Sequential changepoint detection in quality control and dynamical systems. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(4), 613–644.

Naruse, T., Takada, M., Hirata, T., Hagio, Y., & Akechi, Y. (2017). Development and practical application of crack detection technology for steel materials using high-temperature TOFD method. *Tribologist*, 62, 671.

Nozaki, N. (2010). Development of high-speed ultrasonic inspection system for steel products [in Japanese]. *CAMP-ISIJ*, 23, 203.

Sarda, K., Acernese, A., Nolè, V., Manfredi, L., Greco, L., Glielmo, L., & Del Vecchio, C. (2021). A multi-step anomaly detection strategy based on robust distances for the steel industry. *IEEE Access*, 9, 53827–53837. https://doi.org/10.1109/ACCESS.2021.3070659

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267–288.

Wette, S., & Heinrichs, F. (2024). OML-AD: Online machine learning for anomaly detection in time series data. *arXiv*. https://doi.org/10.48550/arXiv.2409.09742

Yamamoto, T., Wada, K., Kanoko, S., & Higuchi, A. (2015). Present status and future of equipment diagnosis technology for iron- and steelmaking plants [in Japanese]. *Shinnittetsu Sumikin Giho*, (402), 2. https://www.nipponsteel.com/tech/report/nssmc/pdf/402-02.pdf

Zhang, Y., & Dudzic, M. S. (2006). Industrial application of multivariate SPC to continuous caster start-up operations for breakout prevention. *Control Engineering Practice*, 14(11), 1357–1375.

BIOGRAPHIES

Yohei Harada was born in Tokyo, in 1994. He received M.S. of Engineering from Tokyo Metropolitan University, Japan, in 2018. He is currently Research Scientist with Cyber Physical System Research and Development Department, in Steel Laboratory, JFE Steel Corporation. He is member of The Iron and Steel Institute of Japan.

Masafumi Matsushita was born in Wakayama Prefecture, Japan, in 1982. He received his Ph.D. in Physics from Rikkyo University, Tokyo, Japan. He began his research career in nuclear physics and accelerator science, working at both RIKEN and the Faculty of Science at the University of Tokyo. He later joined JFE Steel Corporation, where he has been engaged in research and development of cyber-physical systems, with a particular focus on equipment diagnostics and anomaly detection technologies. He is currently working in the Intelligent Technology Development Department at JFE Steel Corporation, promoting digital transformation (DX) through simulation and optimization technologies. His research interests include predictive maintenance, data-driven modeling, and system optimization. He is a member of The Iron and Steel Institute of Japan.

Takehide Hirata received his B.S. and M.S. degrees from Tokyo Institute of Technology, Japan, in 1990 and 1992 respectively. He joined Kawasaki Steel Corporation (currently named JFE Steel Corporation) in 1992. He transferred to JFE Techno-Research Corporation in 2024. He is currently a staff at the Data Science Center.