Establishment of Condition Monitoring and Predictive Anomaly Detection System in No.2 Bloom Mill at Kakogawa Works

Takashi HAGIWARA[*1] · Yoshitaka INOUE[*1] · Kazuhiro MORII[*2] · Hisafumi UNO[*3] · Tsutomu ANDO[*4]

[*1] Wire Rod Production Department, Kakogawa Works, Steel & Aluminum Business
[*2] Digital Innovation Technology Center, Technical Development Group
[*3] Process Control Department, Kakogawa Works, Steel & Aluminum Business
[*4] Equipment Department, Kakogawa Works, Steel & Aluminum Business

Abstract

To ensure the stable operation of the No.2 Bloom Mill at the Kakogawa Works, a data acquisition apparatus has been introduced to develop a system for detecting signs of equipment abnormalities. To accurately capture the signs of abnormalities in the No.2 Bloom Mill, where diverse operating conditions and patterns exist, a method has been adopted that enhances the versatility of models by extracting, learning, and analyzing data at suitable intervals for evaluation. This has led to the development of a system, Mode Oriented Novel Anomaly Detector (MONAD), which enables the steps from data preprocessing to model creation without the need for programming. Standardization of the flow from the model creation to maintenance work using MONAD has enabled the set-up of a system that can be operated by technicians and operators without expertise in machine learning. Efforts to improve the accuracy of the model are being made through its operation at the manufacturing sites.

Introduction

The Kakogawa Works No.2 Bloom Mill (hereafter, No.2 Bloom Mill), which began operation in 1970, produces steel billets destined for the rolling mills that produce our steel wire rod and bar products. In 2017, we implemented a major expansion in conjunction with the consolidation of upstream processes[1)] to increase the production capacity to a level unparalleled throughout the world, at 300,000 tons per month. The No.2 Bloom Mill supplies almost all the steel billets used in Kobe Steel's steel wire rod and bar products. As such, ensuring the stable operation of the No.2 Bloom Mill is critical for ensuring the stable supply of the entirety of the company's steel wire rod and bar products.

Maintenance plans have conventionally been driven by time-based maintenance (TBM) and condition-based maintenance (CBM) approaches, which are based on past experience as well as daily inspections that reveal equipment condition. There was no option other than to apply provisional TBM and CBM standards to the new equipment introduced as part of the 2017 expansion. However, we knew we needed to introduce methods that could more reliably prevent sudden equipment failure. This paper details the initiative that spurred at Kobe Steel in 2018 to develop the information infrastructure and system for continuous condition monitoring and predictive anomaly detection for a wide variety of equipment at the No.2 Bloom Mill.

1. Guidance for developing an anomaly detection system

1.1 Overall system concept

A common method of using machine learning for anomaly detection is to learn the regular patterns in data and regard deviations from these regular patterns as predictive signs of an anomaly.[2), 3)] Many anecdotes report the implementation of anomaly detection based on multivariate analysis of data aggregated from individual plants that are in continuous operation. This method has attracted attention because it can uncover signs of anomalies that are based on unexpected relationships in the data that could not be discerned by humans.[4)]

However, there is a particularly wide variety of operating conditions and operating patterns at the No.2 Bloom Mill. This is because certain circumstances call for manual operation and because parameters such as steel grade and dimensions affect motor loads and the processes executed. In contrast to a plant that is in continuous operation, data across this operation lack regular patterns; therefore, the conditions flagged by machine learning as deviations may be false positives rather than predictive signs of anomalies.

One potential solution is to train individual models for each operating condition and operating pattern. However, this would require a prohibitively large number of models. It would also be impossible to apply these models to new operating conditions and operating patterns. Therefore, we set out to develop a method to increase the model's versatility. We abandoned the focus on operating conditions and operating patterns, instead applying machine learning to subsets of data suitable for evaluation, such as those in which the load is constant. Determining the conditions for data extraction requires considerations beyond operating conditions and operating patterns. Critical as well are equipment structure and conditions related to automatic operation, making the expertise of operators and maintenance technicians essential. Even with the data extraction method described above, exceptional circumstances could lead to false positives. This makes it challenging to define absolute standards for what should occur when an anomaly is detected, such as immediately shutting down equipment without human intervention - a human must ultimately decide the proper action.

We therefore devised the plan below to develop a model creation system for operators, maintenance technicians, and shop floor workers who do not have expertise in machine learning.

- Models should be optimized based on individual equipment rather than multivariate analysis of aggregated data for the entire plant.
- The system's tools for data preprocessing and model creation must not require programming.
- The maximum number of variables should be limited to two so the user can visually comprehend the model's behavior and the relationship between the data and actual events.

1.2 Identifying which equipment to monitor

The No.2 Bloom Mill has several thousand pieces of equipment; it is unrealistic to create anomaly detection models for all of them. With stable operation of the No.2 Bloom Mill as the objective, we narrowed our focus to equipment expected to cause production line downtime of 24 hours or more following malfunction. We put our efforts toward creating a model for predictive anomaly detection and collecting data in support of continuous monitoring.

Continuous monitoring data required for such a model include motor speeds and currents, control signals, and other data acquired by conventional control networks.

2. Implementing a data acquisition apparatus

2.1 Requirements for a data acquisition apparatus

The No.2 Bloom Mill has a complicated control system, with equipment from a variety of manufacturers due to the plant's expansion and iterations of piecemeal upgrades to the control system. Data acquisition apparatuses differ by manufacturer, and data for some signals could not be collected. Together, these issues made it impossible to collect all data synchronously and in a central repository - two criteria that are critical because they make it possible to monitor correlations within data and to detect deviations, which in turn are critical for anomaly detection.

We established the five stipulations below for the data acquisition apparatus to support the development of an anomaly detection system.

(1) The apparatus must communicate with the control networks of major Japanese PLC (programmable logic controller) manufacturers to facilitate deployment at other Kobe Steel factories.
(2) The implementation of requirement (1) must enable synchronized data acquisition and centralized data storage.
(3) Data must be available as a text file to ensure universal data utilization.
(4) Collected data must be accessible via any computer on the network to facilitate data utilization.
(5) Scaling up and the addition of new functionalities must be possible to accommodate potential future needs.

To meet these stipulations, we chose the data acquisition apparatus ibaPDA, which is manufactured by iba AG (hereafter, iba).

2.2 Features of iba's data acquisition apparatus

iba's data acquisition apparatus can connect to the networks of many PLC manufacturers (requirement (1)) and can perform synchronized, central data collection (requirement (2)). It also provides numerous methods for integrating data with higher-level systems, and packaged data can be transferred via a text file (requirement (3)). Collected data can also be viewed via analysis software from any terminal on the company network for utilization (requirement (4)). Functionalities such as vibration analysis and image processing are easy to add to the software, making it highly scalable (requirement (5)). iba's software proved extremely useful in this project, as it processed numerous signals and visualized the correlations (**Fig. 1**), making it easy to conduct preliminary modeling studies.

2.3 Introducing iba's data acquisition apparatus

Connection to a Japanese PLC manufacturer's network was not guaranteed when the system was introduced. Therefore, we conducted a small-scale demonstration to verify that the system could

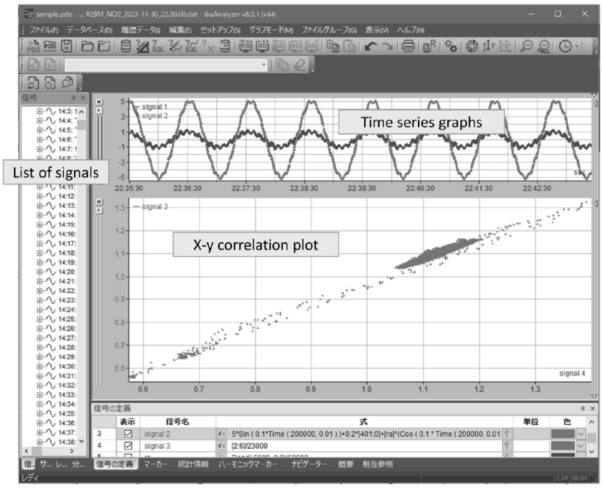

Fig. 1 Sample screen image of analysis software

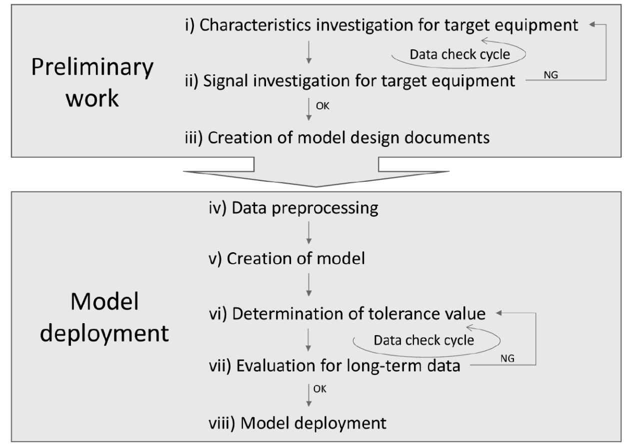

Fig. 2 Workflow for model creation

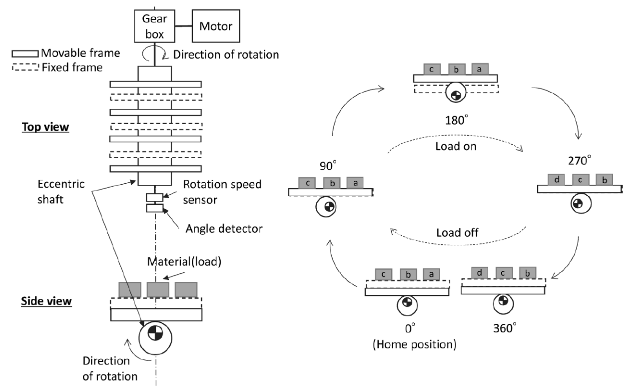

Fig. 3 Schematic diagram of equipment

perform stable data collection and central storage.

In addition, to make it easy to link collected data with Kobe Steel's other systems for complete utilization, we tagged each piece of equipment with a standardized ID as a universal key. We appended variable names and units of measure to important data to facilitate data utilization.

In addition, we developed a standard based on the No.2 Bloom Mill to guarantee that the system would be of equal quality when deployed at other plants.

3. Establishing a method for model creation

3.1 Process flow of anomaly detection model creation

Since the operators and maintenance technicians will spearhead the creation of optimized models for each piece of equipment according to the concept described in Section 1.1, we defined a standard workflow to ensure that each model would be of equal quality and developed with equal efficiency.

There is often little or no data on anomalies for training machine learning. Therefore, a model that allows for a certain degree of false positives is created, and then it is refined by trial and error until the rate of false positives is reduced to an acceptable level. This is why we created the workflow shown in **Fig. 2**, which is based on the "Conceptual Diagram of the Mixed-Standard Learning Lifecycle".[5] In this workflow, the model creation process is divided into a preprocessing phase and a deployment phase. Each phase has a data check cycle to incorporate the concept of agile development, a method that anticipates changes and additions during the development process. The preprocessing phase, which involves characteristic selection and data extraction, is particularly important for improving the accuracy of the model; the next section details this phase by way of example.

3.2 Example-guided explanation of the model creation process

Fig. 3 depicts the model creation process as it was applied to actual conveying equipment. This equipment has a movable frame connected to an eccentric shaft that picks up steel billets and advances them forward. We created the model based on a failure of the eccentric shaft to rotate, which is one of the possible failures of this equipment.

(i) Characteristics investigation for the target equipment

The first step is to select up to two characteristics for detecting rotational failure, in line with the development concept. Motor current can serve as the first characteristic because a failure of the eccentric shaft to rotate results in increased motor torque. This equipment repeatedly starts and stops during operation. Because the current constantly fluctuates alongside changes in speed, we selected motor speed as the second characteristic. The correlation between these two characteristics can be assumed to be constant when the conveying equipment is unloaded. However, while the steel billet is being lifted, factors such as the weight of the steel billet cause variation in the motor current, changing the correlation between the two characteristics. Changes caused by operating conditions, such as the weight of the steel billet, can lead to false positives. As such, the equipment drawings and control circuit schematics must first be consulted so that only no-load data will be extracted. Per Fig. 3, the no-load condition exists when the eccentric shaft angle is between 0° and 90° or between 270° and 360°.

(ii) Characteristic investigation via analysis software

The next step is to use analysis software to review the data distribution from the investigation in (i) for a potential correlation between the two characteristics. The time series data of the two selected characteristics and the angle of the eccentric shaft as the condition for data extraction are displayed to reveal key aspects of the data (**Fig. 4** (a)). In this example, in addition to the expected "with load" section, spikes with low reproducibility occurred immediately after the motor started and immediately before it stopped. Therefore, in addition to the "with load" angle condition, we also excluded the angle range in which these spikes occurred from the scope of data extraction. We used the analysis software to create a scatter plot of the two characteristics extracted under the applicable eccentric shaft angle condition (Fig. 4 (b)). If the data points are clustered, this means the extracted data are highly reproducible, and a model can be created. Conversely, if the points are highly diffuse, reproducibility is low, and it is not possible to create a model. In the latter case, extracted data must be reviewed. These steps constitute the most critical process in model creation. It is important to systematically record the approaches taken and the results of trials run as part of this process.

(iii) Devising a model design specification

Upon data validation, the content of the investigation should be recorded as a set of model design specifications. This makes it possible for people other than the model's creator to verify the model's design concept and validity, which is useful for long-term management of the model and application to maintenance activities.

We used Airtable (trademark Formagrid Inc.) web service to create the model design specification (**Fig. 5**). This service can store information related to the model in the form of text and images, including design information, equipment structure, operating plans, sensor signal information, correlation study results, and model evaluation results. It can also store comments and revision history information, making it easy to keep a record of the model creation process and changes made.

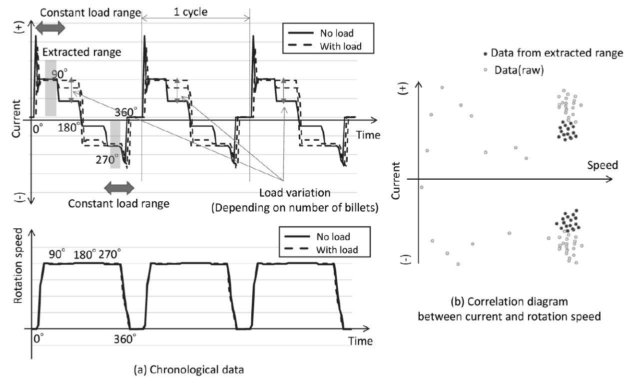

Fig. 4 Example of summary data

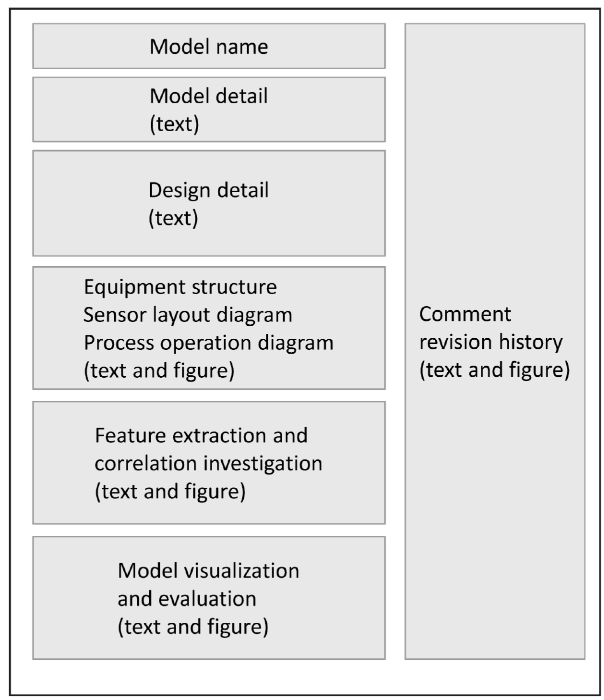

Fig. 5 Schematic image of Airtable

4. Developing an anomaly detection system

We developed the MONAD system (Mode Oriented Novel Anomaly Detector; **Fig. 6**) based on the design procedure standardized in Section 3, "Establishing a method for model creation." MONAD enables simple, no-programming model creation and implementation via the workflow shown in Fig. 2.

MONAD has two core strengths: (1) Only two characteristics are required for anomaly detection. As such, even operators and maintenance technicians without expertise in machine learning can ascertain model behavior via scatter plots and trend graphs, enabling the design of practical anomaly detection logic that capitalizes on their operational expertise. (2) MONAD has extensive

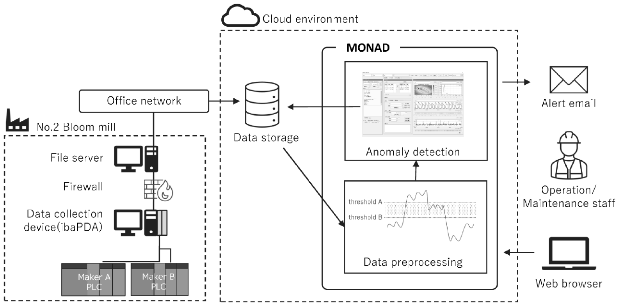

Fig. 6 MONAD system

preprocessing options, enabling simple, user-driven processing of the two variables selected. The following sections introduce MONAD's mode extraction and anomaly detection logic construction functions.

4.1 Mode extraction function

In MONAD, the regularity of data fluctuations occurring alongside operating conditions and operating patterns is called a "mode." To extract the desired mode from the data, the first step is to select up to two variables for the anomalies to be detected and the auxiliary signals, such as ON/OFF signals and command values, to be used for mode extraction. Mode extraction methods include the specification of a range of auxiliary signal values, segmentation based on fluctuation patterns within time-series data, and direct selection of the extraction range on a scatter plot. Users can combine these functions to extract the data they wish to evaluate, such as during constant-load operation, based on the knowledge they have about the equipment (**Fig. 7**).

4.2 Anomaly detection logic construction function

MONAD trains anomaly detection models such as the LOF (local outlier factor)[6] on data after mode extraction. The system then uses the trained models to calculate and graph anomaly scores for the evaluation period. The user can confirm that the model is designed as intended by graphing contour plots of anomaly scores on a scatter plot (**Fig. 8**).

Anomaly evaluation occurs via batch processing (processing of a predefined set of data in bulk) after deployment of the model. An anomaly is reported if the anomaly score exceeds the threshold at a frequency greater than the permissible frequency.

In addition to LOF, MONAD can perform anomaly detection by evaluating monotonically increasing or decreasing trends in time series data and changes in data distributions and waveforms. Users can select from these logics for anomaly detection, and new logics will be added based on user requests and analyses of anomaly cases.

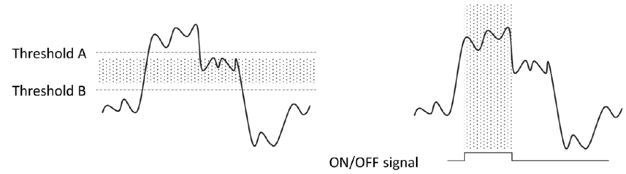

Fig. 7 Examples of mode extraction

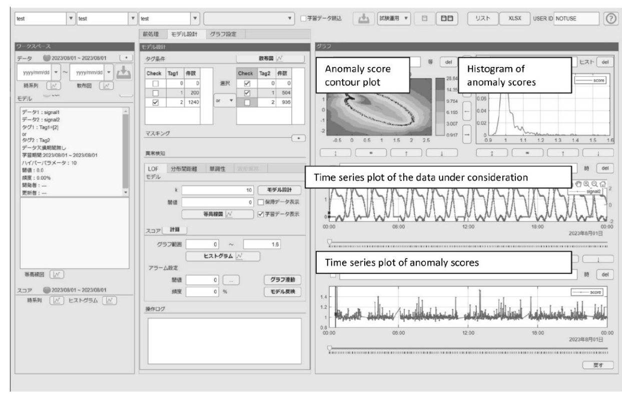

Fig. 8 Example of model design screen

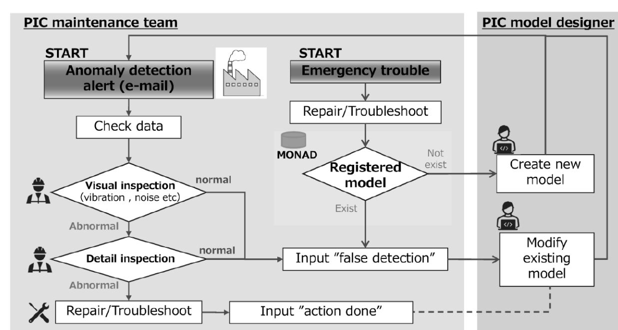

Fig. 9 Workflow of daily maintenance using MONAD

5. Application in maintenance

To use the models in MONAD for each piece of equipment as tools for anomaly detection in maintenance activities (predictive maintenance), the system must be deployed on the production line, and the models must be refined.

To support deployment on the production line, we first had to define the optimal workflow for data and situational validation following an anomaly detection alert (**Fig. 9**). We designed the interface based on this to provide maintenance technicians with the necessary information at each step when using MONAD.

MONAD automatically emails the maintenance technician when equipment is determined to be in an anomalous condition. The maintenance technician checks how the anomaly score is trending, checks the condition at the time of the anomalous data, performs a visual inspection of the equipment, and inspects the equipment for abnormal vibration and noise. If an anomaly is suspected during the visual inspection, the technician decides whether to shut the equipment down and conduct a detailed investigation. The technician records the inspection results and actions taken on the monitoring screen so the pertinent information is available to anyone who needs it, such as the model's creator.

After establishing the framework detailed above, the system went live in production in October 2023. We are validating and improving the accuracy of each model alongside our use of the system per the process flow shown in Fig. 9 The system has already detected early indications of poor bearing lubrication in steel billet conveying equipment; as such, we anticipate similar benefits in the future. For equipment comprising multiple drive components, such as reduction drives, however, we must continue validating the accuracy of the models alongside research into suitable condition monitoring methods based on the unique parameters of each piece of equipment.

Conclusions

This paper describes the efforts to collect data from equipment and introduce an anomaly detection system that uses these data at the No.2 Bloom Mill. Operators and maintenance technicians without expertise in machine learning can use the MONAD system and model creation process introduced in this paper to create equipment-specific models using agile methods. This constitutes a new approach to anomaly detection in plants with diverse operating conditions and operating patterns. Furthermore, MONAD itself is a versatile system that can easily be deployed to other plants alongside an appropriate selection of the data required for monitoring the equipment. We will continue to validate MONAD's effectiveness at the No.2 Bloom Mill and plan to deploy it to other plants within the company.

References

1) T. Taira et al. R&D Kobe Steel Engineering Reports. 2019, Vol.69, No.2, pp.37-41.
2) T. Ide. Introduction to Anomaly Detection using Machine Learning. CORONA PUBLISHING CO., LTD. 2015.
3) K. Yamanishi. Anomaly Detection with Data Mining. KYORITSU SHUPPAN CO., LTD. 2009.
4) T. Soma. Journal of Japan Society for Safety Engineering. 2018, Vol.57, No.2, pp.107-113.
5) National Institute of Advanced Industrial Science and Technology. Machine Learning Quality Management Guidelines. 4th edition. 2023-12-12. https://www.digiarc.aist.go.jp/publication/aiqm/AIQuality-requirements-rev4.1.0.0112-signed.pdf. Accessed 2023-12-13.
6) M. Breunig et al. Proceedings of the 2000 ACM SIGMOD international conference on Management of data. 2000, pp.93-104.