Technical Report

Application of Data-driven Models for Process Control and Operational Action Recommendations

Yasunori KADOYA* Akira MORITA
Satoshi KOSUGI Shota TAKESHIMA

Abstract

Due to progress in information technology and artificial intelligence, initiatives to utilize the large amount of accumulated operational data are progressing in the steel industry. In the steel manufacturing process, process control has been implemented by applying physical models developed based on principles to produce high-quality products. In process control, it is necessary to perform model calculations within a practical calculation time. However, there are problems such as the long computation time required to calculate the exact physical model and some control objects are difficult to build models based on principles. Therefore, improvement of the control accuracy and stabilization of operations by utilizing operational data are highly anticipated. This article introduces technologies and application examples of data-driven models based on operational data developed to realize high accuracy process control and advanced operational action recommendations.

1. Introduction

In recent years, significant progress has been made in information technology and artificial intelligence including machine learning. These developments have enabled large-scale database systems to be built in order to collect data through high-speed networks at low cost and to create added value from big data analysis. In the steel industry, there has been a notable increase in the data capacity and the storage period of accumulated data from operation control systems and production management systems. As a result, sophisticated environments have been built to collect and analyze various information across different systems. The steel industry has also developed technology to ensure stable measurements in its harsh environment, resulting in improved quality of actual operation data. Effective use of this data is expected to improve quality and productivity.

The steel industry produces diverse products with high productivity through various processes with strict operation restrictions. Since improvements have been made through accumulated technological development and operational knowledge, simply applying statistical analysis or machine learning is often insufficient to achieve better yield and higher productivity.

Nippon Steel Corporation has developed physical models based on fundamental principles and operational knowledge based on field experience over many years. To achieve further process improvements, physical models based on physical and chemical knowledge must be used in combination with statistical analysis and machine learning. And highly compatible methods with operational knowledge are also required. The research and development departments of instrumentation and control in Nippon Steel have continued the research and development of data modeling technology that leverages the strengths of both physical and statistical models. They have also tackled the research and development of advanced operational support through statistical analysis and machine learning using actual operation data. This paper introduces the data-driven model technology based on actual operation data and the cases in which the technology is applied to process control and advanced operational support.

In Chapter 2, we introduce our process control initiatives with the data modeling technology that leverages the strengths of physical and statistical models. Chapter 3 describes our advanced operational support based on big data.

2. Process Control by Data Modeling Technology

When it comes to quantifiable quality indicators expressed

* General Manager, Head of Dept., Automation Research Dept., Intelligent Algorithm Research Center, Process Research Laboratories
20-1 Shintomi, Futtsu City, Chiba Pref. 293-8511

through numerical values of dimensions (such as thickness and width) and temperature of steel products, the steel industry widely implements predictions and controlling processes using physical models based on physical and chemical principles. However, application of these physical models based on theories or experiments to actual processes requires adjustment of the models to fit the real equipment and processes. To adjust the errors of physical models and as alternatives to physical models, control models based on regression equations are widely used.

Physical models clearly show how their configurations and mathematical equations correlate with actual processes. They are convincing but cannot represent all of the actual phenomena. Also, they sometimes lack accuracy. On the other hand, statistical models, including machine learning models, are based on actual process data and easily obtain high accuracy. Nevertheless, their correspondence with actual processes is not always clear or is less descriptive. In addition, their low accuracy and reliability for new operating conditions and materials are problematic.

Therefore, gray box modeling by combining physical and statistical models has been developed to improve accuracy while ensuring descriptiveness for processes.[1] This gray box was named by mixing both the black box meaning low descriptiveness of the statistical model and the white box meaning high descriptiveness of the physical model.

Various types of gray box modeling methods can be considered. The following are the main methods applied to actual operations.

First, as shown in **Fig. 1** (a), the errors of the physical model are corrected by the statistical model to improve the prediction accuracy. This method builds the statistical model to predict errors by using the actual data of the physical model prediction errors.

Next, as shown in Fig. 1 (b), the parameters of the physical model are set by the statistical model. This method is suitable when the physical model can express the qualitative characteristics of the process but when it is difficult to set parameters in the actual process. In such a case, the parameters of the physical model are estimated from operation results by some method, and the statistical model is prepared to predict the parameters.

The above two cases are introduced in the following sections.

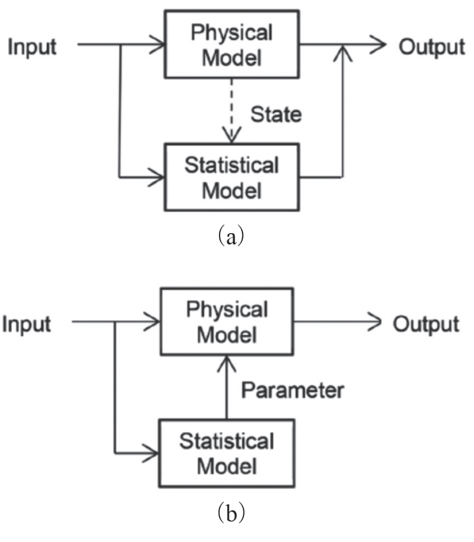

Fig. 1 Various configurations of gray box modeling

2.1 Technology for correcting physical model by statistical model

We introduce two technologies that improve the model prediction accuracy by correcting errors in the physical model using the statistical model in the first configuration of gray box modeling (Fig. 1 (a)).

The first technology uses an automatic stratification control model construction method to build a statistical model.[1–3] This method creates a nonlinear model by dividing the operation factor space into regions and superimposing nonlinear relationship equations (weighted averages) in divided local regions. The divided region structure is characteristic in that the overall picture of the model can be grasped to some extent. However, a certain number of data is required to build the model.

The other is a technology that constructs a statistical model by applying case-based modeling technology for building a model based on past empirical values, actual values, and input-output relationships. This method stores a large amount of data, including explanatory variables, in a database, extracts data with similar conditions from the database when necessary, and creates a local regression model. It is mesh-free (no need for region segmentation) and enables the construction of a model even when the number of data is small. However, it isn't easy to understand the overall aspect of the model.

When applying the technology for correcting errors in a physical model with a statistical model, an appropriate method is used by considering the abovementioned characteristics.

2.1.1 Hot-rolling crown and shape set-up with data modeling technology

Many statistical models used in actual operations use multiple models stratified by production conditions and other conditions to meet nonlinearity of processes. However, constructing these stratified models based on factors such as stratum conditions often relies on human experience and trial and error. As a result, the work load of model adjustment is high, and accuracy deterioration due to the diversification of product types becomes a problem. Therefore, we developed a method for building an automatic stratified control model using past operation data. This method produces a nonlinear model $\hat{y}$ by superimposing (weighted average) linear relationship equations $\hat{y}_i$ within local regions where the operation factor space is divided. The weight function Φ_i is automatically generated from actual data to build a highly accurate control model by minimizing the number of stratifications.

$$\hat{y}=\sum_{i=1}^{M}\underbrace{(w_{i0}+w_{i1}u_1+w_{i2}u_2+\dots+w_{ip}u_p)}_{\hat{y}_i}\,\Phi_i(\underline{u}) \qquad (1)$$

where $\underline{u}=(u_1, \dots, u_p)$ is an explanatory variable, M is the number of region divisions, and w_{ij} is a regression coefficient.

Figure 2 shows an example of a weight function that segments a two-variable space into three regions. The entire model can represent continuous and smooth characteristics by using a weight function that provides smooth region boundaries, as shown in Fig. 2. There is a method of selecting segmentation candidate points from among the points that equally divide the regions or the points that equally divide the number of data in the regions. It is also possible to provide segmentation candidate points according to prior physical knowledge or operational criteria. This approach enables constructing a model incorporating prior knowledge and with improved explainability.

With this method, it may take time to search for segmented re-

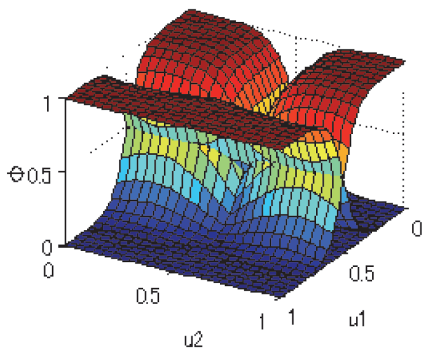

Fig. 2 Example of weight function

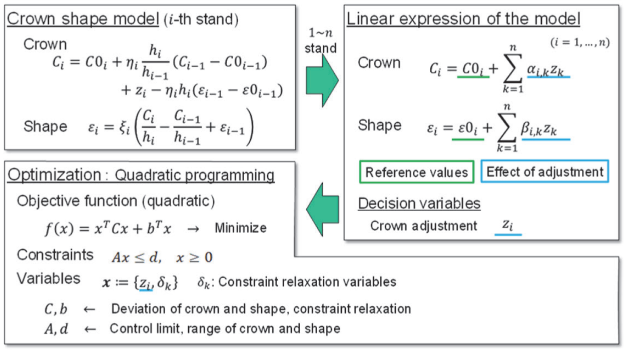

Fig. 3 Crown and shape schedule set-up

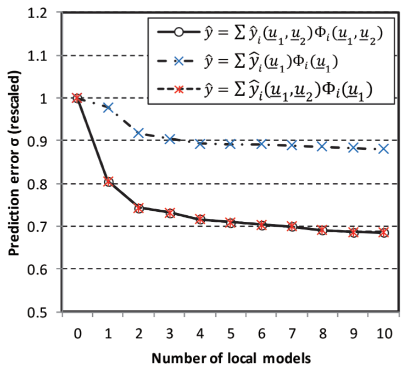

Fig. 4 Comparison of model variations

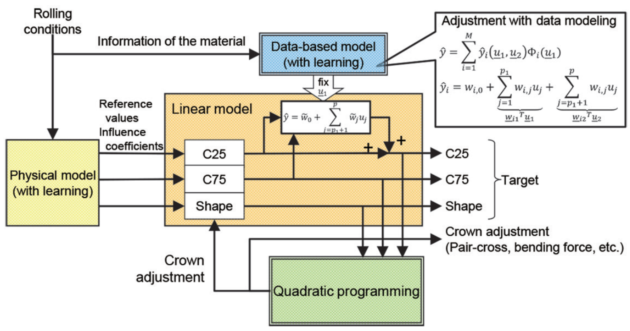

Fig. 5 Combination of data modeling and quadratic programming

gions, but prediction calculations by model equations after the region segmentation is determined are fast. Also, since only the weight functions and regression coefficients need to be calculated, there is no need to keep the actual result data. Furthermore, by recursively learning the regression coefficients, it is possible to follow operation fluctuations.

As an application example of this method, hot-rolling crown and shape set-up technology[3] is described below.

In hot finisher rolling, the strip crown must be held within the product tolerances at the exit of the last rolling stand, while the strip shape is contained within the stable strip threading range at the exit of each rolling stand. It is essential to properly set each rolling stand's crown and shape controllers (pair cross angle, work roll bending force, etc.) or to determine the crown and shape schedules properly. Mathematical optimization effectively sets the strip crown and shape within the controllable range by considering the upstream-to-downstream effects of crown and shape changes. Furthermore, to improve the predictive accuracy of strip crown, we developed crown and shape set-up technology in combination with physical model correction by data modeling technology using past actual results.

The crown and shape schedules from the finishing mill entry side to the finishing mill exit side can be determined appropriately by mathematical optimization such as quadratic programming after deriving linear expressions concerning the strip crown control variable (z_i) of each rolling stand from physical models[4] concerning the strip crown and shape (elongation difference ratio) before and after each rolling stand (**Fig. 3**). If the constraint conditions cannot be satisfied within the controllable ranges, it is possible to find feasible relaxation solutions by introducing constraint relaxation amounts or to find solutions by expanding the problem to include the strip thickness schedule as a decision variable.

The strip crown and shape are predicted according to physical models. In a real mill environment, it is difficult to determine the thermal expansion and wear of the rolls in operation and to obtain sufficient accuracy. Therefore, the physical models were corrected using past actual result data by an automatic stratification control model construction method. However, if this approach is applied as is, nonlinear optimization or linear approximation is required to determine the crown shape schedule.

This time, the explanatory variable $\underline{u} = (u_1, ..., u_p)$ of the correction model was divided into the variable $\underline{u}_1 = (u_1, ..., u_{p_1})$ that does not depend on the strip crown control amount and the control variable $\underline{u}_2 = (u_{p_1+1}, ..., u_p)$ that depends on the strip crown control amount. It was confirmed that the accuracy of the correction model did not change when the operation factor space was divided only by $\underline{u}_1$ (**Fig. 4**). By using this correction model, the crown correction amount to determine the crown and shape schedules can be expressed linearly concerning $\underline{u}_2$ by fixing $\underline{u}_1$. Even with quadratic programming, the optimal solution can be obtained quickly without deteriorating prediction accuracy (**Fig. 5**).

In an actual mill, parameter errors and process variations occur. Therefore, the prediction errors are recursively learned, and the crown correction model is updated. The coefficient w_{ij} of the linear relationship equations in the local regions is recursively corrected with the weight coefficient Φ_i fixed. As a result of application to an actual mill, the crown prediction accuracy was improved by about 30% compared to the physical model (**Fig. 6**).

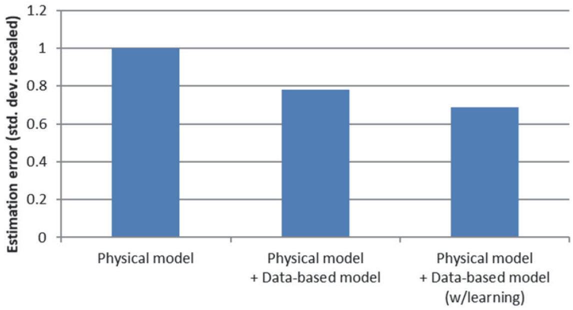

Fig. 6 Comparison of model variations

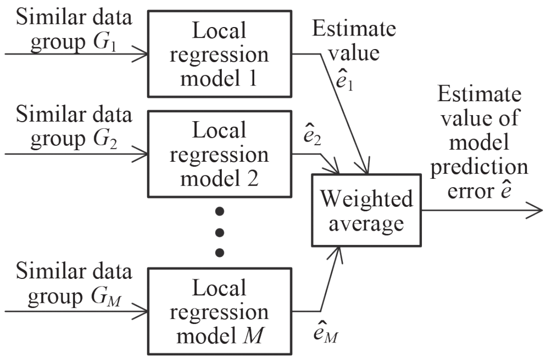

Fig. 7 Outline of ensemble learning using multiple local regression models

2.1.2 Plate cooling temperature control by using data modeling technology

Another method for correcting the physical model by the statistical model is introduced here. We developed a method for building a local regression model to estimate the model prediction error using past operation data with similar production conditions that is extracted from a database where large amounts of operation result data, including past production conditions, are stored. This method does not divide the operation factor space into several regions like the aforementioned automatic stratification control model construction method. When extracting data similar to the current production conditions, the distance d_i between the production conditions of operation result data No. i stored in the database, and the current conditions is evaluated for similarity by the weighted Euclidean distance of Eq. (2).

$$d_i = \sqrt{\sum_{j=1}^{m} w_j \cdot (x_{i,j} - x_{n,j})^2} \tag{2}$$

where m is the number of explanatory variables, j is a subscript for an explanatory variable, $x_{i,j}$ is an explanatory variable for the condition of operation result data No. i, $x_{n,j}$ is an explanatory variable for the current condition, and w_j is a weight factor.

A certain number of data with small distance d is extracted. Using the extracted data set, regression coefficients of the relationship equation (3) between the model prediction error e, which is the response variable, and the explanatory variable x_j are calculated to build a local regression model.

$$e = a_0 + \sum_{j=1}^{m} a_j \cdot x_j \tag{3}$$

where $a_1, a_2, ..., a_m$ are regression coefficients.

When the regression coefficients are obtained by using data sets with similar production conditions, the conditions that serve as explanatory variables are biased. Consequently, this may cause multicollinearity problems where there are highly correlated combinations of explanatory variables. This is not desirable from the viewpoint of application to online control. For this reason, a local regression model is built by applying the partial least squares (PLS) regression method that can eliminate the adverse effects of multicollinearity.

This way, the local regression model that corrects the model prediction error can be built. For industrial applications, it is necessary to eliminate the probability that the estimated value of the local regression model becomes abnormal and to increase the reliability of the local regression model. When estimating the prediction error of the physical model by the local regression model, the regression coefficients do not need to be accurate for the explanatory variables, and only the estimated value needs to be obtained with high accuracy. The idea of ensemble learning[5, 6] is applied here. Ensemble learning is a method of obtaining the final output by using the outputs of multiple models. When the weight w_j of the distance function equation (2) is changed in various ways to extract similar data to the current production conditions, multiple data sets including different data can be extracted. In this way, changing the weight w_j means changing the factors to be emphasized among the explanatory variables when extracting similar data.

Next, as shown in **Fig. 7**, M local regression models are built using the respective similar data sets $G_1, G_2, ..., G_M$. The estimated values $\hat{e}_1, \hat{e}_2, ..., \hat{e}_M$ of the model prediction errors under the current conditions are obtained from the local regression models. $M-2$ average values, excluding the maximum and minimum values from among the estimated values of M model prediction errors, are determined as the final estimated value $\hat{e}$ of the model prediction error.

Temperature control technology for accelerated cooling of steel plates is introduced below as an application example of this method.

Cooling temperature control uses a steel plate temperature prediction model, including a water cooling heat transfer model, to predict the steel plate temperature after cooling. The plate cooling equipment is operated so that this predicted temperature becomes the target temperature. The steel plate temperature prediction model is built based on heat transfer phenomena. However, some factors are difficult to model, such as steel plate surface properties. As a result, temperature prediction errors occur. The steel plate temperature can be controlled with high accuracy by using a local regression model that estimates the temperature prediction error and correcting the calculation results of the temperature prediction model (physical model). As shown in **Fig. 8**, this technology calculates the prediction error of the steel plate temperature prediction model after cooling. Then this calculated prediction error is stored in a database together with production conditions and operation result data. Before the cooling of the next steel plate starts, production conditions and operation results in the preceding rolling process that are similar to the current conditions are extracted from the database, a local regression model is built using similar data, and the model prediction error is estimated. In cooling control of the next steel plate, the target value of the control model is bias corrected for the model prediction error estimated as described before. Then, the plate transfer speed and cooling water flow rate, which are the manipulated variables of the cooling equipment, are determined.

As mentioned, the steel plate temperature prediction model is built based on heat transfer phenomena, but some factors, such as steel plate surface properties, are difficult to model. Therefore, the

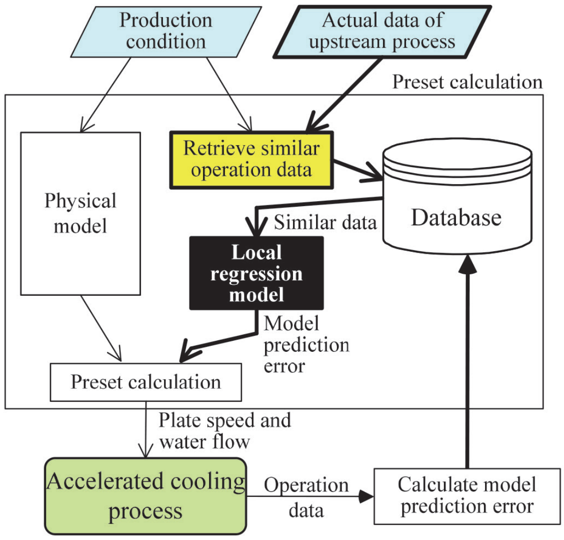

Fig. 8 Correction of model prediction error using local regression model in plate cooling temperature control

factors responsible for the steel plate temperature prediction error are used as explanatory variables for the local regression model. Specifically, the factors (HF) considered responsible for the errors in the heat flux as boundary conditions between the cooling water and the steel plate, and the factors (PP) considered responsible for the errors in thermal conduction and latent heat of phase transformation in the steel plate were used as explanatory variables. Concerning these factors, the weight w_j in the distance function equation (2) are changed for seven conditions, as shown in **Table 1**, and $M=7$ data sets are extracted. The weight of some factors is doubled or quadrupled by reference to the condition W-1, and local regression models are built using the respective data sets.

Using this method, the prediction accuracy of the steel plate temperature after cooling is improved to 10.9°C as compared to 17.2°C by the steel plate temperature prediction model (physical model), as shown in **Fig. 9**. This means that cooling temperature control can be executed with high accuracy.

2.2 Setting parameters of physical model using statistical model

Next, as an example of setting the parameters of a physical model using a statistical model, which is the second configuration of gray box modeling (Fig. 1 (b)), the use of a hierarchical Bayesian model to estimate the phosphorus concentration in the molten steel in the basic oxygen furnace is introduced below.[7)]

In basic oxygen furnace (BOF) blowing control, the total oxygen volume and the fluxes input are specified so that the molten steel temperature and composition at the blow end reach the target values as required for secondary refining. After the carbon concentration and molten steel temperature are determined by sublance sampling and measurement at the blow end, the molten steel composition and temperature are sequentially estimated for dynamic control, and the blowing progress is estimated.

Controlling the phosphorus concentration in the molten steel at the blow end is also essential for steel quality control. It is necessary to estimate the phosphorus concentration as well as the carbon concentration. Because the phosphorus concentration at the blow end is lower than that at the blow start, change in the phosphorus concentration $[P]$ in the molten steel can be approximated by the following

Table 1 Weight coefficient of explanation variable for ensemble learning simulation

Item No.	W-1	W-2	W-3	W-4	W-5	W-6	W-7
1 (HF)	1	2	1	1	1	2	2
2, 3 (HF)	1	1	1	1	1	1	1
4–6 (HF)	1	1	2	1	4	2	1
7–10 (HF)	1	1	1	1	1	1	1
11–24 (PP)	1	1	1	2	1	1	2

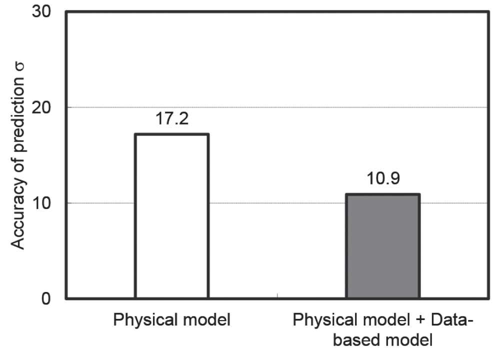

Fig. 9 Comparison of model prediction accuracy of plate finish cooling temperature

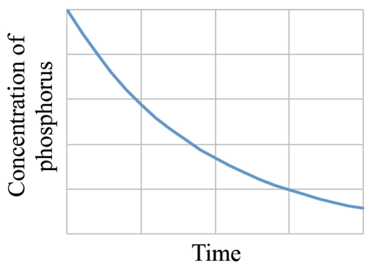

Fig. 10 Dephosphorization model

first-order reaction rate equation:

$$\frac{d[P]}{dt} = -k_p \times [P] \tag{4}$$

The proportional constant k_p is called the dephosphorization rate coefficient. Let $[P]_{ini}$ be the molten pig iron phosphorus concentration at the blow start and k_p be the dephosphorization rate coefficient. Then, the molten steel phosphorus concentration at any time can be estimated from the first-order reaction rate equation by $[P]=[P]_{ini}\times \exp(-k_p \times t)$ (**Fig. 10**).

If the end point phosphorus concentration $[P]_{end}$ and blowing time t_{end} are measured after blowing, the dephosphorization rate coefficient can be calculated by

$$k_p = \frac{1}{t_{end}} \ln \frac{[P]_{ini}}{[P]_{end}} \tag{5}$$

This dephosphorization rate coefficient k_p becomes a value that characterizes the dephosphorization behavior. To estimate the molten steel phosphorus concentration for dynamic control, it is necessary to accurately predict the dephosphorization rate coefficient during blowing. The following multiple regression model is constructed where the operation factor X_i known until the sublance measurement is an explanatory variable and the dephosphorization rate coefficient

k_p that can be calculated by Eq. (5) after blowing is an objective variable.

$$k_P = a_0 + \sum_{i=1}^{I} a_i \times X_i \tag{6}$$

where a_i (i=0,...,I) is a multiple regression parameter, and I is the number of explanatory variables. Using this multiple regression model made it possible to estimate the molten steel phosphorus concentration from the sublance measurement until the blow end.

However, when the measured phosphorus concentration at the blow end and the estimated phosphorus concentration were compared, the estimated phosphorus concentration was lower than the measured phosphorus concentration in the high phosphorous concentration range. It was feared that conditions unfavorable for dephosphorization were not sufficiently reflected in the multiple regression parameters.

Therefore, we concentrated on the molten steel temperature known to affect dephosphorization qualitatively and studied how to reflect this consideration in the multiple regression model. We assumed that the molten steel temperature at the blow end is sufficiently close to the target temperature and that the multiple regression parameters interact with the target temperature. In addition, one multiple regression parameter should have the same symbol as long as it remains within the target temperature range. If the target temperature is different, the contribution of one operation factor to dephosphorization should be qualitatively the same. These were set as improvement requirements.

As a hierarchical Bayesian model is suitable for satisfying these requirements, the following model was constructed.

$$\begin{aligned} k_{P,n} &\sim N\left(a_0 + \sum_{i=1}^{I} a_i \times \mathrm{X}_{i,n}, \sigma_Y^2\right) \quad (n = 1, ..., N) \\ a_i &= \bar{a}_i + b_i \times Z_n \quad (i = 0, ..., I) \\ b_i &\sim N(0, \sigma_i^2) \\ \sigma_i &\sim N_+(0, v_i^2) \end{aligned} \tag{7}$$

where $N(\mu,\sigma^2)$ represents the normal distribution of the mean μ and the variance σ^2, N_+ is a half-normal distribution, N is the number of data, and Z_n is the target temperature. A noninformative prior distribution and a weakly informative prior distribution were set for σ_Y and $\bar{a}_i$, respectively. A predefined positive constant was set for v_i. Furthermore, the maximum a posteriori probability estimate of the interaction parameter was used to estimate the molten steel phosphorus concentration during dynamic control.

This change was confirmed to improve the prediction accuracy of the phosphorus concentration at the blow end by 26% in the root mean square error (RMSE) (**Fig. 11**).

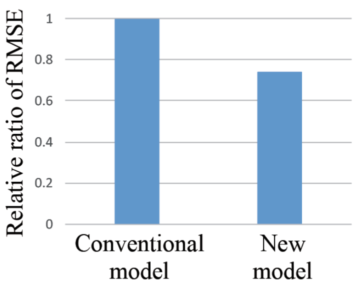

Fig. 11 Comparison of model prediction accuracy of phosphorus concentration

3. Advanced Operational Support Based on Data

It is difficult to construct models for judging events such as operation and equipment anomalies based on physical models. In modeling, relatively small numbers of operation and equipment anomaly data are mixed in with large numbers of normal operation and equipment data. A modeling method is required to detect abnormal data accurately when the number of normal and abnormal data is imbalanced (imbalanced data). When an anomaly occurs, various factors are considered to be intricately intertwined. It is thus difficult to gain higher performance with a simple linear regression model. Even if the occurrence of an operation or equipment anomaly is successfully learned and detected using a flexible and complex model, why the model has reached such a judgment is black boxed and is poor in interpretability. There are problems with the inability to identify factors and formulate improvement actions.

Considering the above issues, we developed a method for accurately modeling operation fluctuations and equipment anomalies that occur infrequently with reasonable accuracy, an interpretation method for extracting anomaly factors from model judgment results, and an advanced operation support framework for recommending appropriate change action amounts of main factors. The developed framework is described below by using the analysis of coke clogging as an example.

3.1 Framework to support analysis of coke clogging

During coke oven operation, the carbonization of coal and pushing of coke after carbonization are conducted daily and on an oven-by-oven basis. Let us consider the case where the load of the pusher is affected by the process results such as coal properties, oven temperature, carbonization time, and the oven condition in a pushing operation i, x_i=(x_{i1}, x_{i2}, ..., x_{ip}). The anomaly state when the pushing load exceeds the threshold value is labeled y_i=1, and the normal state when the pushing load does not exceed the threshold value is labeled y_i=0. If coke clogging occurs, it significantly damages the operating cost and equipment regarding its maintenance. When this anomaly state is reached, appropriate operation action improvement is required.

The developed framework is shown in **Fig. 12**. First, a model m that outputs the anomaly probability is learned from the past operation results variable data $\mathbf{X}$=[x_1, x_2, ..., x_n]^T and the corresponding past anomaly state $\mathbf{y}$=[y_1, y_2, ..., y_n]. Next, the anomaly probability Pr[y_j=1|x_j] is determined by inputting the operation variable x_j of the pushing operation j through the model m. Next, the contribution degree of the operation variables x_{j1}, ..., x_{jp} to the anomaly probability is determined. The operation variables with high contribution de-

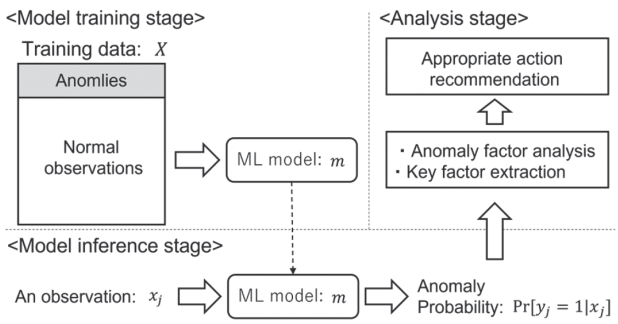

Fig. 12 Proposed framework of data-driven operation analysis

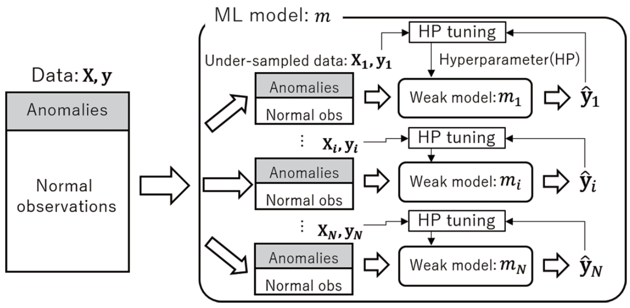

Fig. 13 Training of ensemble ML model considering imbalanced data

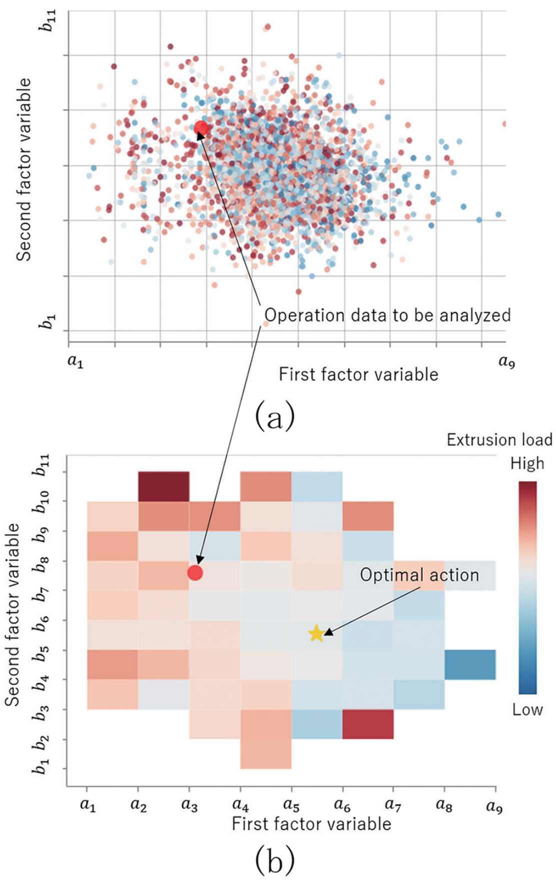

Fig. 14 Visualization and recommendation of optimal action

grees are extracted as a main factor group. Lastly, the action amount that reduces the pushing load for each main factor is visualized and recommended. The details are explained in the following sections.

3.2 Coke clogging anomaly detection machine learning model considering imbalanced data

In machine learning, the parameters are optimized to minimize the loss function using all available data during learning. In our present case study, the number of anomaly data is small and whole data is regarded as imbalanced data. The loss function becomes small if most of the normal data ($y=0$) is correctly detected. This might prevent the anomaly data from being detected.

With the proposed method, as shown in Fig. 12, the normal data that accounts for the majority of imbalanced data is undersampled. Sub-datasets are created by increasing the proportion of anomaly data compared to the normal data and are used for learning. This approach is expected to improve the detection accuracy of anomaly data. However, there is a possibility that the model will overfit due to the reduction of the data in sub-datasets caused by undersampling. Therefore, as illustrated in **Fig. 13**, hyperparameter tuning is employed to create simpler models. These multiple models are combined to form a majority voting ensemble model to avoid overfitting.

3.3 Extraction of coke clogging anomaly factors using machine learning explanation method

The anomaly probability $\Pr[y_j=1\,|\,x_j]$ is obtained by using the operation variable x_j of the analysis target operation j as input to the model m trained as described in the previous section. The contribution degree to a certain operation variable x_{jk} is calculated by considering the degree by which the anomaly probability decreases when x_{jk} is replaced with the average value of normal operation data. By calculating the contribution degree of $x_{j1}, \dots, x_{jp}$, the influence of a single operation variable on the anomaly probability can be evaluated.

Additionally, when the anomaly probability increases with a combination of multiple operation variables, the contribution degree decomposition by SHAP (SHapley Additive exPlanation)[8] is applied. The prediction generated from the proposed machine learning model is difficult to interpret. It is decomposed into the sum of the contribution degrees φ for the respective operation variables, as shown in Eq. (8). This allows for extracting major factor groups according to the relative size of their contribution degrees.

$$\begin{aligned}\Pr[y_j=1\,|\,x_j] &= m(x_{j1}, x_{j2}, \dots, x_{jp}) \\ &= \varphi_0 + \varphi_{j1} + \dots + \varphi_{jp}\end{aligned} \tag{8}$$

3.4 Visualization and recommendation of appropriate action amounts

For the contribution degrees $\varphi_{j1}, \dots, \varphi_{jp}$ of the operation variable x_j of the analysis target operation j, for example, the l operation values with decreasing contribution degrees are selected as the main factor group. The action amount is determined for each of the l main factor group variables as improvement action targets. First, using $p-l$ variables outside the action targets, operations similar to the analysis target operation j are selected from past normal operations. Since the units and dimensions may differ for each operation variable in determining the similarity, each variable is normalized. For example, the L2 norm is calculated as the similarity criteria and the data below the threshold are considered similar operations. **Figure 14**(a) shows a plot example of the analysis target operations and the normal similar operations when $l=2$. Next, as shown in Fig. 14(b), the normal similar operation points are grouped into rectangular regions on the main factor plane. The average pushing load value in the group is displayed on a color scale.

To recommend the appropriate action amount, the standard error of the average pushing load is calculated from the number and variability of similar normal operation data included in these groups. The action amount for the next operation should be the lowest pushing load average value among the groups whose standard error is below the threshold (indicated by a star in Fig. 14(b)).

4. Conclusions

We introduced typical cases where data-driven models based on operation result data have been applied to steelmaking process control and advanced operational support. We have promoted the development of operation result data utilization technologies in combination with process and operational knowledge, and physical models. In this way, we have contributed to improving yield and productivity. The development of artificial intelligence technologies, including machine learning, is remarkable. We will continue to incorporate such innovative technologies to contribute to the advancement of our operations.

References

1) Morita, A. et al.: Nippon Steel & Sumitomo Metal Technical Report. (121), 94 (2019)
2) Morita, A. et al.: CAMP-ISIJ. 26, 854 (2013)
3) Morita, A. et al.: CAMP-ISIJ. 35, 379 (2022)
4) Rolling Theory Committee, ISIJ: Theory and Practice of Plate Rolling
5) Schapire, R.E.: The Strength of Weak Learnability. Machine Learning. 5-2, 197 (1990)
6) Ueda et al.: Analysis of Generalization Error on Ensemble Learning. IEICE Transactions D-2. J80-D-2-9, 2512 (1997)
7) Ota, N. et al.: CAMP-ISIJ. 36, 146 (2023)
8) Lundberg, S.M. et al.: A Unified Approach to Interpreting Model Predictions. Advances in Neural Information Processing Systems. 4765 (2017)

Yasunori KADOYA
General Manager, Head of Dept.
Automation Research Dept.
Intelligent Algorithm Research Center
Process Research Laboratories
20-1 Shintomi, Futtsu City, Chiba Pref. 293-8511

Satoshi KOSUGI
Ph. D., Senior Researcher
Automation Research Dept.
Intelligent Algorithm Research Center
Process Research Laboratories

Akira MORITA
Chief Manager
Digitalization Engineering Section-II
Digitalization Technology Dept.
Systems & Control Engineering Div.
Plant Engineering and Facility Management Center

Shota TAKESHIMA
Senior Manager
Digital Innovation Div.