The Arc Loss Challenge: A Novel Industrial Benchmark for Process Analytics and Machine Learning

Ibrahim Yousef [1a], Lee D. Rippon[a,b], Carole Prévost[c], Sirish L. Shah[d], R. Bhushan Gopaluni[a]

[a]*Department of Chemical and Biological Engineering, University of British Columbia, Vancouver, B.C., Canada*

[b]*Spartan Controls, Burnaby, B.C., Canada*

[c]*BBA Engineering Consultants, Mont-Saint-Hilaire, Q.C., Canada*

[d]*Department of Chemical and Materials Engineering, University of Alberta, Edmonton, A.B., Canada*

Abstract

Rapid development in data-driven process monitoring has provided a rich selection of models and data preprocessing strategies for applications such as fault detection and diagnosis. However, the development, comparison, and selection of process monitoring algorithms can become complicated and unnecessarily onerous. As a result, numerous publicly available benchmark datasets have emerged in the literature. Unfortunately, benchmark literature often suffers from problems such as low fidelity, inconsistent usage, and lack of transparency. This paper presents a benchmark challenge based on a large-scale industrial dataset that aims to enhance the evaluation and comparison of learning algorithms and overall data preprocessing workflows. We introduce the arc loss challenge, a machine learning benchmark with data from a large-scale mining and pyrometallurgy operation. By providing a supervised learning challenge based on large quantities of raw industrial process data with transparent and consistent evaluation procedures, the arc loss challenge is a unique contribution to fault detection benchmarking.

Keywords: Benchmarking, Big data analytics, Fault detection and diagnosis, Machine learning, Process monitoring, Pyrometallurgy

1. Introduction

Benchmarking has allowed the process systems engineering (PSE) community to make valuable contributions toward addressing problems related to process monitoring, including fault detection and diagnosis (FDD). Benchmark datasets are of great value to the community as they form the basis for developing, testing, and comparing different algorithms for process monitoring. In addition, benchmark datasets have played a critical role in advancing the field by enabling direct progress tracking of proposed methods [1].

The literature on data-driven methods for FDD often relies on idealistic simulated datasets to evaluate monitoring algorithms. Unfortunately, this approach makes it difficult to compare the practical utility of different algorithms [2, 3, 4]. In addition, recent advances in computational hardware and algorithmic efficiency have made traditional FDD benchmark datasets, released decades ago, seem "simple and easy to solve." For example, the widely used Tennessee Eastman process (TEP) dataset [2] has been solved with increasingly accurate results (96% to 100%) using advanced machine learning (ML) algorithms [5]. Therefore, a turn toward modern benchmarks incorporating raw industrial data is essential to provide a realistic platform for data-driven process monitoring.

This paper aims to improve process monitoring benchmark research by introducing the arc loss challenge, an ML benchmark with operating data from a pyrometallurgy plant. It consists of one year of daily exports from multiple sources, including process measurements, valve positions, and laboratory measurements. The dataset captures many of the non-trivial challenges faced by industrial practitioners, such as multimodality, class imbalance, and irregular sampling rate. The arc loss challenge also provides the PSE community with an open-source competition for fault detection with standardized evaluation procedures to facilitate the comparison of different fault detection workflows.

The arc loss benchmark dataset can be downloaded from the arc loss challenge website. The website contains the raw data stored in Apache Parquet (.parquet) files and standardized testing software. The testing software is designed to establish consistency, encourage transparency, and prevent over-fitting by providing a rigorous evaluation procedure to compare different monitoring workflows. The main contributions of the arc loss challenge are as follows: i) to provide a publicly available dataset from an industrial process historian, ii) to present a formal procedure for consistently evaluating submissions to track progress in FDD, iii) and to facilitate more impactful research in FDD.

This paper is organized as follows: in Section 2, background information is provided to differentiate the arc loss challenge from existing FDD benchmarks; in Section 3, we present the metallurgical process in question, Section 4 introduces the arc loss challenge, Section 5 describes the dataset, Section 6 provides insights and observations from the industrial data, Section 7 reviews the submission and evaluation procedures and why these procedures were chosen, and Section 8 presents a demonstration of how the benchmark dataset can be used for the development of FDD workflows. Finally, the paper ends with concluding remarks in Section 9.

2. Benchmarking in Process Monitoring

Since its release in 2010, the ImageNet dataset [6] has provided the computer vision community with a benchmark dataset to develop and test proposed models. In turn, the accuracy of the state-of-art model has catapulted from 50.90% to 90.88% in just 11 years [7]. This rapid growth in computer vision was partly enabled by the community-wide adoption of an open-source standardized dataset (i.e., ImageNet dataset). Over the same period, the data available for FDD has grown substantially due to the increased digital transformation of the process industries. This section briefly reviews existing FDD benchmark datasets and places the arc loss benchmark dataset relative to other works.

In general, FDD benchmark datasets fall into two categories depending on the method used for data acquisition. The first category is simulated data or data that is artificially generated to replicate the real-world but with known underlying patterns [8]. Simulated data exists on a spectrum ranging from low fidelity to high fidelity, depending on the degree of precision and realism portrayed in the simulation models [9]. For example, simulation models based solely on first-principles (e.g., TEP [2, 10] and PenSim [11, 12]) exhibit a low level of fidelity. This is because such models tend to omit real-world characteristics and model them as Gaussian random variables or linear piece-wise functions. Next, simulation models for process equipment and instrumentations are derived from the analytical description of the physical behaviour of the equipment, such as thermodynamics and mechanics. For instance, the DAMADICS (Development and Application of Methods for Actuator Diagnosis in Industrial Control Systems) benchmark provides a simulation model of electro-pneumatic actuators in a sugar production process [13]. To more accurately represent the real-world effects, hybrid models integrate measurements from a real process with the simulation models. Hybrid models (e.g., CSTH [14]) provide a platform for FDD applications under realistic conditions of measurement noise, process disturbances and constraints [15].

Real-world data constitute the second category and are obtained from an actual process. Real-world data can be acquired from one or more sources. Possible data sources include readings from physical sensors, alarm events, laboratory results, valve positions, images, and video records. For instance, the PRONTO benchmark [16] includes heterogeneous data from disparate sources in an industrial-scale multi-phase flow facility. In contrast, the 3W benchmark [17] only includes eight process measurements related to the operation of offshore oil wells. The heterogeneity of real-world data presents a number of opportunities and challenges for robust and reliable FDD applications. Figure 1 shows a conceptual comparison of some of the most prominent FDD benchmarks. For a more comprehensive analysis and comparison of accessible benchmarks for process monitoring, readers can refer to the review paper by Melo et al. [18].

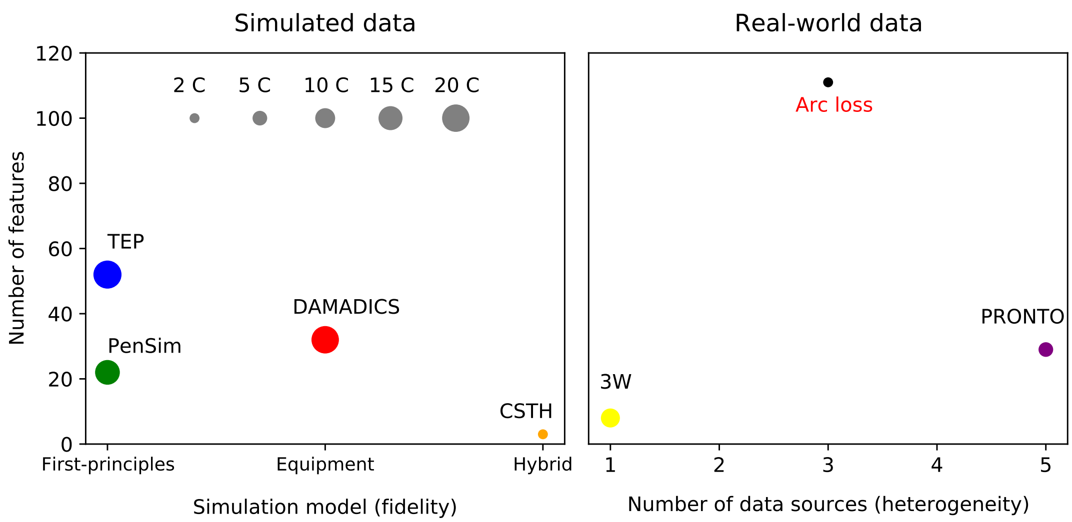

Figure 1: A conceptual comparison of several FDD benchmarks available in the literature. The following criteria are used to compare the benchmarks: 1) the origin of the data (simulations or actual physical sensors), 2) the number of features, 3) the number of classes, 4) the type of simulation models in the case of simulated data, and 5) the heterogeneity of the data in the case of real-world data. The size of blobs is proportional to the number of classes; a legend is displayed in the top left corner, spanning from 2 to 20 classes. Both plots share the same y-axis.

In the PSE community, the first form has become more prevalent than the second. This is due to the difficulty of acquiring high-quality data with recurring and identifiable faults in industrial settings [19]. For instance, i) industrial processes often have an inherently long mean time between failures (MTBF) [20], ii) faults often prevent operation entirely, and iii) most data owners refuse to publish open-source data due to confidentiality and intellectual property concerns [18]. As a result, most faults in existing FDD benchmark datasets are artificially induced by abnormally deviating a characteristic property of the underlying simulation models. In this work, we present a novel benchmark challenge that we believe will have a significant impact on FDD research, similar to the positive impact that the ImageNet challenge had on computer vision. The arc loss benchmark dataset, obtained from three different data sources - process measurements, valve positions, and lab results - falls under the category of real-world data.

3. Pyrometallurgical Smelting Process

The arc loss benchmark dataset was acquired from a large-scale open pit mine and pyrometallurgical plant. The high-level mining and pyrometallurgical operations are depicted in Figure 2. In such processes, high-grade oxidized ore deposits are converted into refined base metals to be processed by shotting and packaging units before being shipped to customers [21]. In this section, the relevant mining and pyrometallurgical operations are described.

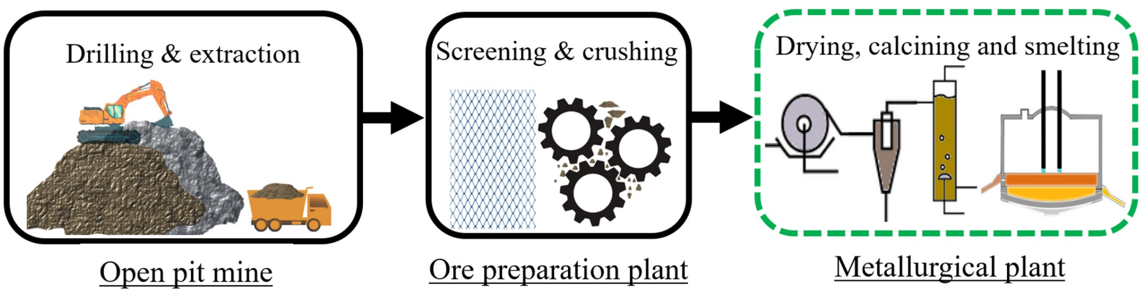

Figure 2: An illustration of the broader mining and metallurgical processes [22].

Saprolite ore is mined from multiple open pits in the massif using hydraulic shovels. The extracted ore is then loaded onto dumper trucks and transported to the ore preparation plant. Due to the friable nature of the run-of-mine ore, waste rocks and undersized particles are removed via crushing and screening operations in the ore preparation plant [23]. The crushed ore is conveyed on an overland belt conveyor to the metallurgical plant, where it is further processed.

Figure 3 shows a simplified schematic of the metallurgical plant in question. Processing operations such as milling, drying, calcining, reduction, and smelting are used in the metallurgical plant. Ore from the preparation plant is wet (i.e., it has a moisture content of up to 40%); therefore, the ore is dried as the first step in the metallurgical plant [24]. Hammer mill flash dryers are used to produce fine ore with less than 1% free moisture content. Next, the dried ore is fed into a series of calciner cyclones operating at 1000 °C where it is dehydrated [25]. The calcined ore is then pre-reduced in fluidized bed reducers using pulverized coal and hot gases at 1000 °C to remove oxides [26]. This step is critical to ensure the operational stability of the subsequent smelting operation in the direct current electric arc furnace (DC EAF) unit. The focus of this challenge is on the operation of the DC EAF utilized as a smelter to refine ores into base metals.

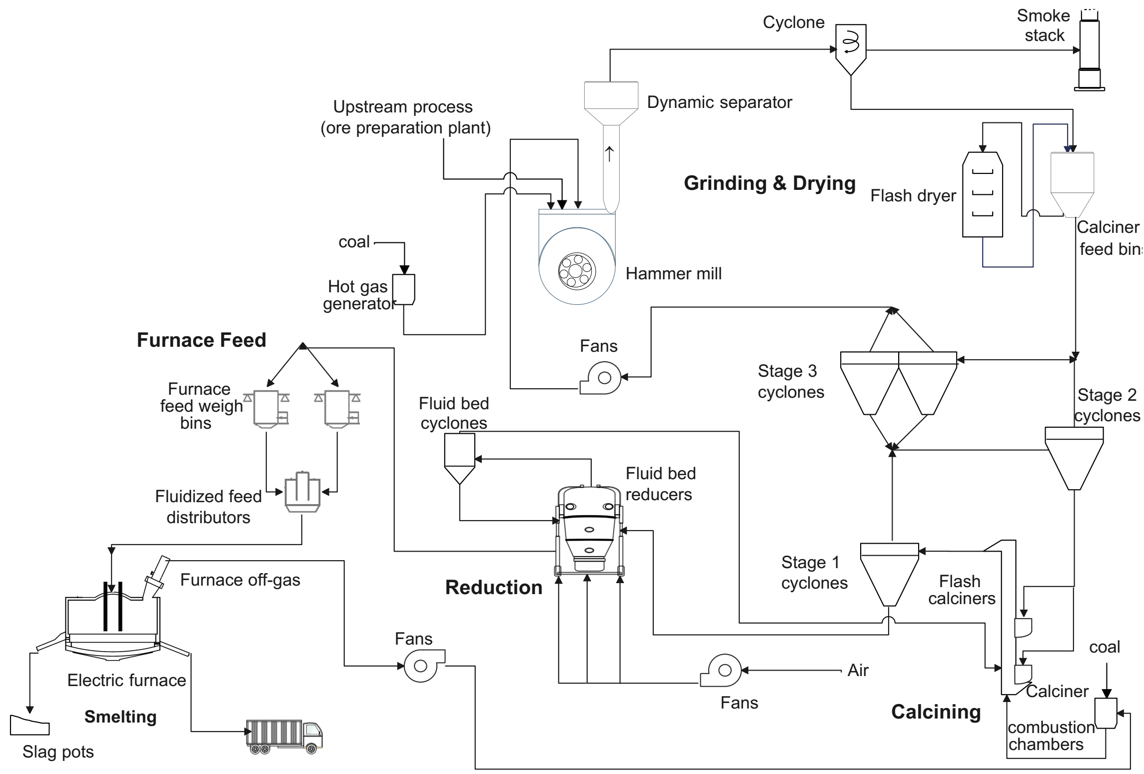

Figure 3: A simplified schematic process flow diagram

A schematic diagram of the DC EAF is illustrated in Figure 4. The smelter unit consists of a refractory-lined cylindrical vessel with water-cooled sidewalls, a conical roof, and a twin hollow graphite electrode located vertically in the center of the roof [27]. The steel vessel contains a molten mixture with a dense metal phase below a lighter slag phase. Refined ore is fed to the furnace through multiple feed ports positioned on the roof using weight bin feeders (c.f., Figure 5). The slag and metal are tapped intermittently from the furnace through launders. Two plasma arcs span from the bottom tip of the electrodes to the top surface of the molten bath, serving as cathode and anode, respectively. The graphite electrodes are connected to a large DC power supply, providing the electrical power required for the DC EAF unit operation [28].

The high-temperature plasma arcs are developed by means of the direct electrical current

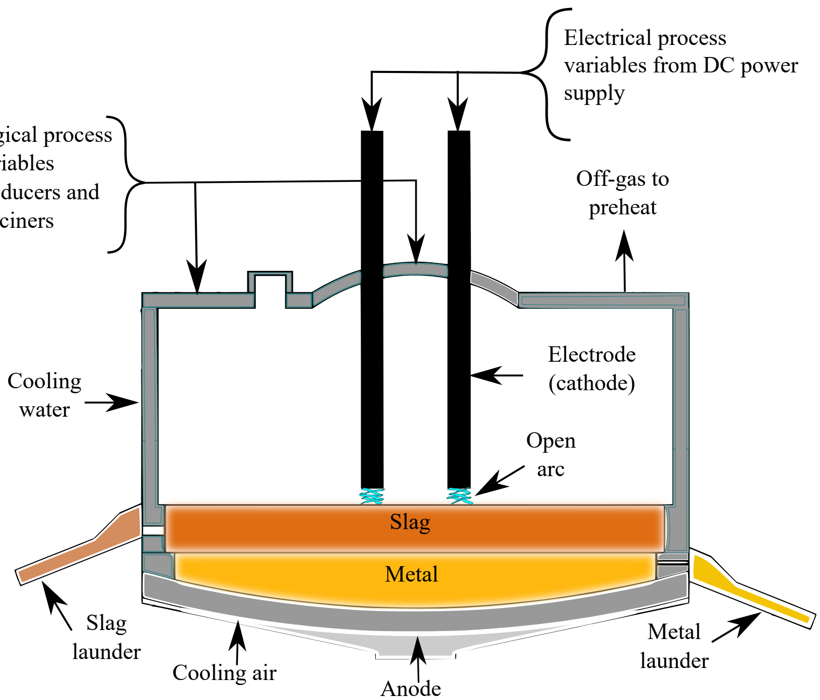

Figure 4: A typical view of the DC EAF unit.

transfer from the cathodes to the anode. The arcs serve as the main heating component in the DC EAF unit [27]. The plasma arcs convert the electrical energy attained from the DC power supply (up to 80 MW) into thermal energy (above 1500 °C) required to maintain the metal-to-slag ratio at desirable operating levels [29]. Since the operation of the DC EAF is directly related to the thermal energy transferred into the molten bath, the presence of the plasma arcs is critical to ensure operational stability and maximum production efficiency. However, the loss of the plasma arc is an unforeseen process fault that adversely impacts the efficiency and stability of the DC EAF unit.

4. Problem Definition for the Arc Loss Challenge

This section defines the process fault and its impact on the operation. In addition, an overview of the principal challenge is provided.

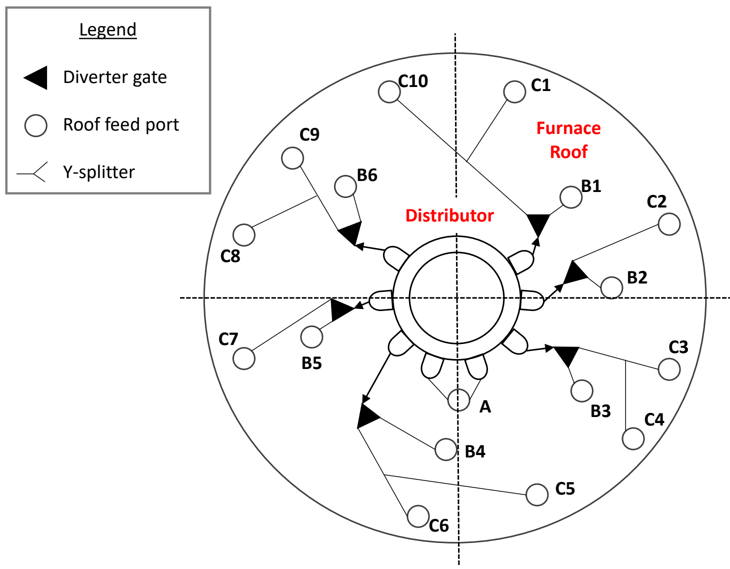

Figure 5: Top furnace view of feed ports.

4.1. The process fault: loss of furnace plasma arc

The arc in the DC EAF is a plasma-based jet with a high temperature and velocity that conducts electricity efficiently [30]. It serves as the primary heating element in the DC EAF, transforming electrical energy from the DC power supply into thermal energy and transmitting it into the molten bath. The DC electric circuit running through the furnace connects the arc and slag bath in series, splitting the total operating voltage between them depending on the electrical properties of the slag (i.e., electrical resistivity) [27].

The slag is often of high electrical resistivity, resulting in a significant voltage drop across the slag layer of the circuit [31]. As a result, the available voltage could be less than the required voltage for the arc to develop, leading to an arc loss event. Other suspected causes of arc loss include upstream process disturbances, electrical disturbances, and extra-long arcs.

Due to the harsh process conditions and safety-critical nature of the operation, obtaining a visual recording of the plasma arc is infeasible in this scenario. Therefore, the daily raw exports have no arc loss labels. Fortunately, unexpected power fluctuations can help define arc loss events. Specifically, a loss of arc in an electrode at time t occurs if and only if the following three conditions related to its power are satisfied: i) a power drop of 10 MW or more relative to the one recorded at $t - 0.6$ min, ii) the power must be steady for approximately 11.25 minutes within a standard variation of 2 MW before the power drop

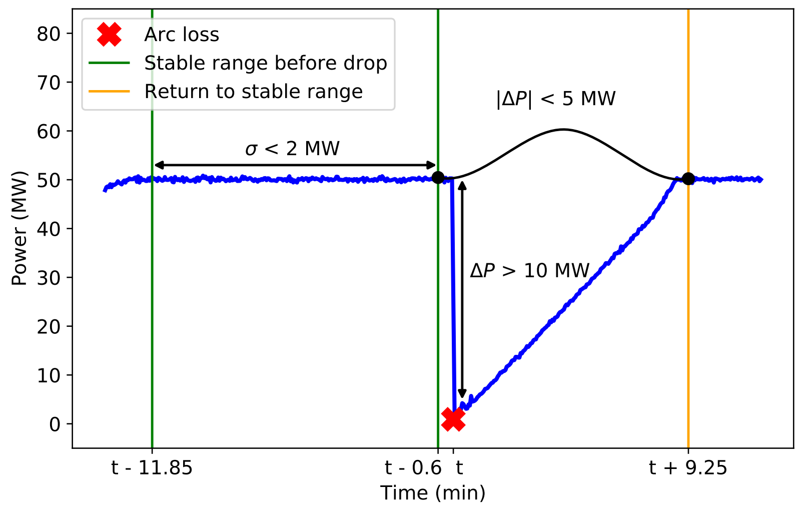

Figure 6: Timeline chart of power and arc loss conditions

(i.e., from $t - 11.85$ min to $t - 0.6$ min), and iii) the power must return to within $\pm$ 5 MW of the initial stable range in approximately 9.85 minutes after the power drop ($t + 9.25$ min) [32]. Figure 6 illustrates the three conditions that constitute an arc loss event. For each sample, these conditions are used to produce output labels that are binary indicators of an arc loss (i.e., $0 \rightarrow$ no arc loss and $1 \rightarrow$ arc loss) in the corresponding electrode.

The unexpected occurrence of arc loss severely impacts the DC EAF. In addition to having an effect on electrical efficiency, recovering from an arc loss often requires a temporary feed reduction and a power increase. Figure 7 illustrates the severity of the arc loss event on the process. It displays a side-by-side comparison of the process measurements recorded during stable and faulty operating regimes. Since the DC EAF operation is directly related to the presence of an open arc, having a reliable predictive alarm that alerts operators of the onset of arc loss would be of great economic and environmental value.

4.2. Arc loss prediction: a supervised learning challenge

The proposed benchmark fuses data preprocessing, feature engineering, supervised learning, and fault detection into a single challenge. Most of the existing FDD benchmarks investigate the use of conventional and advanced ML on clean simulated data. The proposed challenge offers a different form of contribution as it focuses on comparing and validating end-to-end data analytics workflows on raw historical data taken from a large-scale industrial process. Figure 8 illustrates the typical layout of an FDD process analytics

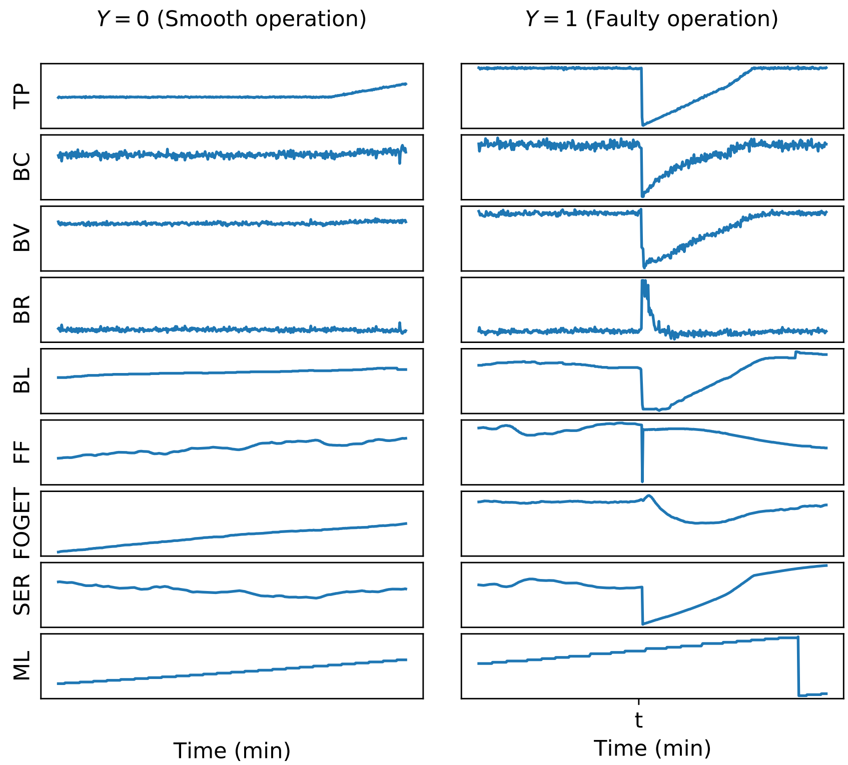

Figure 7: Visual demonstration of the arc loss on the process operation. The left-hand plot represents a relatively stable operation, while the right-hand plot represents a faulty operation. The time t marks the time at which an arc loss event occurred. Both plots share the same y-axis. Readers are referred to Table A1 for variables' descriptions.

workflow. Note that there may also be hidden feedback connections between the modules as such frameworks often progress iteratively.

The objective of this challenge is to utilize information from a full year of high-frequency operating data to correctly predict the onset of an arc loss event (i.e., the target variable Y). For the purpose of this challenge, the target variable Y is defined as an event where either one of the following conditions is satisfied: i) a loss of the arc in electrode A only, ii) a loss of the arc in electrode B only, or iii) arc losses in both electrodes A and B. Mathematically speaking, Y can be described as follows:

$$Y = \begin{cases} 0, & A^{Loss} + B^{Loss} = 0 \\ 1, & A^{Loss} + B^{Loss} > 0 \end{cases} \tag{1}$$

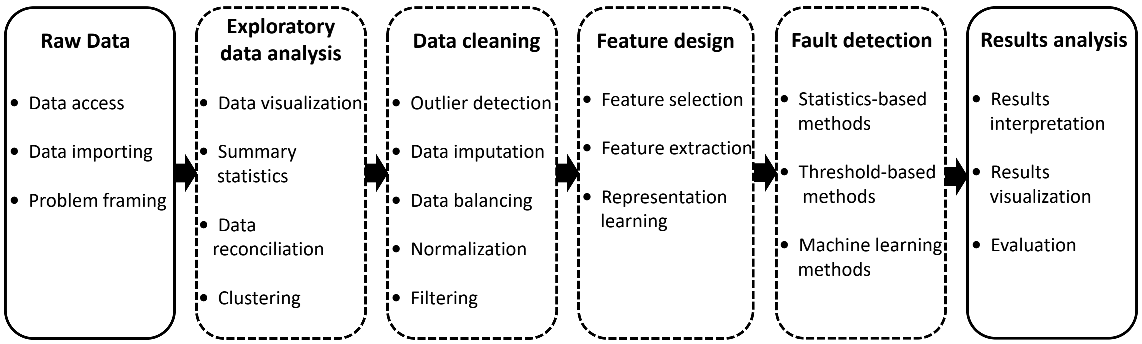

Figure 8: Typical structure of an overall process analytics workflow for FDD.

where A^{Loss} and $B^{Loss} \in [0, 1]$ are binary indicators of an arc loss in electrodes A and B respectively. The target variable Y has been pre-computed as part of the minimal pre-processing performed on the dataset.

5. Industrial Data for the Arc Loss Challenge

The raw data consists of one year of high-frequency operating data collected from various metallurgical process units (e.g., calcining, smelting, reduction, etc.). This section summarizes the process parameters and pre-processing methods carried out to make the data easier to use.

5.1. Data structure

Each raw daily export captures a day of operation in 2022 and is stored with over 200 columns and roughly 30,000 rows. The columns in the daily exports correspond to process variables (PVs) and their associated timestamps. As shown on the left-hand side of Figure 9, the PV sampling rates vary. The number of samples collected from a particular PV throughout the day is represented by the number of rows (i.e., the height of blue and red bars). Some PVs have high-frequency measurements (e.g., a sample every three seconds), while laboratory measurements can have sampling periods greater than three hours. Overall, the total daily exports have an uncompressed size of 17.4 gigabytes (GB), making the data unwieldy for practitioners to load, analyze, and process.

In addition to the varying sampling rates, the raw data is riddled with other problematic artifacts such as bad inputs, outliers, and irrelevant or misleading data. The arc loss benchmark is concerned with workflows that take in raw data and yield models that pro-

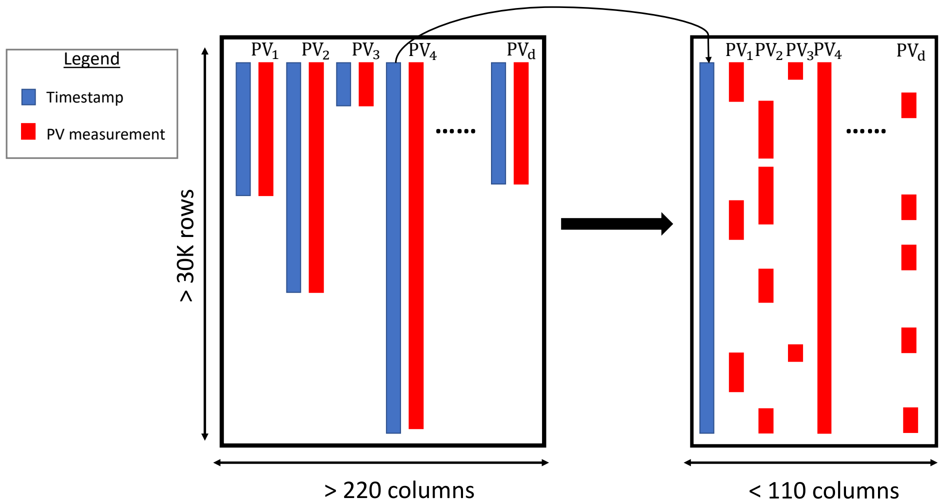

Figure 9: Illustration of structuring a raw daily export

vide operational insights and predictions. Therefore, the arc loss dataset published with this paper is available in a raw form with minimal pre-processing to structure the data.

To make the data easier to utilize, we pre-processed all raw daily exports to follow a structured format where all PVs have a unified timestamp. As illustrated in Figure 9, data structuring was performed by identifying the most frequently sampled PV and using the corresponding timestamp as a reference to reindex the remaining PVs. The unified sampling frequency is three seconds. Additionally, all PVs with categorical values were replaced by numerical indicators (e.g., close and open became 0 and 1, respectively). Finally, problematic entries (e.g., tag not found, access denied, bad data, etc.) were substituted with not a number (NaN) values.

Overall, the data consists of a time index and 111 PVs. Although it is too much information to present here, the tag, description, range, and unit of measurement for each PV are provided in Table A1. There are 92 PVs with high-frequency samples representing physical process properties (e.g., feed rate, temperature, etc.), 14 variables with discrete values (label encoded) that correspond to valve positions, and five laboratory measurements with a sampling period greater than one hour.

5.2. Accessing the arc loss data

For the purpose of this challenge, the data is chronologically split into training/validation (Jan.-Oct.) and testing (Nov.-Dec.) sets. The daily exports are bundled and stored in two compressed (.parquet) files. Participants are allowed to use any data in the training/validation set for tuning the hyperparameters of their candidate model. The arc loss challenge data can be downloaded from here in the form of (.parquet) files:

1. `trainval.parquet`: This 1.21 GB dataset (Jan.-Oct.) is intended for training and validating the process monitoring workflows
2. `test.parquet`: This 0.60 GB dataset (Nov.-Dec.) is for testing the monitoring algorithm and benchmarking with performance metrics.

6. Data Exploration and Analysis

Process knowledge and data exploration can guide participants to discover valuable information for their analysis. This section provides insightful summaries and observations regarding the industrial data and the underlying process. The content provided in this section was collected from process experts and previous data analysis.

6.1. Dataset statistics

Table 1 summarizes relevant statistics for the arc loss benchmark dataset. Figure 10 shows the daily arc loss events over an entire year of operation. It can be observed that arc loss faults are most frequent from May through September. Additionally, Figure 11 presents a circular graph for the arc loss rate in June. The chart displays a time series of 30 days, starting from June $1^{\text{st}}, 2022$ at 00:00 at the innermost circle, and moving clockwise until reaching June $30^{\text{th}}, 2022$ at 23:00 at the outermost ring. Dark blue is used to indicate periods with a higher fault rate, while light blue denotes periods with the fewest faults. Both Figures 10 and 11 imply that the process faults are characterized by randomness, seasonality, and variability, making fault detection a non-trivial task.

Table 1: Summary of arc loss dataset statistics.

Total number of samples	10483200
Total number of PVs	111
Sampling period	3 sec
Total number of categorical variables	14
Training/validation set ratio	83.5% (Jan.-Oct.)
Testing set ratio	16.5% (Nov.-Dec.)
Number of classes	2
Class ratio (normal:faulty)	3337:1

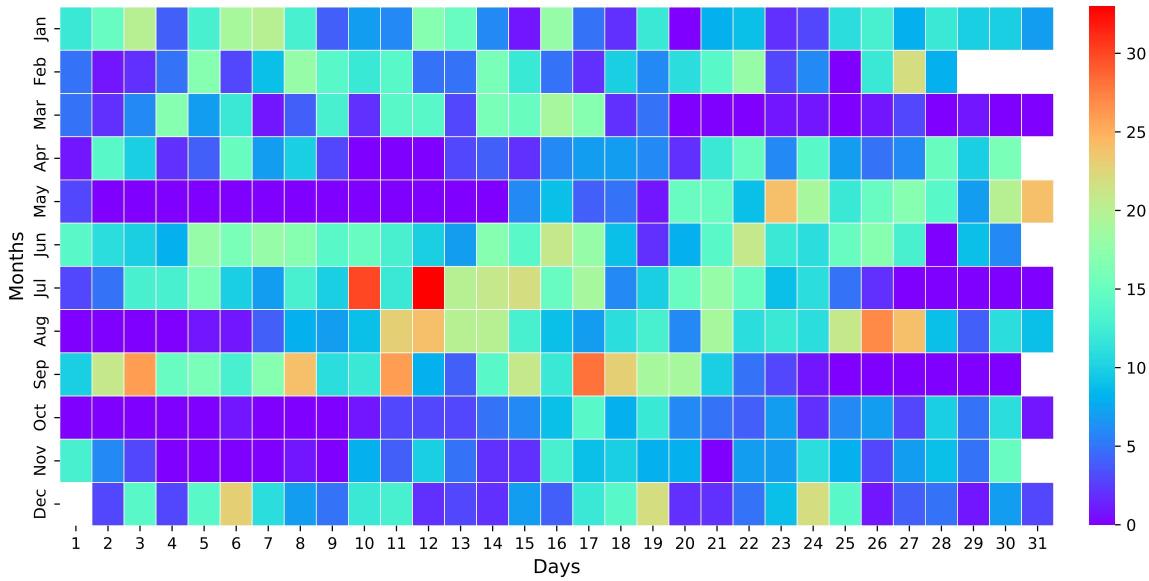

Figure 10: The number of arc loss events per day throughout a year of operation

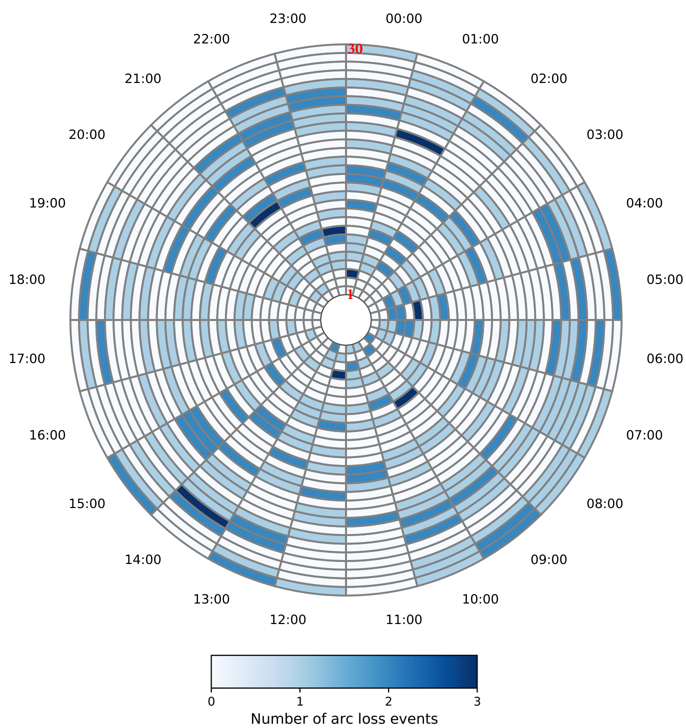

Figure 11: The distribution of arc loss events over operating hours in June

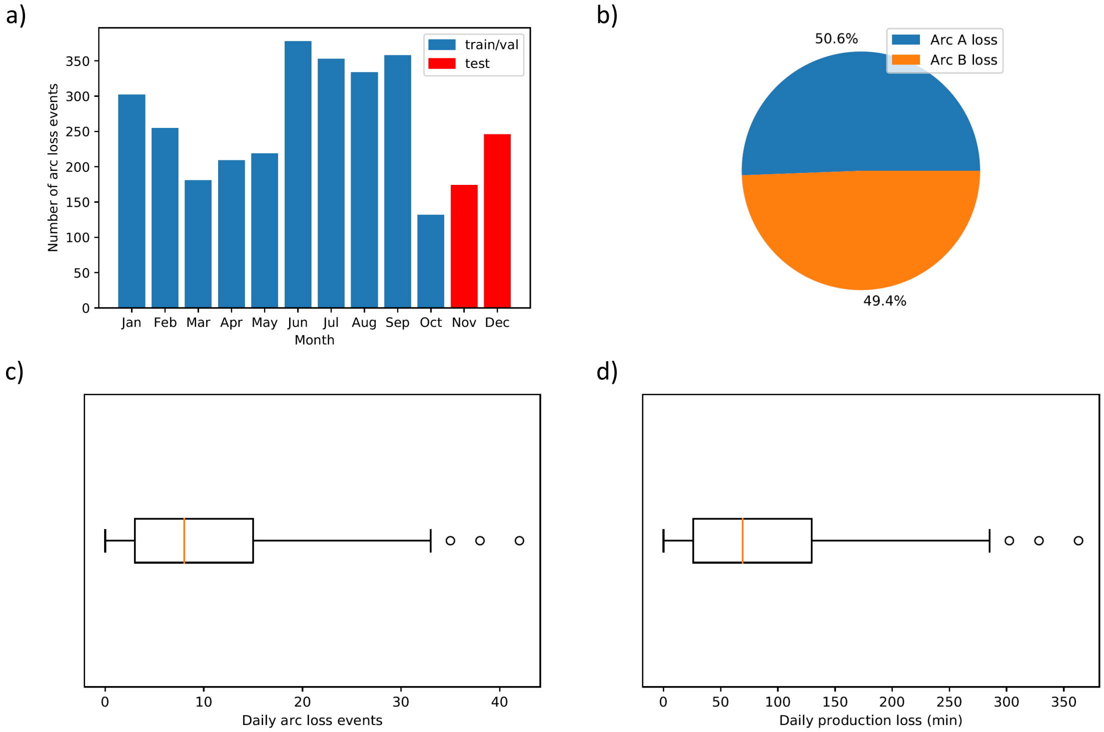

Figure 12: Statistics of the arc loss benchmark dataset.

In one year of operating data, the furnace experiences 3,141 arc loss events. As illustrated in Figure 12a, the month that saw the most arc loss events was June, closely followed by September, and October was the month that witnessed the fewest arc loss events.

Figure 12b shows that both subclasses (arc loss A and arc loss B) that constitute the faulty class (Y=1) are represented approximately equally in the arc loss data. This indicates that the arc loss event is independent of which electrode is operating. Even though arc loss events typically last less than one minute, a single arc loss fault can cause up to 10 minutes worth of production loss involving significant material and energy inefficiencies. On average, 9.4 arc loss events occur daily, resulting in 82 minutes of lost production as shown in Figures 12c and d, respectively.

6.2. Process shutdowns

A shutdown is a period of time during which a process is taken from a normal to an idle state of operation to carry out all necessary maintenance. Even though shutdowns are scheduled in advance, their duration is not known *a priori* with high certainty. If samples corresponding to shutdown periods are not carefully imputed, the trained predictive models could be biased toward an irrelevant process state. Data from shutdowns do not possess any meaningful information from the process perspective. Therefore, it is essential to identify shutdown periods to prevent model degradation.

The selection of relevant PVs for shutdown identification is a non-trivial problem and often requires the support of operators or domain experts. According to process experts, a time period $[t_a, t_b]$ corresponds to a process shutdown period if and only if the total power (TP) drawn from both electrodes A and B are less than 10 MW (AP + BP $\leq$ 10 MW) for at least ten hours ($t_b - t_a > 10$ hrs) as shown in Figure 13. Participants are encouraged to use their judgment and select any strategy they deem appropriate for dealing with shutdown data.

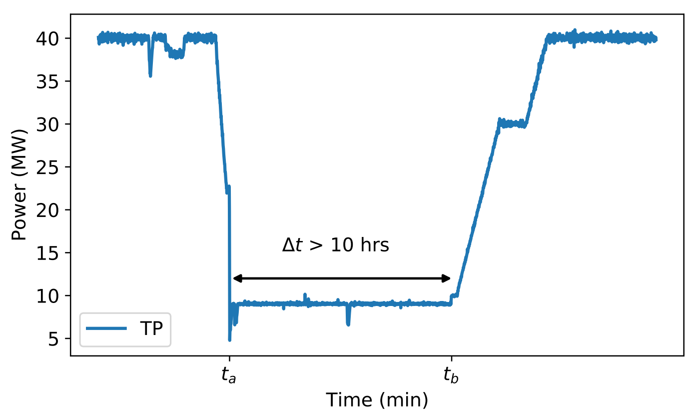

Figure 13: Illustration and quantitative definition of shut down periods.

6.3. Sampling frequency

Statistical process monitoring methods for fault detection are generally designed for uniformly sampled data. However, like any other industrial process, the pyrometallurgy process in question is a multi-rate sampling system. In other words, PVs are measured in a non-uniform fashion and differ in sampling rate. The variables with the highest sampling rates are primarily temperature-related, while those with the lowest are laboratory measurements.

Fusing multi-rate sampled process measurements has resulted in a significant amount of missing data. Missing data are encoded as blanks or not a number (NaN) values. In fact, 38 PVs have over 90% of their values missing as demonstrated in Figure 14. The predictive

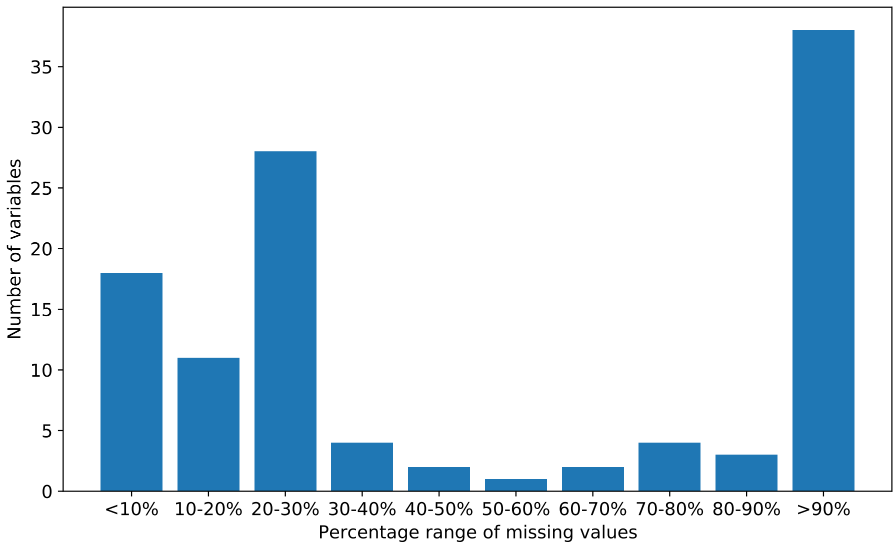

Figure 14: The distribution of missing values for PVs.

power and quality of a data-driven model can be adversely impacted by training it on data with lots of missing values. Hence, it is important to address the missing data problem. Participants are free to employ any data imputation technique to tackle the missing values problem.

6.4. Observations from the data

Several observations from the dataset that might be of interest to participants when developing their monitoring algorithms are provided in the following list:

- December 1^{st} records are missing.
- December 4^{th} has the highest percentage of missing values, with over 83%.
- W1F, W2F, W3F, W6F, W7F, W8F, and TMF measurements are unreliable due to instrumentation errors (i.e., incorrect calibration, malfunction, etc.).
- The minimum time interval between two successive arc loss events is 10 minutes.
- PV data does not always lie within the range provided by process experts and listed in Table A1. For instance, almost all crucible heat loss data are outside the limit as shown in Figure 15. Participants may have to evaluate data quality and perform sanity checks to improve the reliability and robustness of monitoring algorithms (i.e.,

relying on process knowledge for data cleaning is not always an infallible strategy).

- Different PVs are highly coupled and may exhibit strong cross-correlations. A PV may also have an autocorrelation with its previous values. These correlations reveal the underlying characteristics of the process and they may change during abnormal operations (e.g., arc loss). The electrical parameters pertaining to electrode A are depicted in Figure 16. The orange plot represents the measured voltage, while the blue plot corresponds to the theoretical voltage computed using Ohm's law. Such theoretical relationships can be used for data reconciliation to enhance data quality.

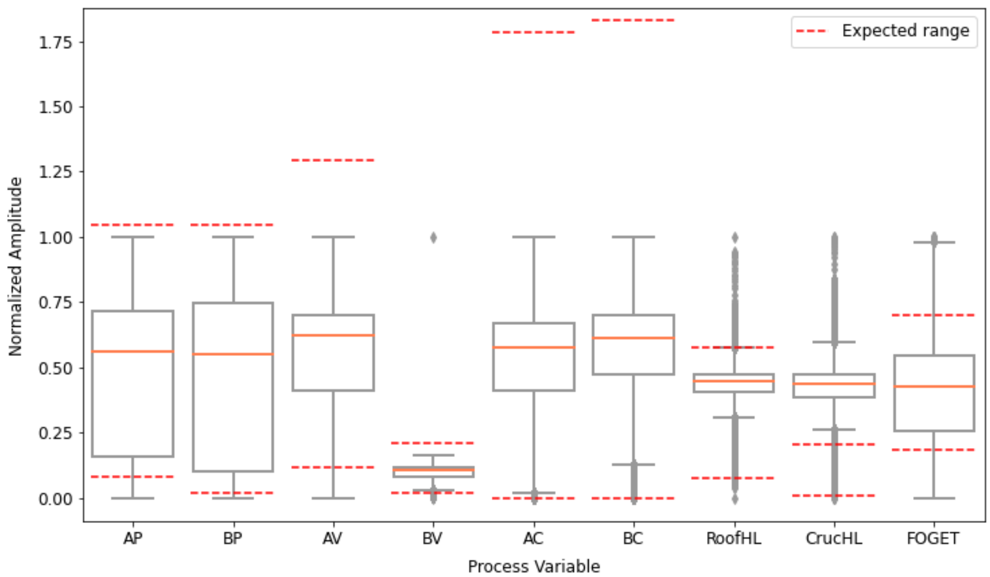

Figure 15: Box plots illustrating PV data distributions. Expected range lines are shown for each PV, normalized to be consistent with the normalized values.

7. Submission and Evaluation

This section outlines the procedures for submitting entries to the benchmark challenge and discusses the evaluation process. The submission and scoring procedures for the arc loss challenge are designed to ensure fairness through transparency and consistency. The evaluation also aims to prevent over-fitting by emphasizing soft sensor durability over two consecutive months of unseen operating data. The following section explains how submissions are scored. We encourage participants to post their questions and queries in our community discussion forum.

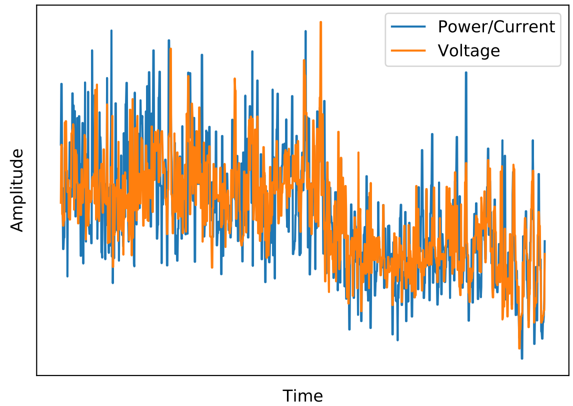

Figure 16: Observations from the benchmark dataset: comparing the actual vs. theoretical voltage drawn by electrode A.

7.1. Submission of entries

The arc loss challenge is divided into two stages. Firstly, participants should use the *labeled* training/validation set to develop a monitoring workflow, optimize hyperparameters, and estimate the performance measures. In the second stage, participants are required to run their workflow on the *unlabeled* testing set (Nov. and Dec. data) to classify whether each timestamp in the testing set corresponds to stable operation ($Y = 0$) or an arc loss event ($Y = 1$). More specifically, participants are required to only submit the timestamps that belong to the faulty class ($Y = 1$). The scoring metric (described in subsection 7.2) is computed specifically on the instances that are important for the problem, i.e., identifying arc loss events. It is worth noting that since the minimum time interval between two consecutive arc loss events is 10 minutes, submitted timestamps should be at least 10 minutes apart. Entries must be submitted as a `.zip` file containing the following:

- A one-column `.csv` file containing the predicted timestamps of arc loss events. The entries should be in the form of (yy-mm-dd hh:mm:ss) (e.g., 22-12-01 14:04:12)

To reduce the likelihood of over-fitting, we will not release the ground-truth labels of the testing set until the challenge is over. Releasing the true testing labels could lead to some "optimistic" results, where participants test several configurations on the testing set and only report the best result. This risk is present in any benchmarking initiative where the ground truth is made available.

Participants may have a maximum of 20 attempts over the course of the competition. Upon submission, participants will receive their performance score on the held-out testing set. The arc loss challenge comprises an unofficial phase and an official phase. During the unofficial phase, participants may receive scores for up to five attempts. This will allow participants to examine the testing software and compare different methodologies without being officially counted for the competition. In the official phase, participants are allowed to submit up to 15 attempts, and the best-performing attempt will be officially counted. Note that unused attempts during the unofficial phase can not be carried over to the official stage. Attempts that can not be scored due to incompleteness (e.g., improper formatting, missing components, etc.) will not be counted toward the attempt limits. The challenge deadlines are listed in Table 2 (check our website for further updates). The data will be publicly available for approximately a year before the submission deadline. This will allow substantial time for data cleaning and ensure that participants with limited access to computational resources have an equal chance of winning. After the competition deadline, the fully labeled data will be made publicly available to facilitate the development of data-driven process monitoring algorithms by the community.

Table 2: Challenge deadlines

Phase	Start date	Deadline	Entry limit
Unofficial phase	June 1$^{\text{st}}$, 2023	August 31$^{\text{st}}$, 2023	5
Official phase	September 1$^{\text{st}}$, 2023	July 1$^{\text{st}}$, 2024	15

7.2. Scoring

Evaluating results on the two-class arc loss dataset poses several challenges, including the non-uniform prior distribution across the classes, with $P(Y = 0) > 0.9997$ and $P(Y = 1) < 0.0003$. For this reason, a simple performance measure like accuracy is not proper because it fails to distinguish between the number of correctly classified samples of different classes. Moreover, the cost of misclassifying a positive sample (i.e., a false negative or missed alarm) is greater than the cost of incorrectly classifying a sample from the negative or majority class (i.e., a false positive or false alarm). These factors must be taken into account when designing a suitable scoring metric. In addition to these challenges, it is critical to predict arc loss events before they occur to enable operators to take preventive measures. Hence, a novel scoring metric is designed to address the imbalanced classes and associated costs, reward early predictions, and penalize late predictions.

For this challenge, a binary prediction $\hat{Y}(t)$ is made for each time step t, i.e., :

$$\hat{Y}(t) \begin{cases} \hat{Y}_N(t) = 0, \text{a negative prediction at time } t \text{ during a normal operating period (N)} \\ \hat{Y}_N(t) = 1, \text{a positive prediction at time } t \text{ during a normal operating period (N)} \\ \hat{Y}_F(t) = 0, \text{a negative prediction at time } t \text{ during a faulty operating period (F)} \\ \hat{Y}_F(t) = 1, \text{a positive prediction at time } t \text{ during a faulty operating period (F)} \end{cases}$$

Figure 17 illustrates the scoring functions assigned to each $\hat{Y}(t)$. During faulty operating periods (F) (i.e., $t_{early} \leq t \leq t_{late}$), early arc loss detection is crucial. Therefore, arc loss predictions made within 7 mins before the onset time of arc loss t_{loss} are rewarded with a maximum reward of (+1.0) given at $t_{optimal}$ = t_{loss} – 5 mins. However, positive predictions that are more than 7 mins before t_{loss} (i.e., t < t_{early}) are penalized (-0.5). Additionally, late arc loss predictions are unhelpful, and missed arc loss predictions are significantly harmful. To reflect this, arc loss predictions made up to 2 minutes after t_{loss} (i.e., t > t_{late}) are slightly penalized (-0.1), while no arc loss predictions or missed alarms are heavily penalized (-2.0).

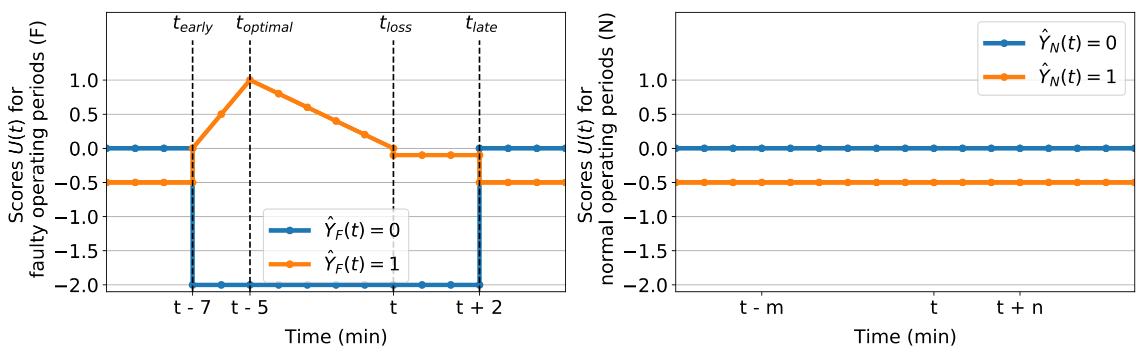

Figure 17: The scoring functions for negative and positive predictions during normal (right) and faulty (left) operating periods. True positives are rewarded based on their time stamps, while false negatives (i.e., models that fail to predict arc loss before t_{late}) receive a penalty of -2.0. Late predictions (t_{loss} < t < t_{late}) are slightly penalized with -0.1, and false positives are penalized with -0.5. True negatives are neither rewarded nor penalized.

During normal operating periods (N), when no arc loss events occur within 10 mins (i.e., $t < t_{early}$ or $t > t_{late}$), false alarms can lead to decreased confidence in monitoring models and poor allocation of operational attention and resources. To address this, false alarms (i.e., $\hat{Y}_N(t) = 1$) are penalized with a score of -0.5. However, negative predictions (i..e, $\hat{Y}_N(t) = 0$) are neither rewarded nor penalized.

The total score of a model's predictions is calculated by summing the scores across all time steps t:

$$U_{total} = \sum_{t \in T} U(t) \tag{2}$$

To facilitate interpretation, we normalize Equation 2 such that all scores fall within the range [0 - 1], using the following Equation:

$$U_{normalized} = \frac{U_{total} - U_{inactive}}{U_{optimal} - U_{inactive}} \tag{3}$$

where $U_{optimal}$ denotes the unnormalized optimal score while $U_{inactive}$ denotes the unnormalized score for a completely inactive classifier (no positive predictions). The highest possible normalized score $U_{normalized}$ is 1, which is awarded to the optimal model that accurately predicts all arc loss events with a 5-min warning and generates no false alarms. Conversely, an inactive model that only outputs negative predictions and does not detect any arc loss events will receive a normalized score $U_{normalized}$ of 0. The winner of the challenge will be determined based on the model with the highest normalized score $U_{normalized}$ on the unseen testing set.

8. Demonstration of an Arc Loss Prediction Workflow

This section provides a demonstration example of the arc loss prediction workflow using the benchmark dataset. The example serves as a baseline framework for the arc loss challenge. The demonstration starts with an overview of the methods used to clean the raw data. It is followed by a description of the predictive model used for workflow development and its performance evaluation. Participants can use the baseline Python implementation available on the challenge website as a template for developing their own submissions.

8.1. Data preprocessing

Data preprocessing is a crucial procedure that involves cleaning and preparing data for analysis and modelling. The quality of data preprocessing can significantly impact the generalization performance and effectiveness of the final model or analysis. When done properly, data preprocessing can enhance the accuracy of the analysis and improve the model's performance. However, inadequate data preprocessing can lead to inaccurate or biased results. In this demonstration example, data preprocessing tasks include removing outliers, handling shutdown data, imputing missing data, and scaling or normalizing data.

Outliers are data points that lie far away from the rest of the data points in a dataset. To remove outliers from the raw dataset, the Z-score method was used. This method involves calculating the Z-score for each data point, which is the number of standard deviations away from the mean that the point lies. Data points with a Z-score greater than a specified threshold were considered outliers and were replaced with NaN (Not a Number). The threshold value used in this method was set to 3. Next, the data from shutdown periods (i.e., periods of zero or low activity) were carefully handled using the zero-imputation technique. It involves identifying periods of time during which power values are below a certain threshold for more than a certain time threshold (refer to section 6.2) and replacing these values with zeros. To address missing data (represented by NaN values), a forward-fill operation was performed. This method involves filling missing values with the last measured value. However, in cases where missing values occur at the beginning of the time series with no previous measurements, a backward-fill approach was used, where missing values were filled with the first measured value in the time series. Finally, the data were scaled to have zero mean and unit variance. The scaling was performed by subtracting the mean and dividing it by the standard deviation of the training data.

8.2. Arc loss prediction: baseline performance

Given the binary characteristic of the problem at hand, logistic regression was chosen as the natural baseline model to demonstrate the feasibility of our benchmark in validating data-driven process monitoring workflows. Logistic regression is a widely used and well-established algorithm for binary classification tasks, providing an effective and interpretable starting point for our analysis. The logistic regression model estimates the probability of a time stamp belonging to the faulty class ($Y = 1$), using a logistic or sigmoid function. This probability is then converted into a binary prediction by applying a threshold, typically 0.5. To address the class imbalance, cost-sensitive learning was used. This involves assigning different costs to both classes during the training process inversely proportional to their frequencies. Specifically, misclassifying faulty class instances (false negatives) has a higher cost than misclassifying normal class instances (false positives). A hold-out strategy was employed for model evaluation. The train/val set was split into two subsets: a training set comprising January-August data used for training the model, and a validation set comprising September-October data used for validating the models' hyper-parameters. Next, a random search over a manually predefined search space is performed to find a well-performing model configuration. Finally, the model with the hyper-parameter configuration that achieved the best performance on the validation set during the random search was tested on the testing set.

Table 3 summarizes the performance evaluation of the baseline model reported on the testing set. To assess the model's performance, we use three key terms: true positive (TP), false negative (FN), and false positive (FP) predictions. In our case, we define a TP prediction as a positive prediction made between the time interval of t_{early} and t_{loss}. On the other hand, an FN prediction is a negative prediction made during $t \in [t_{early} - t_{loss}]$. Finally, an FP is a positive prediction made at $t \notin [t_{early} - t_{loss}]$. To help illustrate a TP prediction made by the baseline model on the testing set, readers are referred to Figure 18.

Table 3: Performance evaluation of the baseline workflow. Abbreviations: TP- true positives, FP - false positives, FN - false negatives.

	TP	FP	FN	Precision	Recall	F_2	$U_{normalized}$
Baseline score	247	499	173	0.3311	0.5881	0.5091	**0.2077**

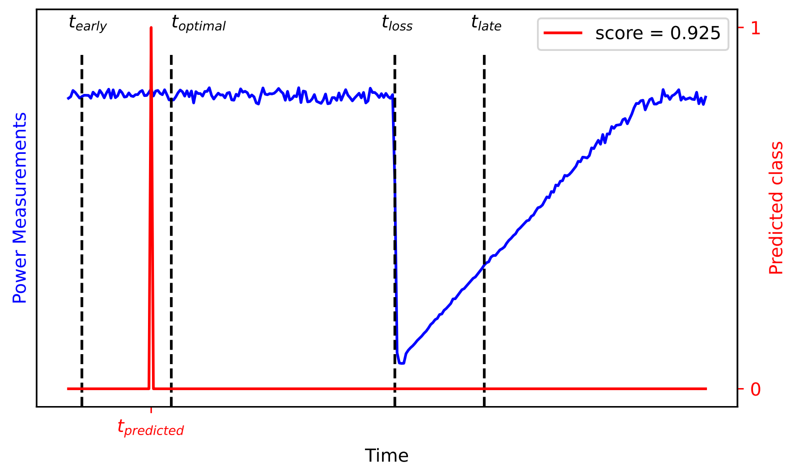

Figure 18: An illustration of a TP prediction.

The baseline workflow presented in this section is intended to serve as a starting point rather than a complex solution. Participants are encouraged to use it as a reference point to measure and compare their own workflows. The results of the baseline demonstrate the suitability of the benchmark dataset for testing data-driven process monitoring workflows. The well-defined problem and well-documented faults provide a suitable testing ground. For more complex solutions, readers can refer to our previous work [32] where traditional and contemporary approaches to representation learning and binary classification were compared in a comprehensive analysis for their ability to predict arc loss. Additionally, Table 4 summarizes various approaches for addressing different challenges that can be tested and validated using this benchmark dataset. By exploring these approaches, researchers can develop a deeper understanding of the strengths and limitations of different methods, and develop more effective and robust monitoring workflows.

Table 4: Comparison of baseline and suggested methods for various challenge tasks: a guide for participants

Tasks	Baseline approach	Suggested methods
Outliers removal	Z-score method	Interquartile range method [33] Distance-based methods [34, 35] Domain knowledge (PV limits)
Shutdown data handling	Zero-imputation	Exclusion (removal from analysis) Integration (include in the analysis) Segmentation (separate analysis)
Missing data imputation	Forward/ backward filling	Regression imputation [36, 37] K-nearest neighbor imputation [38] DL imputation [39, 40]
Feature selection/ extraction	N/A	Dimensionality reduction [41] Embedded methods [42] Deep representation learning [43]
Fault detection & diagnosis	Linear ML model	Threshold-based methods [44] Statistics-based methods [45] DL models [46, 47]

9. Conclusion

The arc loss benchmark dataset is a valuable resource that contains a large amount of historical industrial data from a large-scale metallurgical process with an unexpected process fault, namely an arc loss. The dataset spans an entire year of operation data collected from various metallurgical process units with fast sampling rates. The paper describes the process in detail and presents an overview of the data characteristics, such as data types, variables, and sampling rates. Additionally, the paper introduces the arc loss challenge, which is an open-source challenge that provides the community with a standardized dataset and evaluation framework for comparing different industrial fault detection methods. The primary objective of the arc loss challenge is to catalyze research in the fast-growing field of process monitoring and FDD, and to investigate whether the success of deep learning and ML in computer vision and natural language processing can be replicated in the process industries. The authors believe that the arc loss benchmark dataset will be a valuable asset to the FDD community and will inspire new ideas for real-world industrial data applications.

Acknowledgment

The authors thank BBA Engineering Consultants, the National Science and Engineering Research Council of Canada and the Izaak Walton Killam Memorial Fund for funding this research through an Engage grant and a Killam Pre-Doctoral Memorial Fellowship. This work is supported in part by the Institute for Computing, Information and Cognitive Systems (ICICS) at UBC.

Appendix A. Process Variable Information

Table A1: Overview of process parameters.

	Tag	Description	Range	Unit
Progress parameters				
1.	t	Time [yyyy-mm-dd hh:mm:ss]		
Smelting parameters				
2.	AP	Electrode A power	0-60	MW
3.	BP	Electrode B power	0-60	MW
4.	TP	Total power (AP+BP)	0-100	MW
5.	APSP	Electrode A power set point	0-60	MW
6.	BPSP	Electrode B power set point	0-60	MW
7.	AC	Electrode A current	0-100	kA
8.	BC	Electrode B current	0-100	kA
9.	ACSP	Electrode A current set point	0-100	kA
10.	BCSP	Electrode B current set point	0-100	kA
11.	AV	Electrode A voltage	<2200	Volts
12.	BV	Electrode B voltage	<2200	Volts
13.	AVSP	Electrode A voltage set point	≤ 2200	Volts
14.	BVSP	Electrode B voltage set point	≤ 2200	Volts
15.	AR	Resistance around electrode A	≥ 0	mΩ
16.	BR	Resistance around electrode B	≥ 0	mΩ

To be continued

Table A1 (continued)

	Tag	Description	Range	Unit
17.	ARSP	Resistance around electrode A set point	≥ 0	mΩ
18.	BRSP	Resistance around electrode B set point	≥ 0	mΩ
19.	SER	Specific Energy Ratio	410-440	W/ton
20.	AL	Arc A length		mm
21.	BL	Arc B length		mm
22.	CrucHL	Crucible (the wall) heat loss	0-5	MW
23.	RoofHL	Roof heat loss	0-5	MW
24.	PCHL	Plain cooler heat loss	≥ 0	MW
25.	UWCHL	Upper chilled water heat loss	≥ 0	MW
26.	LWCHL	Lower chilled water heat loss	≥ 0	MW
27.	HFansHL	Hearth (Fans) heat loss	≥ 0	MW
28.	HTCsHL	Hearth (Technological Control System) heat loss	≥ 0	MW
29.	SL	Slag level		mm
30.	ML	Metal level		mm
31.	FOGET	Off-gas temperature	180-630	°C
32.	TPA	Slag tap A valve opening	0-100	%
33.	TPB	Slag tap B valve opening	0-100	%
34.	CO2	CO_2 volume	0-25	%
35.	ST	Slag temperature after being tapped		°C
Furnace feed parameters				
36.	FF	Furnace feed rate	0-200	tons/hr
37.	FFDiv5APos	Furnace Feed Inlet Diverter 5A Position 1	[0,1]	
38.	FFDiv5BPos	Furnace Feed Inlet Diverter 5B Position 1	[0,1]	
39.	FFDiv5CPos	Furnace Feed Inlet Diverter 5C Position 1	[0,1]	
40.	FFDiv5DPos	Furnace Feed Inlet Diverter 5D Position 1	[0,1]	
41.	FFDiv5EPos	Furnace Feed Inlet Diverter 5E Position 1	[0,1]	
42.	FFDiv5FPos	Furnace Feed Inlet Diverter 5F Position 1	[0,1]	
43.	W1OC	Weir 1 valve opening	[0,1]	
44.	W2OC	Weir 2 valve opening	[0,1]	
45.	W3OC	Weir 3 valve opening	[0,1]	
46.	W4OC	Weir 4 valve opening	[0,1]	
47.	W5OC	Weir 5 valve opening	[0,1]	
48.	W6OC	Weir 6 valve opening	[0,1]	
49.	W7OC	Weir 7 valve opening	[0,1]	

To be continued

Table A1 (continued)

	Tag	Description	Range	Unit
50.	W8OC	Weir 8 valve opening	[0,1]	
51.	W1F	Weir 1 flow rate	0-200	tons/hr
52.	W2F	Weir 2 flow rate	0-200	tons/hr
53.	W3F	Weir 3 flow rate	0-200	tons/hr
54.	W6F	Weir 6 flow rate	0-200	tons/hr
55.	W7F	Weir 7 flow rate	0-200	tons/hr
56.	W8F	Weir 8 flow rate	0-200	tons/hr
57.	W1	Weir 1 flow rate	≥ 0	mA
58.	W2	Weir 2 flow rate	≥ 0	mA
59.	W3	Weir 3 flow rate	≥ 0	mA
60.	W4	Weir 4 flow rate	≥ 0	mA
61.	W5	Weir 5 flow rate	≥ 0	mA
62.	W6	Weir 6 flow rate	≥ 0	mA
63.	W7	Weir 7 flow rate	≥ 0	mA
64.	W8	Weir 8 flow rate	≥ 0	mA
65.	TMF	Total microwave flow rate	≥ 0	mA
66.	W1T	Feed temperature after leaving weir 1		°C
67.	W2T	Feed temperature after leaving weir 2		°C
68.	W3T	Feed temperature after leaving weir 3		°C
69.	W4T	Feed temperature after leaving weir 4		°C
70.	W5T	Feed temperature after leaving weir 5		°C
71.	W6T	Feed temperature after leaving weir 6		°C
72.	W7T	Feed temperature after leaving weir 7		°C
73.	W8T	Feed temperature after leaving weir 8		°C
74.	IT1	Feed temperature in the distribution bin 1		°C
75.	IT2	Feed temperature in the distribution bin 2		°C
76.	IT3	Feed temperature in the distribution bin 3		°C
77.	IT4	Feed temperature in the distribution bin 4		°C
78.	PAT	Feed temperature entering port A		°C
79.	PB1T	Feed temperature entering port B1		°C
80.	PB2T	Feed temperature entering port B2		°C
81.	PB3T	Feed temperature entering port B3		°C
82.	PB4T	Feed temperature entering port B4		°C
83.	PB5T	Feed temperature entering port B5		°C

To be continued

Table A1 (continued)

	Tag	Description	Range	Unit
84.	PB6T	Feed temperature entering port B6		°C
85.	PC1T	Feed temperature entering port C1		°C
86.	PC2T	Feed temperature entering port C2		°C
87.	PC3T	Feed temperature entering port C3		°C
88.	PC4T	Feed temperature entering port C4		°C
89.	PC5T	Feed temperature entering port C5		°C
90.	PC6T	Feed temperature entering port C6		°C
91.	PC7T	Feed temperature entering port C7		°C
92.	PC8T	Feed temperature entering port C8		°C
93.	PC9T	Feed temperature entering port C9		°C
94.	PC10T	Feed temperature entering port C10		°C
95.	WBCVPF	Weigh bin cone valve position feedback	≥ 0	mA
96.	WBCVOMV	Weigh Bin Cone valve opening (Measured Value)	0-100	%
97.	WBCVPCO	Weigh Bin cone valve position controller output	≥ 0	mA
98.	CVPCSP	Cone valve position controller set point	0-100	%
99.	FFPAPAF	Furnace feed pipe A, A-port flow	≥ 0	tons/hr
Reduction parameters				
100.	FBRLSP	Fluidized bed reducer level set point		m
101.	FBRL	Fluidized bed reducer level		m
102.	FBRCMV	Fluidized bed reducer level		mA
103.	FBRLCO	Fluidized bed reducer level controller output	≥ 0	mA
104.	FBRCVC	Fluidized bed reducer cone valve control	0-100	%
105.	FBRBedT	Fluidized bed reducer temperature	800-1100	°C
Calcining parameters				
106.	CalcFR	Calciner feed rate	≥ 0	tons/hr
107.	CoalFeed	Coal feed rate	≥ 0	tons/hr
Laboratory parameters				
108.	AL2O3	Al_2O_3 concentration in the slag		ppm
109.	FeO	FeO concentration in the slag		ppm
110.	MgO	MgO concentration in the slag		ppm
111.	Ni	Ni concentration in the slag		ppm
112.	SiO2	SiO_2 concentration in the slag		ppm

References

[1] A. Paullada, I. D. Raji, E. M. Bender, E. Denton, A. Hanna, Data and its (dis)contents: A survey of dataset development and use in machine learning research, Patterns 2 (2021) 100336.

[2] J. Downs, E. Vogel, A plant-wide industrial process control problem, Computers Chemical Engineering 17 (1993) 245–255. Industrial challenge problems in process control.

[3] P. F. Odgaard, J. Stoustrup, M. Kinnaert, Fault-tolerant control of wind turbines: A benchmark model, IEEE Transactions on Control Systems Technology 21 (2013) 1168–1182.

[4] J. Van Impe, G. Gins, An extensive reference dataset for fault detection and identification in batch processes, Chemometrics and Intelligent Laboratory Systems 148 (2015) 20–31.

[5] X. Yang, D. Feng, Generative adversarial network based anomaly detection on the benchmark tennessee eastman process, in: 2019 5th International Conference on Control, Automation and Robotics (ICCAR), pp. 644–648.

[6] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. S. Bernstein, A. C. Berg, L. Fei-Fei, Imagenet large scale visual recognition challenge, CoRR abs/1409.0575 (2014).

[7] Z. Dai, H. Liu, Q. V. Le, M. Tan, Coatnet: Marrying convolution and attention for all data sizes, CoRR abs/2106.04803 (2021).

[8] R. Olson, W. La Cava, P. Orzechowski, R. Urbanowicz, J. Moore, Pmlb: A large benchmark suite for machine learning evaluation and comparison, BioData Mining 10 (2017).

[9] A. Feinstein, H. Cannon, Fidelity, verifiability, and validity of simulation: Constructs for evaluation, Developments in Business Simulation and Experiential Learning 28 (2001) 57–67.

[10] E. L. R. L H Chiang, R. D. Braatz, Fault detection and diagnosis in industrial systems, Measurement Science and Technology 12 (2001) 1745.

[11] G. Birol, C. Ündey, A. Çinar, A modular simulation package for fed-batch fermentation: penicillin production, Computers Chemical Engineering 26 (2002) 1553–1565.

[12] J. Van Impe, G. Gins, An extensive reference dataset for fault detection and identification in batch processes, Chemometrics and Intelligent Laboratory Systems 148 (2015) 20–31.

[13] M. Bartyś, R. Patton, M. Syfert, S. de las Heras, J. Quevedo, Introduction to the damadics actuator fdi benchmark study, Control Engineering Practice 14 (2006) 577–596. A Benchmark Study of Fault Diagnosis for an Industrial Actuator.

[14] N. F. Thornhill, S. C. Patwardhan, S. L. Shah, A continuous stirred tank heater simulation model with applications, Journal of Process Control 18 (2008) 347–360. Festschrift honouring Professor Dale Seborg.

[15] L. von Rueden, S. Mayer, R. Sifa, C. Bauckhage, J. Garcke, Combining machine learning and simulation to a hybrid modelling approach: Current and future directions, in: M. R. Berthold, A. Feelders, G. Krempl (Eds.), Advances in Intelligent Data Analysis XVIII, Springer International Publishing, Cham, 2020, pp. 548–560.

[16] A. Stief, R. Tan, Y. Cao, J. R. Ottewill, N. F. Thornhill, J. Baranowski, A heterogeneous benchmark dataset for data analytics: Multiphase flow facility case study, Journal of Process Control 79 (2019) 41–55.

[17] R. E. V. Vargas, C. J. Munaro, P. M. Ciarelli, A. G. Medeiros, B. G. do Amaral, D. C. Barrionuevo, J. C. D. de Araújo, J. L. Ribeiro, L. P. Magalhães, A realistic and public dataset with rare undesirable real events in oil wells, Journal of Petroleum Science and Engineering 181 (2019) 106223.

[18] A. Melo, M. M. Câmara, N. Clavijo, J. C. Pinto, Open benchmarks for assessment of process monitoring and fault diagnosis techniques: A review and critical analysis, Computers Chemical Engineering 165 (2022) 107964.

[19] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, J. Lin, Machinery health prognostics: A systematic review from data acquisition to rul prediction, Mechanical Systems and Signal Processing 104 (2018) 799–834.

[20] J. Jiao, M. Zhao, J. Lin, K. Liang, A comprehensive review on convolutional neural network in machine fault diagnosis, Neurocomputing 417 (2020).

[21] X. Liang, J. Tang, L. Li, Y. Wu, Y. Sun, A review of metallurgical processes and purification techniques for recovering mo, v, ni, co, al from spent catalysts, Journal of Cleaner Production 376 (2022) 134108.

[22] L. Rippon, I. Yousef, B. Hosseini, J.-F. Beaulieu, C. Prévost, S. Shah, B. Gopaluni, Process analytics and machine learning to predict arc loss in an electric arc furnace, in: 59th Conference of Metallurgists 2020 hosting the 4th International Uranium Conference.

[23] G. Qu, S. Zhou, H. Wang, B. Li, Y. Wei, Production of ferronickel concentrate from low-grade nickel laterite ore by non-melting reduction magnetic separation process, Metals 9 (2019).

[24] I. Kotzé, Pilot plant production of ferronickel from nickel oxide ores and dusts in a dc arc furnace, Minerals Engineering 15 (2002) 1017–1022.

[25] E. Keskinkilic, Nickel laterite smelting processes and some examples of recent possible modifications to the conventional route, Metals 9 (2019).

[26] W. Meihack, The potential role of fluidized beds in the metallurgical industry, Journal of the Southern African Institute of Mining and Metallurgy 86 (1986) 153–160.

[27] Q. G. Reynolds, C. J. Hockaday, D. T. Jordan, I. J. Barker, Arc detection in dc arc furnaces, in: P. J. Mackey, E. J. Grimsey, R. T. Jones, G. A. Brooks (Eds.), Celebrating the Megascale, Springer International Publishing, Cham, 2016, pp. 157–167.

[28] R. Jones, Reductive smelting for the recovery of nickel in a dc arc furnace, European Metallurgical Conference, EMC 2013 (2013) 1019–1026.

[29] R. T. Jones, Dc arc furnaces — past, present, and future, in: P. J. Mackey, E. J. Grimsey, R. T. Jones, G. A. Brooks (Eds.), Celebrating the Megascale, Springer International Publishing, Cham, 2016, pp. 129–139.

[30] Q. Reynolds, R. Jones, B. Reddy, Mathematical and computational modelling of the dynamic behaviour of direct current plasma arcs, Journal of the Southern African Institute of Mining and Metallurgy 110 (2010).

[31] H. Pauna, T. Willms, M. Aula, T. Echterhof, M. Huttula, T. Fabritius, Electric Arc Length-Voltage and Conductivity Characteristics in a Pilot-Scale AC Electric Arc Furnace, Metallurgical and Materials Transactions B 51 (2020) 1646–1655.

[32] L. D. Rippon, I. Yousef, B. Hosseini, A. Bouchoucha, J.-F. Beaulieu, C. Prévost, M. Ruel, S. Shah, R. B. Gopaluni, Representation learning and predictive classification: Application with an electric arc furnace, Computers & Chemical Engineering 150 (2021) 107304.

[33] H. P. Vinutha, B. Poornima, B. M. Sagar, Detection of outliers using interquartile range technique from intrusion dataset, in: S. C. Satapathy, J. M. R. Tavares, V. Bhateja, J. R. Mohanty (Eds.), Information and Decision Sciences, Springer Singapore, Singapore, 2018, pp. 511–518.

[34] K. Zhang, M. Hutter, H. Jin, A new local distance-based outlier detection approach for scattered real-world data, in: T. Theeramunkong, B. Kijsirikul, N. Cercone, T.-B. Ho (Eds.), Advances in Knowledge Discovery and Data Mining, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 813–822.

[35] M. Sugiyama, K. Borgwardt, Rapid distance-based outlier detection via sampling, in: C. Burges, L. Bottou, M. Welling, Z. Ghahramani, K. Weinberger (Eds.), Advances in Neural Information Processing Systems, volume 26, Curran Associates, Inc., 2013.

[36] M. J. Azur, E. A. Stuart, C. Frangakis, P. J. Leaf, Multiple imputation by chained equations: what is it and how does it work?, International Journal of Methods in Psychiatric Research 20 (2011) 40–49.

[37] S. Khan, A. Latiful Haque, Sice: an improved missing data imputation technique, Journal of Big Data 7 (2020).

[38] D. M. P. Murti, U. Pujianto, A. P. Wibawa, M. I. Akbar, K-nearest neighbor (k-nn) based missing data imputation, in: 2019 5th International Conference on Science in Information Technology (ICSITech), pp. 83–88.

[39] Z. Che, S. Purushotham, K. Cho, D. Sontag, Y. Liu, Recurrent neural networks for multivariate time series with missing values, Scientific Reports 8 (2018).

[40] S. J. Choudhury, N. R. Pal, Imputation of missing data with neural networks for classification, Knowledge-Based Systems 182 (2019) 104838.

[41] N. Sakthivel, B. B. Nair, M. Elangovan, V. Sugumaran, S. Saravanmurugan, Comparison of dimensionality reduction techniques for the fault diagnosis of mono block centrifugal pump using vibration signals, Engineering Science and Technology, an International Journal 17 (2014) 30–38.

[42] T. N. Lal, O. Chapelle, J. Weston, A. Elisseeff, Embedded Methods, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 137–165.

[43] J. Qian, Z. Song, Y. Yao, Z. Zhu, X. Zhang, A review on autoencoder based representation learning for fault detection and diagnosis in industrial processes, Chemometrics and Intelligent Laboratory Systems 231 (2022) 104711.

[44] R. Isermann, On fuzzy logic applications for automatic control, supervision, and fault diagnosis, IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 28 (1998) 221–235.

[45] S. Joe Qin, Statistical process monitoring: basics and beyond, Journal of Chemometrics 17 (2003) 480–502.

[46] D.-T. Hoang, H.-J. Kang, A survey on deep learning based bearing fault diagnosis, Neurocomputing 335 (2019) 327–335.

[47] S.-Y. Shao, W.-J. Sun, R. Yan, P. Wang, R. Gao, A deep learning approach for fault diagnosis of induction motors in manufacturing, Chinese Journal of Mechanical Engineering 30 (2017) 1347–1356.