

# Robust health indicator and rankings for railway point machines using motor current curves

Susanne Reetz<sup>1,\*</sup> , Thorsten Neumann<sup>2</sup> , Douwe Buursma<sup>3</sup>, Wessel Roodenburg<sup>4</sup>

<sup>1</sup>Institute of Transportation Systems, German Aerospace Center e.V. (DLR), Lilienthalplatz 7, D-38108 Braunschweig, Germany

<sup>2</sup>Institute of Transportation Systems, German Aerospace Center e.V. (DLR), Rutherfordstr. 2, D-12489 Berlin, Germany

<sup>3</sup>Strukton Systems, Jan Tinbergenstraat 110, NL-7559 SP Hengelo, the Netherlands

<sup>4</sup>Strukton Rail, Westkanaaldijk 2, NL-3542 DA Utrecht, the Netherlands

\*Corresponding author. Institute of Transportation Systems, German Aerospace Center (DLR), Lilienthalplatz 7, D-38108 Braunschweig, Germany.

E-mail: [susanne.reetz@dlr.de](mailto:susanne.reetz@dlr.de)

Special issue Title: Smart Railway Infrastructure

## Abstract

Digitization and artificial intelligence for monitoring the health conditions of switches and point machines in the railway infrastructure has become increasingly important. By using a large unlabeled dataset of motor current curve measurements collected over two years from 58 electro-mechanical point machines in the Netherlands—containing over 1.7 million samples—this paper introduces an unsupervised health indicator. In contrast to many existing approaches, the indicator is comparable across point machines and operating temperatures. It also features transparent training and prediction pipelines, and its outputs are easily interpretable by maintenance personnel. The resulting ranking can be used to quickly gain an overview and prioritize planned maintenance between machines. Evaluation of the health indicator on the large dataset demonstrates behaviors that highlight the importance of condition monitoring systems that go beyond the analysis of individual current curves. Instead, short-, mid-, and long-term observations should be synthesized into a coherent overall health assessment.

**Keywords** railway switch, point machine, condition monitoring, current curve, prognostics and health management, health indicator

## 1. Introduction

Switches and their point machines are vital components of the railway infrastructure, influencing both network availability and safety. A large portion of delays in railway operations can be traced back to infrastructure failures, particularly those involving switches, see [Swier \(2021\)](#). Conventional maintenance strategies that depend on preventive actions at fixed intervals fail to provide the high availability required without incurring excessive costs. Consequently, operators are increasingly pursuing condition-based and predictive maintenance strategies to overcome these challenges.

### 1.1. State of the art

The authors in [Hamadache et al. \(2019\)](#) present a comprehensive introduction to and overview of various fault detection and diagnosis methods for railway point machines. A widely used approach involves measuring the current or power curves of the electric motors that drive the point machines and applying data-driven techniques. This perspective is echoed in [Hu et al. \(2023\)](#), where the authors offer an updated review of data-driven—in particular, machine-learning—approaches.

Existing fault detection and diagnosis techniques typically operate on a per-sample basis, that is, each switch movement is evaluated independently without leveraging historical observations from the same point machine. Environmental factors such as temperature are sometimes incorporated, and some approaches are trained independently for each point machine and direction, whereas others are built on a broader dataset that includes multiple point machines. Comprehensive labeled datasets are rarely available and are typically of limited size, making supervised learning impracticable.

For instance, the studies [Huang et al. \(2018\)](#), [Narges et al. \(2021\)](#), and [Guo et al. \(2020\)](#) all follow a two-stage workflow. They first perform a fault detection step and then classify faults using an ensemble of small, fault-specific classifiers. The work in [Narges et al. \(2021\)](#) relies on labeled data, 6,000 samples selected from data collected of thirty-three point machines over one year, covering four fault types and the normal state. [Guo et al. \(2020\)](#) combine unsupervised machine learning with expert assessments during training, with an unlabeled training set of 19,104 samples and 500 labeled test samples. [Huang et al. \(2018\)](#) adopt a fully unsupervised approach, on a dataset of 817 switch movements of four point machines. In [Zhang et al. \(2021\)](#), the authors implement a two-stage diagnostic system, based on limited labeled data. The

Received: 23 February 2026. Revised: 2 April 2026. Accepted: 21 April 2026

© The Author(s) 2026. Published by Oxford University Press and Southwest Jiaotong University.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

first stage distinguishes between normal operation and several fault modes based on power curves. This is then extended in a second stage by incorporating information about the control circuit, allowing the classification of multiple fault sub-classes. The dataset contains eight fault types, each with twelve to fifty samples, and fifty normal samples. None of these approaches explicitly incorporate environmental factors.

In contrast, the authors in [Narezo et al. \(2019\)](#) introduce an anomaly detection approach that accounts for temperature effects using a simple normalization technique. Their dataset consists of motor current curves from one direction of seven switches, air temperature, and reported switch failures over five years. It is supplemented with expert assessments of a limited amount of samples for test purposes. The models are trained separately for each point machine and direction and evaluated per current curve, using a healthy training dataset. The authors also examined the dependency of their features on humidity and precipitation data from local weather stations, finding no correlation. The integration of this methodology into an operational pipeline is described in [Narezo and Guzman et al. \(2022\)](#).

In [Reetz et al. \(2024\)](#), a diagnostic model based on Bayesian networks is presented that uses expert knowledge and measurement data to determine the root cause of point machine faults. Human-interpretable motor current curve features and other information sources are used to provide a fuzzy multi-label classification as decision support for maintenance engineers. The model has been manually validated on fifty-five different scenarios, but the robust automatic calculations of evidences from motor current curves (and other measurement data) have yet to be implemented. While the model generally works on a per-motor-current-curve basis, it also contains evidence nodes that aggregate information on past switch behavior to model cascading failures. In theory, this approach could also be applied when directly modeling faults by aggregating information from several recent measurements into evidence nodes, rather than relying on a single measurement.

[Böhm \(2017\)](#) presents a supervised method for predicting the remaining useful life of point machines by framing the task as a multi-class classification problem of distinct failure-lead-time classes. The dataset consists of motor power curves, temperature and humidity, and switch fault reports from twenty-nine switches over two years, albeit with recording gaps. The author acknowledges that failure reports do not exactly match the true switch state, which remains unknown, but are used as a substitute.

The authors in [Atamuradov et al. \(2019\)](#) construct a health indicator for dry sliding-chair failure modes by fusing point machine force, proximity, voltage, and current data and estimate remaining useful life. The methodology is validated on field-test data from a point machine with artificially induced sliding-chair degradation, yielding sixty samples in total.

## 1.2. Challenges

[Hu et al. \(2023\)](#) state that 'In the context of developing an intelligent point machine condition monitoring and fault detection system, it is imperative that the system exhibits trustworthiness, robustness, and a high degree of generalization and transferability', and that the core essence of most of their eight derived requirements for condition monitoring system is to effectively address the challenges posed by limited labeled

and abundant unlabeled training data in diverse and complex operational scenarios.

Much of the existing research does not yet take full advantage of the vast amounts of unlabeled data that commercial implementation of sensors in the field generate. The usability of results for operators is overlooked, as per-sample outputs must be aggregated during normal operations, creating an additional barrier to practical deployment. Many approaches lack transparency and interpretability, which reduces the trustworthiness of the system in real-world applications. Further, any method that seeks to benefit from unlabeled data must account for at least temperature as an environmental factor and the differences between individual point machines.

## 1.3. Contributions

This paper introduces an unsupervised health indicator for point machines based on motor current curve measurement data that are comparable across point machines and temperatures, using a large unlabeled dataset of 1.76 million current curves of 58 point machines collected over two years. In line with the requirements of maintenance operators, the entire pipeline is fully transparent, and the resulting outputs are easily interpretable by humans. The derived ranking can be used to quickly gain an overview and prioritize between point machines. Further, the evaluation of the health indicator on the large dataset of the study reveals behaviors that emphasize the importance of condition monitoring systems that go beyond the analysis of individual current curves. Instead, short-, mid-, and long-term observations should be synthesized into a coherent overall health assessment.

The remainder of this paper is organized as follows. First, the dataset and its characteristics are described. Next, the feature engineering steps and their results are presented. The final selected feature is comparable across point machines. Because the feature's distribution shape changes with temperature, the next section introduces a transparent health indicator methodology that addresses this challenge. The health indicator ranges from 0 to 1 and enables comparison across temperatures. The approach's sensitivity to hyperparameters and its comparison with existing methods are discussed. In the application section, the health indicator's behavior and the ranking is evaluated on the dataset and lessons for design of condition monitoring approaches for railway point machines are discussed. The paper concludes with a summary and an outlook on future work.

## 2. Data

The dataset for this study comprises measurement data of fifty-eight single switches in the Netherlands from the beginning of April 2023 until the end of March 2025. All switches are equipped with a single electro-mechanical point machine of the type NSE2 (see [VRS Railway Industry \(2022\)](#)), operating at a nominal voltage of 136 V DC. Their motor current curves are measured in the relay house using the POSS monitoring system (see [Strukton Rail \(2026\)](#)) with a sampling frequency of 50 Hz; the air temperature is measured outside of the relay house. In total, the dataset contains about 1.76 million current curves. The dataset is unlabeled, i.e. there is no information on the state of the point machines, representing the typical situation in operations. Supply voltage is not measured.

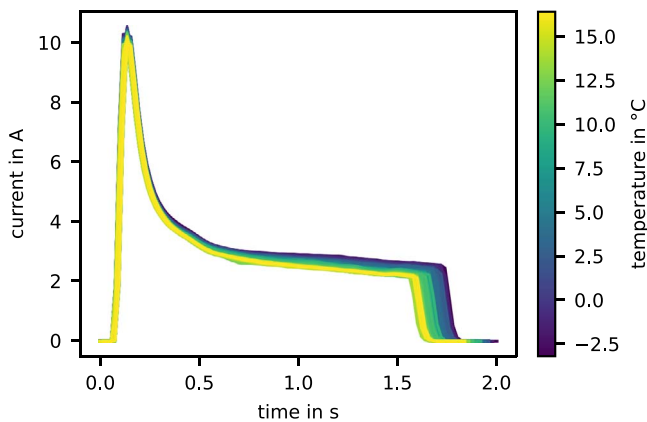


**Figure 1** Raw motor current curves of one direction of a point machine in good condition over 10 days at different temperatures.

Figure 1 shows exemplary data of a healthy point machine at different temperatures. As stated in the previous studies Böhm (2017) and Narezo et al. (2019), the motor current curves of this type of point machines exhibit a temperature dependence. The lower the temperature, the higher the current and the longer the duration of the switch movement. The actual asset temperature of a point machine can differ from the air temperature measured at the relay house, potentially reducing the accuracy of temperature-based assessments for some point machines.

Moreover, motor current curves can be impacted by their supply voltage. Different no-load voltage levels affect current curves in a manner similar to temperature. Transient changes in the operational voltage can also alter the shape of current curves, as demonstrated in the dedicated analysis in Reetz et al. (2026). Their main dataset consists of six months of data from a small railway yard with six NSE2 point machines operating at a nominal voltage of 136 V DC, all connected to a shared voltage supply. The no-load voltage is usually around 136 V or 140 V, with a total range from 134 V to 143 V. As this information is not available for this study, any variance in the data created by voltage fluctuations remains unexplainable. The switches are not all in the same location, so some variation of at least the same magnitude as in Reetz et al. (2026) is to be expected.

A point machine has two throw directions—normal and reverse, which can produce markedly different current curves. The purposeful combination of these characteristics can inform diagnostics. For this reason, they are presented separately throughout this paper.

It should be noted that the switches in this dataset were selected as part of another research study and include several units that the maintenance operator deemed to be of particular interest. Consequently, the dataset may contain a higher proportion of faulty switches than the average switch population in the Netherlands. This potential bias, however, does not compromise the overall viability of the presented methodology or the general interpretation of the results.

### 3. Feature engineering

Motor current curves can reveal faults by deviating from their characteristic shape. For an NSE2-type point machine, this shape is shown in Figure 1. After the inrush peak, the current flattens out

and decreases gradually until the switch blades reach their end positions. At this point, the motor shuts off and the current drops to zero. Generally, an increase in mechanical resistance during the switch movement increases the current, whereas a decrease in resistance reduces it. Faults within the electric motor circuit can also impact the current profile. A list of potential features for NSE-series point machines and some preliminary implementations are presented in Narezo and Guzman et al. (2022) and Neumann et al. (2023).

The objective of the feature engineering in this study is to quantify the ‘humpedness’ of a current curve, that is, local increases in current caused by heightened mechanical resistance, using a transparent calculation that yields interpretable results, ensures comparability between point machines, and remains robust across the entire dataset.

#### 3.1. Preprocessing

Preprocessing starts with a validity check, to sort out any data that cannot be processed as a current curve. For this, each measurement of a switch movement is compared with a dynamically defined corridor. Only if the measurement is within the limits of that corridor, it is considered a valid current curve and fit for processing. After calibration of the corridor hyperparameters, the validity checks in the preprocessing discard about 0.01% of the data. The discarded data have been checked manually, and fully consist of measurements that are not suitable for the feature engineering, as they are highly atypical due to sensor faults, blockages of the switch blades, and similar reasons. On the other hand, all valid current curves that produced outlier feature values were checked manually, and there were no measurements that should not have passed the validity checks. Thus the calibration of the preprocessing is considered successful, even though the full dataset cannot be labeled manually for testing purposes due to its size.

An example of an invalid measurement is shown in Figure 2. Presumably, the switch was blocked at the end of its first throw (0–9.4 s), reversed back to the starting position (9.4–10.6 s), and then succeed in the second attempt (10.6–12.3 s). As the point machine motor never shuts off between the movements, the sensor system does not separate out the current curves and instead considers them as a single measurement. Even if separated out, the first throw would not be considered valid as there is a blockage in the end and the second throw is shorter than usual, as the switch blades are not locked in their end position at the beginning of the throw and thus the point machine needs a shorter time to reverse. Only the last throw is a typical switch throw that moves the switch blades from end position to end position, for which features could be calculated. In operations and for diagnostic models such as in Reetz et al. (2024), preprocessing outliers provide valuable information, but need to be treated separately from the remaining valid data, as discussed later.

After the validity checks, the current curves are up-sampled from 50 Hz to 200 Hz, using an Akima-spline, to improve resolution both in time and current. This approach was verified using 500 Hz motor current curves from the same point machine type obtained from another test dataset. The improved resolution is especially important to determine the exact timing of the inrush peak and the turn off point, where the point machine shuts down and the current drops to zero.

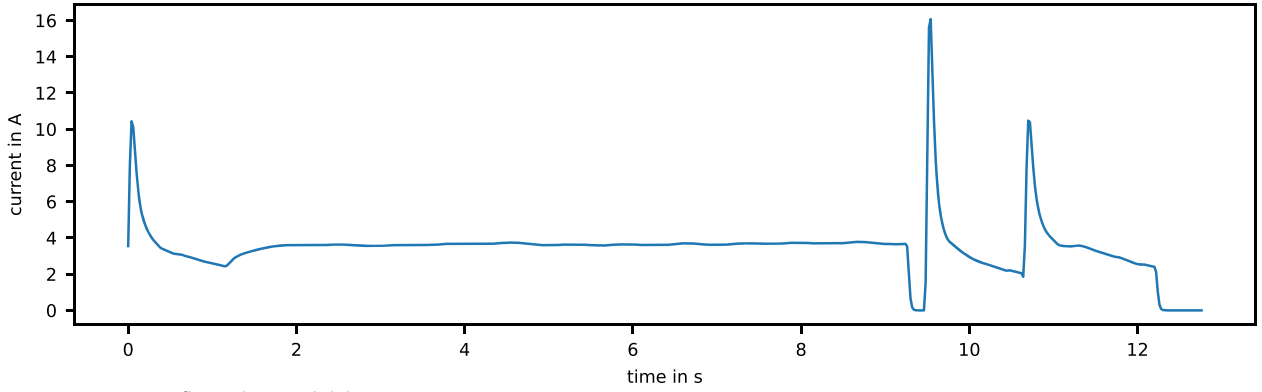


Figure 2 A measurement flagged as invalid during preprocessing.

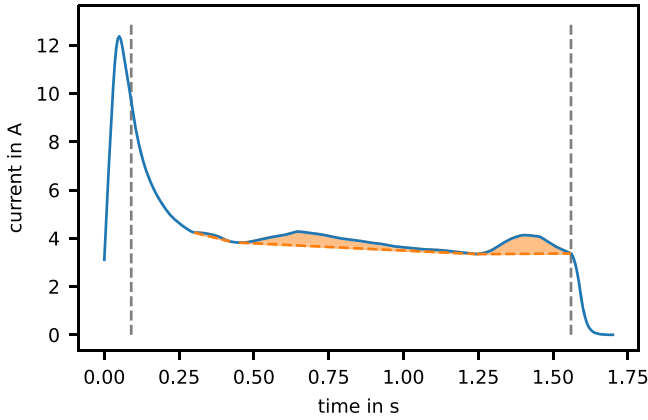


Figure 3 A preprocessed motor current curve (blue) with its detected humps and their baselines (orange) within the hump search borders (gray).

### 3.2. Feature construction

Humps—local increases in the current due to heightened mechanical resistance—can occur at any time between the start of the unlocking of the switch blades and the end of the locking. First, hump search borders are defined, then all humps within these search borders are detected.

The left hump search border is located at a fixed time interval after the inrush peak. The right hump search border is the elbow point toward the end of the current curve, right before the current drops to zero. Figure 3 shows a preprocessed current curve with its hump borders and the humps that are detected in the next step.

#### Box 1 Hump detection

Let  $P_1 = (t_1, z_1), \dots, P_n = (t_n, z_n)$  denote the points of the preprocessed current curve between the hump search borders as a time series with  $t$  as time and  $z$  as the current value. The straight line connecting two points  $P_i$  and  $P_j$  is notated as  $\overline{P_i P_j}$ . The detection of the humps works as follows.

1. Find a sample  $P_i$  that is an initial hump, that is to say, it lies above the straight line between its direct neighbors  $z_i > \overline{P_{i-1} P_{i+1}}(t_i)$ . The initial hump edge

indices are  $e_l = i - 1$  and  $e_r = i + 1$  and the baseline of the hump is  $\overline{P_{e_l} P_{e_r}}$ .

2. Alternately expand the left and right hump edges. For this, define the new hump edge as the sample outside of the old hump edge that is closest to the hump search border, such that all samples between the new hump edges are above the baseline  $z_j > \overline{P_{e_l} P_{e_r}}(t_j)$  for  $e_l < j < e_r$ . The final hump edges are found when neither the left or right edge can be expanded while remaining inside the search borders  $e_l \geq 1, e_r \leq n$ .
3. Repeat the two preceding steps until no additional humps can be identified within the hump search borders. Hump can share edge indices, but by definition they do not overlap.

The concept underlying this detection algorithm was first proposed in Neumann et al. (2023). In contrast to a previously implemented variant in Narezo and Guzman et al. (2022), the baseline of a hump is not defined as a horizontal line, but as the segment between two points, and thus can be sloped, compare Figure 3. This way, humps can also be detected during the unlocking of the switch blades, shortly after the inrush peak, and on naturally sloped current curves.

### 3.3. Feature characteristics

Based on the detection of humps, the next step is to accumulate this information into features. For diagnosis, several features that indicate different properties are required. To assess the overall humpiness of a current curve, a one-dimensional feature is the goal. In both cases, comparability between different point machines is essential so that large datasets can be exploited rather than analyzing each point machine individually.

An intuitive way to describe the overall humpiness of a current curve is to cumulate the areas of all humps of the curve, that is, the sum of the areas between hump baseline and current curve, illustrated in orange in Figure 3, into a single value. This is called the absolute cumulative hump area, measured in Ampere-seconds. The relative cumulative hump area is obtained by dividing the absolute cumulative hump area by the area under the current curve between both hump search borders, i.e. it denotes the fraction of the current curve that belong to humps. If no humps are

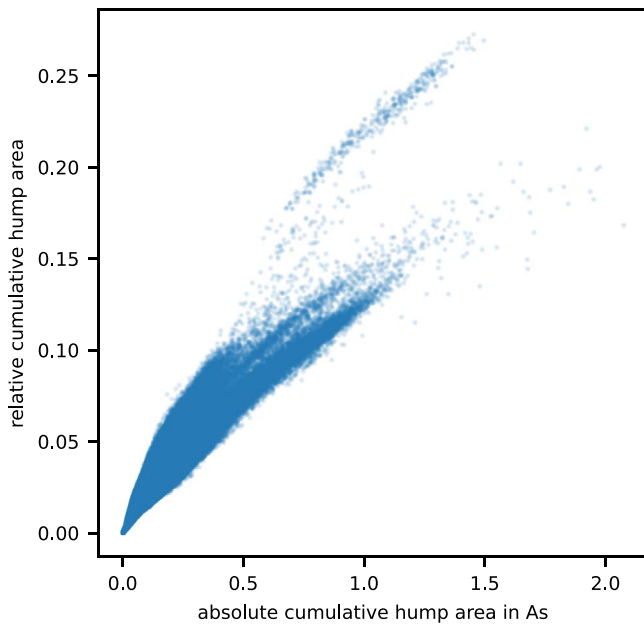


**Figure 4** Relation between the absolute cumulative hump area feature and relative cumulative hump area feature on the whole dataset.

detected, both features are zero. [Figure 4](#) shows the relationship of these two features across the entire dataset after preprocessing and hump detection.

Although all point machines in this study are of the same manufacturing type, their current curves systematically differ in length, height, and slope. These differences cannot be readily explained by switch state, environmental conditions, age, switch blade size, tangent, or other metadata available for these switches. Similar individualistic behavior has been observed in current curves of point machines of other manufacturing types, see [Hu et al. \(2023\)](#). To address this, for example, the SIDIS W system defines its extracted features for S700K point machines relative to the characteristic power level measured at installation, see [Böhm \(2017\)](#). However, there is no evidence that the individual behavior of a point machine remains consistent over its entire service life, whether due to gradual changes or

major maintenance actions. Consequently, simple features such as the area under the whole current curve are generally unsuitable for direct comparison between point machines, which is why the absolute cumulative hump area differs markedly from the relative cumulative hump area. To ensure comparability between point machines, the relative cumulative hump area serves as the basis for developing the health indicator in the next section and is often referred to simply as ‘the feature’.

Other unsupervised learning approaches generally do not address this issue and instead generate their features automatically, for example, through autoencoders in [Guo et al. \(2020\)](#). This does not guarantee comparability across different point machines and directions, and it would fail to meet the traceability and interpretability requirements imposed by operators.

For diagnostic purposes—such as for the model in [Reetz et al. \(2024\)](#)—the position of humps, i.e. where the hump mass is distributed along the current curve, is a crucial factor. Accurately determining this position requires additional feature engineering. Although these aspects lie beyond the scope of the present study, the methodology of the health indicator below can be employed to handle the dependence of an arbitrary feature on an influencing factor, provided that the feature is comparable across different switches.

### 3.4. Feature temperature dependence and distribution

[Figure 5](#) shows the relative cumulative hump area values of a point machine over time. The two directions are shown in light and dark blue. The feature data have two obvious change points, which are caused by maintenance actions. The 28-day running means of the two directions are visualized in light and dark orange. In case of a manually labeled change point in the last 28 days, only the data after the change point are used to calculate the running means. The gray background shows the normalized empirical probability density of the whole dataset, i.e. of all point machines. The relative cumulative hump area is temperature dependent, with higher values in the winter than in the summer, i.e. the current curves in this dataset are more humped relative to their overall area, compared with summer. At colder temperatures, point machines

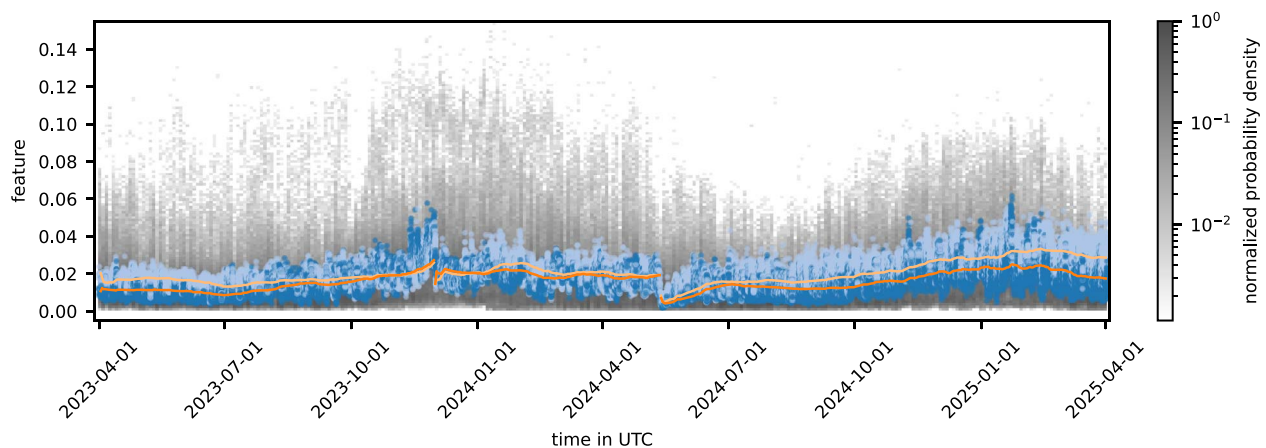


**Figure 5** Scatter plot of the relative cumulative hump area values of all measurements of the two directions of an example point machine in light and dark blue over time, with their 28 days running averages plotted on top in light and dark orange. The gray background shows the normalized empirical probability density of the whole dataset. The feature axis is limited to 0.15 to improve visibility, there are outliers up to 0.27.

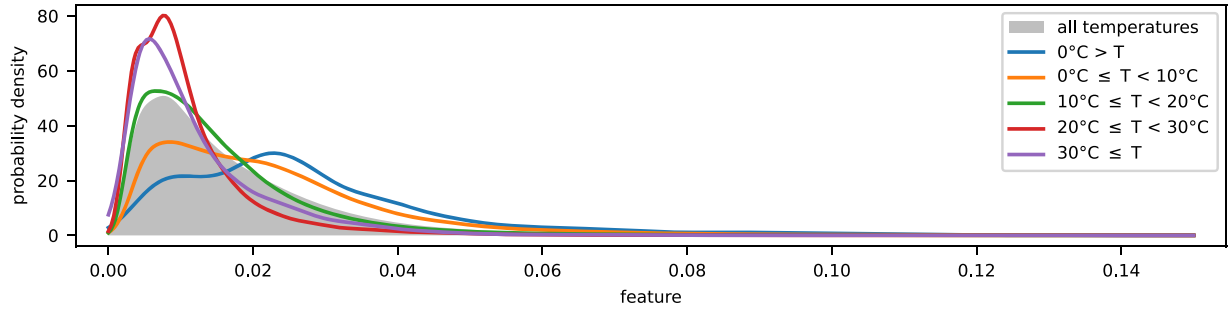


**Figure 6** Probability distribution of the relative cumulative hump area at different temperatures, calculated using kernel density estimation (Gaussian kernel, Scott's rule for bandwidth calculation). The feature axis is limited to 0.15 to improve readability, there are outliers up to 0.27.

require more effort to move and faults that result in mechanical resistance are exaggerated. While the feature values of all point machines and directions are temperature dependent, the extension of this effect can differ. Thus feature values at different temperatures cannot directly be compared with each other without further modeling.

The shape of the overall distribution of the relative cumulative hump area of all point machines varies with temperature and does not match any standard analytical distribution. Figure 6 shows the distribution in different temperature bins, against the distribution over all temperatures. Simple normalization techniques—for example, using the temperature-dependent 95%-quantile to rescale the data to  $[0, 1]$ —would not be sufficient to make the data comparable among different temperatures (see comparison below).

All change points—typically introduced by maintenance actions—are manually annotated for every point machine, which improves the accuracy of running-average calculations. In practice, these points would need to be identified automatically. If detailed digitized records of the relevant maintenance activities exist, the exact timestamps of those events could be employed directly.

## 4. Health indicator

### 4.1. Methodology

The following method is designed to transform a previously constructed feature, which depends on an influencing factor, into a health indicator that is invariant with respect to that factor. Both the feature and the influencing factor must be continuous variables. Beyond this, no assumptions are made about their distribution or other properties.

#### Box 2 Health indicator

Let  $X$  denote the influencing factor and  $Y$  the target variable, i.e. the feature value. The whole dataset consists of observations  $(x_i, y_i)$  for  $i = 1, \dots, n$ . The health indicator algorithm comprises the following steps:

1. Custom binning of  $X$ : for a chosen number of bins  $m$  and the dataset  $(x_1, \dots, x_n)$ , define  $x_{e_k}^{(u)}$  as the uniformly spaced bin edges and  $x_{e_k}^{(q)}$  as quantile-based bin edges, that is, the empirical  $k/m$ -quantiles, for  $k = 0, \dots, m$ .

For the blend factor  $\lambda \in [0, 1]$ , define the bin edges as

$$x_{e_k} = \lambda \cdot x_{e_k}^{(q)} + (1 - \lambda) \cdot x_{e_k}^{(u)}, \quad k = 0, \dots, m$$

and the respective bin center locations as

$$x_{c_k} = \frac{x_{e_{k-1}} + x_{e_k}}{2}, \quad k = 1, \dots, m$$

The resulting custom binning smoothly interpolates between uniformly spaced binning when  $\lambda = 0$  and quantile binning when  $\lambda = 1$ . The number of bins and the blend factor are hyperparameters that should be chosen such that each bin has sufficient samples while a good accuracy with regard to the influence of  $X$  on  $Y$  is retained.

2. Binned empirical distribution of  $Y$ : calculate the empirical cumulative distribution function  $\hat{F}_{Y,k}$  of  $Y$  for each bin  $k = 1, \dots, m$  of  $X$ . Alternatively, a kernel density estimate can be used. Otherwise, if the data in every bin can be adequately described by a parametric analytic distribution (with its own shape parameters per bin), the empirical cumulative distribution functions can be replaced by their analytic counterpart, requiring only the storage of the bin-specific parameters.
3. Smoothed conditional percentile ranks of  $Y$ : for each sample  $(x_i, y_i)$ , the smoothed conditional percentile rank is defined as

$$\hat{p}_i = \begin{cases} \hat{F}_{Y,1}(x_{c_1}) & \text{if } x_i \leq x_{c_1}, \\ \frac{(x_{c_k} - x_i) \cdot \hat{F}_{Y,k-1}(x_{c_{k-1}}) + (x_i - x_{c_{k-1}}) \cdot \hat{F}_{Y,k}(x_{c_k})}{x_{c_k} - x_{c_{k-1}}} & \text{if } x_{c_{k-1}} \leq x_i \leq x_{c_k}, \\ \hat{F}_{Y,m}(x_{c_m}) & \text{if } x_i \geq x_{c_m} \end{cases} \quad i = 1, \dots, n$$

This calculation blends the binned temperature distributions by using the distance from a sample to the two nearest bin centers as weights for their conditional percentile ranks.

4. Beta transformation: the smoothed conditional percentile ranks  $\hat{p}_1, \dots, \hat{p}_n$  are, by construction, approximately uniformly distributed on  $[0, 1]$ . The original

observations  $y_1, \dots, y_n$ , however, possess a characteristic shape—skewness, heavy tails,...—that is lost when the data are reduced to ranks. To recover a comparable shape while keeping the resulting score on the unit interval, an inverse beta cumulative distribution function  $F_{\text{Beta}(\alpha, \beta)}^{-1}$  is applied:

$$s_i = F_{\text{Beta}(\alpha, \beta)}^{-1}(\hat{p}_i), i = 1, \dots, n$$

The transformation always preserves the ordering of the ranks in the health indicator—if  $\hat{p}_i < \hat{p}_j$ , then  $s_i < s_j$ . The hyperparameters  $\alpha$  and  $\beta$  can be chosen in several ways. One option is to trim the outer tails of the original data  $y_1, \dots, y_n$ , rescale the remaining observations to  $[0, 1]$ , and then fit a Beta distribution to these scaled values. Alternatively, a quantile-matching approach can be employed. In both cases, the aim is to approximate the shape of the distribution of  $Y$  across all values of  $X$ .

The resulting health indicators for different bins of the influencing factor can now be combined and compared directly. In contrast to common normalization techniques—for example, removing the mean and dividing by the standard deviation for normally distributed data—this method can handle data in which the target variable has differently shaped and scaled distributions for different influencing factor values. This is achieved by enforcing a common distribution across all bins of the influencing factor. On the downside, the spacing between target variable samples can become distorted in some bins when the bin's original distribution shape diverges significantly from that of the applied beta transformation. In simple terms, the health indicator can be interpreted as comparing a given feature value with the distribution of feature values observed at the same level of the influencing factor. It thus describes health relative to the population under identical conditions. In case of more than one influencing factor, the same approach can be applied by binning each influencing factor and forming a multidimensional grid. Each cell must contain enough data to estimate its distribution reliably, so the grid must be defined carefully when data are sparse.

For practical operations, a large dataset that represents the whole population of target assets in all their health states, i.e. the dependent variable, and for all influencing factors is used as a training dataset. This dataset is used to manually determine all hyperparameters  $(m, \lambda, \alpha, \beta)$  and to train and store a model with Steps 1 and 2, while Steps 3 and 4 only need to be performed on the data that are used in application. As this approach has a low runtime, the model training step can be performed at either the start of an application software, or the model can be trained and stored in the background. The model needs to be retrained if its training data may not be representative any more. For the retrospective analysis of a large historical dataset, a representative dataset is chosen for model training—often, this will be the whole dataset—and the evaluation steps are then executed on all data.

While the presented approach is simple and illustrative, any binning approach has limited numerical precision, particularly in the outer regions of influencing factors that exhibit a tailed distribution and exert a strong influence on the target variable.

Generally, any method to estimate the conditional distribution  $F(Y|X)$  that suits the data may be selected, to then compute the conditional percentile curves and finally the ranks for the samples. The subsequent beta-transformation step remains unchanged. More smooth estimation methods, such as kernel regression, typically have higher runtime for larger datasets, requiring the model to be pretrained and stored before application. The trade-off between runtime and implementation complexity versus numerical accuracy must be evaluated in accordance with the intended use of the health indicator.

## 4.2. Implementation

The methodology presented above is implemented for the relative hump area feature, with the temperature as influencing factor.

As is expected for such a dataset, the number of samples per point machine is highly imbalanced and their usage is not uniform over time. Some point machines are employed heavily throughout the day, with the highest daily average reaching 147 movements for both directions of a single point machine. Other point machines are usually moved only a few times a week, aside from short periods of frequent usage—for instance, when common train routes are diverted due to maintenance or construction. To prevent overrepresentation of a specific switch state, the training dataset can be limited to a fixed maximum number of movements per day for each point machine and direction, with the samples selected at random. For this study, the limit is set to 20, reducing the training dataset from about 1.76 million to 0.9 million samples. This approach can also reduce runtimes in the case of very large datasets that would constitute a representative training set on their own, without impacting the accuracy. The other hyperparameters are set to  $m = 15$ ,  $\lambda = 0.4$  for the custom binning and  $\alpha = 1.1$ ,  $\beta = 4$  for the Beta distribution.

Figure 7 demonstrates the development of the health indicator from the feature values step by step. The chosen binning is a mixture between the quantile and uniform bins; the top plots show the number of samples in each bin. Even though the edge bins are rather large, they still contain only a relatively small number of samples compared with the other bins. The second plot shows the histograms of the feature values per bin. For each bin, the feature has a differently shaped, skewed, nonnegative distribution with outliers. This behavior makes conventional scaling approaches, e.g. a rescaling with a smooth maximum or 90%-quantiles to  $[0, 1]$ , on these data unreliable in application. The third plot shows the resulting conditional percentile ranks, with the same custom binning after their calculation for visualization purposes. The limited number of samples at very low and high temperatures results in a not completely uniform distribution of the ranks in these regions. At low temperature, the ranks are skewed to higher values, at higher temperature to lower values. The last plot displays the health indicator values, again binned by temperature after their calculation for visualization purposes. Their distributions at the outer temperatures are also not as smooth as in the middle and skewed the same way as the conditional percentile ranks. The background in the lower three plots is generated using a two-dimensional histogram of the data, normalized by its maximum value, producing a dimensionless relative density map ranging from 0 to 1. The color shows how populated each region is compared with the most populated region; it represents a probability density in shape but not in value.

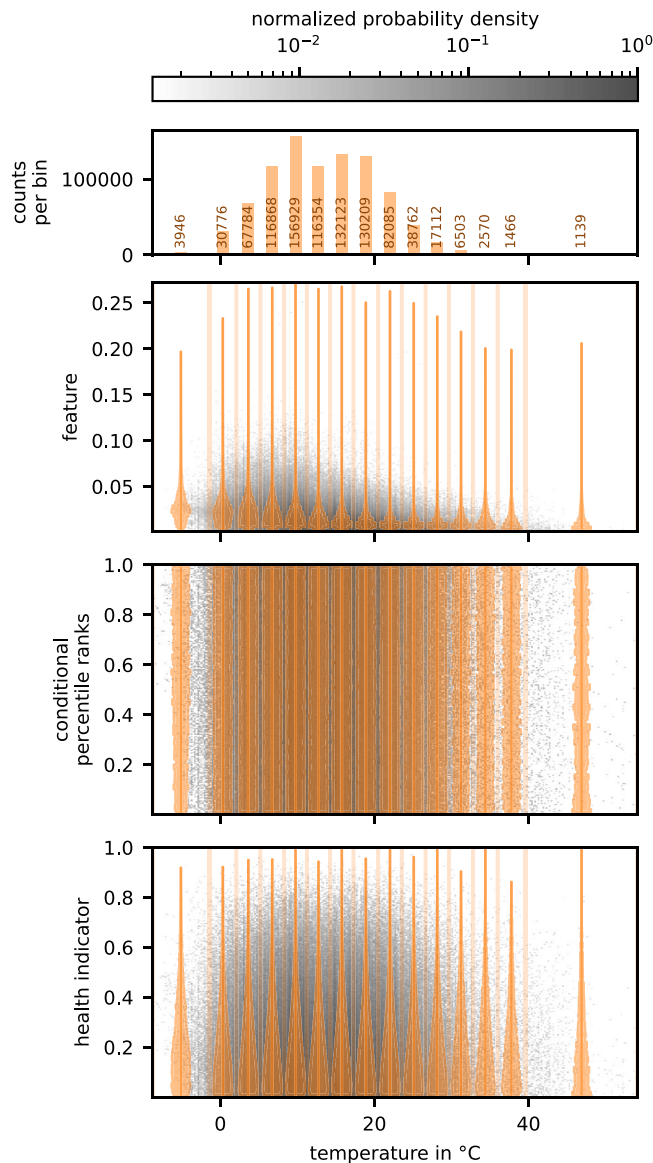


**Figure 7** Step-wise development process on the training dataset. The plot at the top shows the number of samples in each temperature bin. The lower three plots show the binned empirical probability densities of the relative cumulative hump area feature, conditional percentile rank, and health indicator in orange. The gray backgrounds show their normalized empirical probability densities.

### 4.3. Sensitivity to hyperparameters

Table 1 lists the root mean square deviation (RMSD) between health indicators that were produced with different hyperparameter settings on the entire dataset. The variation of the binning hyperparameters  $m$  and  $\lambda$  has little effect on the indicator values, indicating that the methodology is robust. Whether the training data are limited to at most 20 measurements per day per point machine or the full dataset is used for training makes an even smaller difference, suggesting that over-presentation of specific switch states is not a problem for this dataset.

### 4.4. Comparison with other normalization methods

Because the health indicator aims to eliminate temperature dependence from the feature values, the baseline for comparison

is a straightforward temperature-dependent normalization of the feature. The feature is, by design, confined to the interval  $[0, 1]$  and has a right-skewed distribution with a long tail toward higher values, see Figure 6. Suitable normalization schemes are smoothed temperature-dependent min-max scaling, min-95%-quantile scaling, and min-99%-quantile scaling.

Table 2 reports Spearman's correlation coefficient  $\rho$  (see Kornbrot (2014)) between the temperature and the original feature, the health indicator, and the normalized features. The feature has a negative correlation of  $-0.357$  with the temperature, as current curves are prone to increased humpedness at lower temperatures. None of the baseline normalization schemes reduce this dependence by more than 10.78%, with a remaining correlation of at least  $-0.3185$ . By contrast, the health indicator's correlation with temperature is  $-0.0083$ , a reduction of 97.68% relative to the original feature. Thus the health indicator sufficiently removes temperature dependence, enabling direct comparison of data collected at different temperatures, whereas the baseline normalization schemes do not.

## 5. Application

### 5.1. Results

To analyze the results, the previously trained health indicator is evaluated across the entire dataset. Figure 8 shows the output for some example point machines. For each point machine, the health indicator values of its two directions are plotted against the normalized probability density of all point machines over time. The 28-day running means are overlaid on top; if a change point occurs within the last 28 days, only the data following that change point are used to compute the running means.

Point machine 1 exhibits two change points, after which the health indicator level drops noticeably for a brief period before gradually returning to its baseline. These drops are attributable to maintenance actions, which, while improving the switch state shortly, seem to be not successful in the long term. The change points of point machine 2, on the other hand, are persistent, with a significant improvement of the health state first in one direction, then the other. The point machine then remains in a healthy state for the next year. As these are caused by maintenance actions as well, this is an example for long-term successful maintenance. Point machine 3 shows a consistent trend of worsening health over the two year period. The two directions are at markedly different health levels, indicating that the point machine has substantially more additional mechanical resistance in its throw in one direction than the other. All these point machines are representative cases in the dataset, with several other point machines exhibiting similar behavior or combinations thereof.

Point machine 4 exhibits a notable pattern, one direction shows health worsening during the winter months, while the other direction remains largely stable aside from a mild overall upward trend in the health indicator over the two year period. This behavior is an outlier in the dataset and points to a temperature-dependent fault that affects only one side of the point machine, most likely an adjustment issue. When both directions of a point machine exhibit temperature dependence in the indicator, other root causes—such as a sensor fault, discrepancies between measured air temperature and the actual asset temperature, or faults in the switch heating system—are also possible. As it is completely

Table 1 RMSD between health indicators with different hyperparameter combinations on the entire dataset.

| RMSD         |         | Training set |           |           |        |        |        |         |        |        |        |        |        |        |        |        |        |       |
|--------------|---------|--------------|-----------|-----------|--------|--------|--------|---------|--------|--------|--------|--------|--------|--------|--------|--------|--------|-------|
|              |         | All          |           |           |        |        |        | Limited |        |        |        |        |        |        |        |        |        |       |
|              |         | m            |           |           |        |        |        | m       |        |        |        |        |        |        |        |        |        |       |
|              |         | 15           |           |           |        |        |        | 20      |        |        |        |        |        |        |        |        |        |       |
| Training set | all     | m            | 15        | $\lambda$ | 0.2    | 0.4    | 0.6    | 0.2     | 0.4    | 0.6    | 0.2    | 0.4    | 0.6    |        |        |        |        |       |
|              |         |              |           |           | 0.0    | 0.0147 | 0.0254 | 0.0173  | 0.0247 | 0.0319 | 0.0012 | 0.0142 | 0.0247 | 0.017  | 0.0243 | 0.0315 |        |       |
|              |         |              |           |           | 0.4    | 0.0147 | 0.0    | 0.017   | 0.006  | 0.0162 | 0.0249 | 0.0152 | 0.0016 | 0.0161 | 0.0055 | 0.0156 | 0.0244 |       |
|              |         | 20           | $\lambda$ | 0.6       | 0.0254 | 0.017  | 0.0    | 0.0147  | 0.0042 | 0.0146 | 0.0259 | 0.0176 | 0.0024 | 0.0152 | 0.0048 | 0.0139 |        |       |
|              |         |              |           | 0.2       | 0.0173 | 0.006  | 0.0147 | 0.0254  | 0.0173 | 0.0247 | 0.0319 | 0.0012 | 0.0142 | 0.0247 | 0.017  | 0.0243 | 0.0315 |       |
|              |         |              |           | 0.4       | 0.0247 | 0.0162 | 0.0042 | 0.0135  | 0.0    | 0.0152 | 0.0252 | 0.0168 | 0.0035 | 0.014  | 0.0018 | 0.0145 | 0.0224 |       |
|              | Limited | m            | 15        | $\lambda$ | 0.6    | 0.0319 | 0.0249 | 0.0146  | 0.023  | 0.0152 | 0.0    | 0.0324 | 0.0256 | 0.0154 | 0.0236 | 0.0159 | 0.0022 |       |
|              |         |              |           |           | 0.2    | 0.0012 | 0.0152 | 0.0259  | 0.0178 | 0.0252 | 0.0324 | 0.0    | 0.0147 | 0.0252 | 0.0175 | 0.0248 | 0.032  | 0.032 |
|              |         |              |           |           | 0.4    | 0.0142 | 0.0016 | 0.0176  | 0.0068 | 0.0168 | 0.0256 | 0.0147 | 0.0    | 0.0167 | 0.0062 | 0.0163 | 0.025  | 0.025 |
|              |         | 20           | $\lambda$ | 0.6       | 0.0247 | 0.0161 | 0.0024 | 0.0137  | 0.0035 | 0.0154 | 0.0252 | 0.0167 | 0.0    | 0.0142 | 0.0034 | 0.0148 | 0.0148 |       |
|              |         |              |           | 0.2       | 0.017  | 0.0055 | 0.0152 | 0.0013  | 0.014  | 0.0236 | 0.0175 | 0.0062 | 0.0142 | 0.0    | 0.0134 | 0.023  | 0.023  |       |
|              |         |              |           | 0.4       | 0.0243 | 0.0156 | 0.0048 | 0.0129  | 0.0018 | 0.0159 | 0.0248 | 0.0163 | 0.0034 | 0.0134 | 0.0    | 0.0152 | 0.0    |       |

**Table 2** Temperature correlation measured using Spearman’s  $\rho$  and bootstrapping.

|                                                          |                         | Feature            | Health indicator   | min-q95 scaling    | min-q99 scaling   | min-max scaling    |
|----------------------------------------------------------|-------------------------|--------------------|--------------------|--------------------|-------------------|--------------------|
| Temperature correlation                                  | Spearman’s $\rho$       | −0.357             | −0.0083            | −0.3185            | −0.3211           | −0.3546            |
|                                                          | 95% confidence interval | (−0.3556, −0.3583) | (−0.0068, −0.0097) | (−0.3171, −0.3198) | (−0.3197, 0.3224) | (−0.3532, −0.3559) |
| Reduction in temperature correlation relative to feature |                         | -                  | 97.68%             | 10.78%             | 10.06%            | 0.67%              |

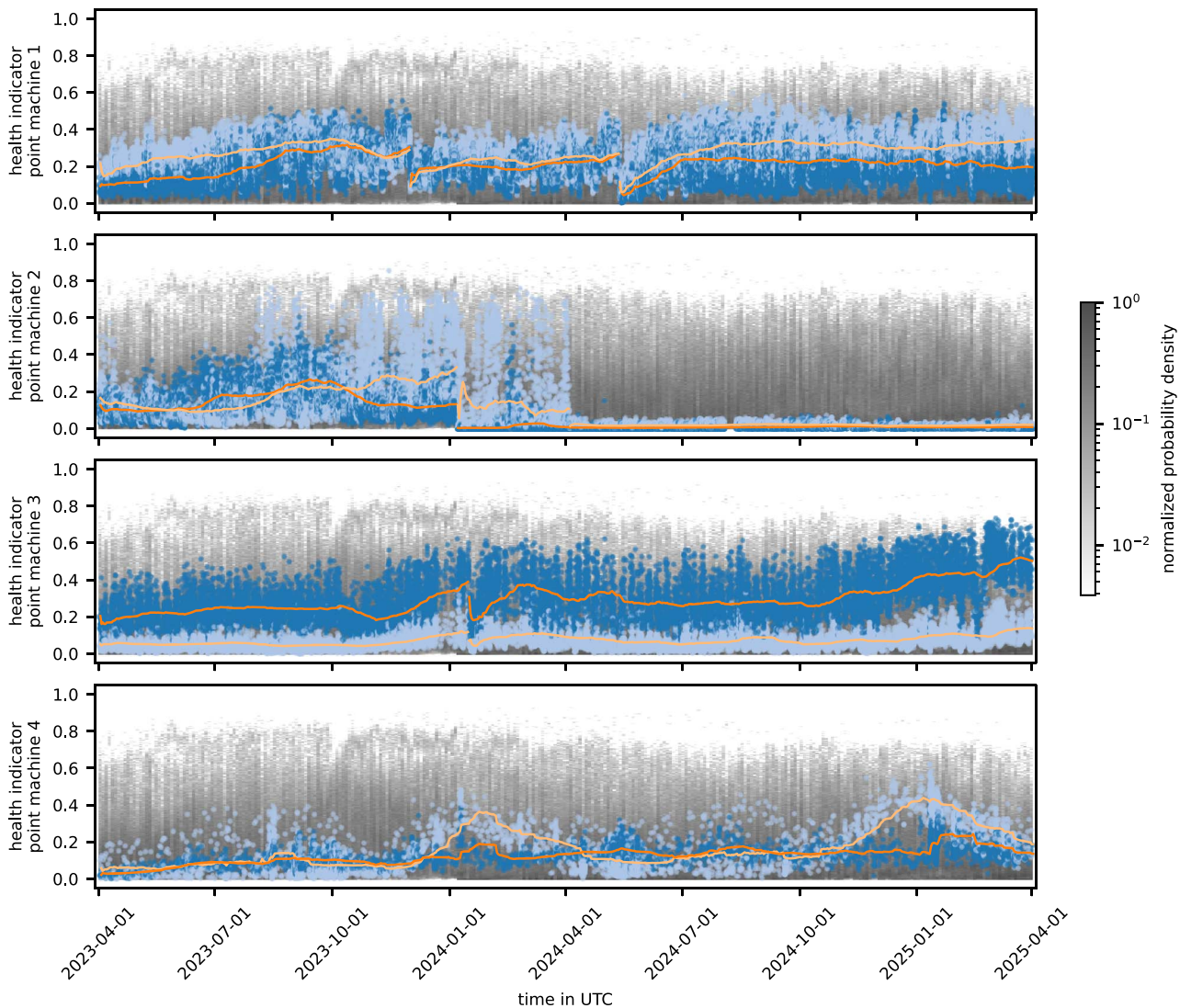


**Figure 8** Scatter plot of the health indicator of all measurements of the two directions of example point machines in light and dark blue over time, with their 28 days running averages plotted on top in light and dark orange. The gray backgrounds show the normalized empirical probability density of the whole dataset.

unsupervised, the health indicator captures only the relationship between the feature and temperature on the training dataset that contains all point machines and directions. It indicates how a sample relates to other samples at about the same temperature. If the feature values of a point machine have an average temperature dependence and the switch health is otherwise stable, the health indicator is as well. When a point machine deviates from this

expected pattern, the anomaly becomes apparent in the indicator. Although this complicates long-term trend analysis—because trends span a two-dimensional plane of temperature-time—the heightened temperature dependence is not obvious from raw measurements alone. It only becomes evident after observing the point machine at various temperatures and comparing its behavior with that of other point machines. In these cases, the

health indicator delivers a genuine gain in information that is readily interpretable and was not available before.

## 5.2. Time horizons

Considering the results above, condition monitoring for point machines has three different time horizons to consider:

- **Short-term:** immediate failure prevention and analysis of outliers. In the dataset, there are examples where the health indicator shows a sudden exponential rise over hours or days, followed by maintenance or repair. To trigger alerts automatically, each indicator per point machine and direction must be monitored for increases that exceed its usual mean and variance. Because the underlying motor current changes affect multiple features, anomaly detection methods such as those in [Narezo et al. \(2019\)](#) and [Narezo and Guzman et al. \(2022\)](#)—trained per point machine and direction—can be adapted to flag these sudden shifts and raise alarms. From an operator's perspective, such alerts enable rapid intervention, thereby mitigating the impact of point machine failures on train operations.

Further, measurements identified as outliers during preprocessing may be of interest. Apart from sensor system errors that can be disregarded, these outliers may include attempted switch movements that indicate the aforementioned failures, as well as incomplete switch movements caused by for example blockages, such as a stone from the track ballast obstructing the blades. In such cases, the data can provide additional information about the root cause of the failure.

- **Mid-term:** robust assessment of the current health state of the point machine. The health of most point machines (with regard to mechanical issues) is a relatively stable state (except for the aforementioned short-term failures), which can include some nonfatal faults. Planned maintenance should target these faults, prioritizing point machines that are in the worst condition and that are most critical to train operations. Because current curve measurement data contain inherent variance, a point machine's current state is best described by aggregating the health indicator values of recent measurements.

For robustness, the time of the last maintenance action should be incorporated, using only data collected after that event. By comparing the average health indicator values immediately before and after maintenance, operators can assess the immediate effect of the intervention and gauge its success, within the limited perspective offered by current curve measurements.

- **Long-term:** additional diagnostics and trends. In the long run, the health indicator values reveal both trends and temperature dependence of the switch state comparative to the population. The latter is an important diagnostic information. The former, together with detailed maintenance records, provides a deeper insight into the long-term effectiveness of maintenance actions and into the overall remaining useful lifetime of the point machine. Point machines that are repeatedly repaired without lasting success are either approaching the end of their service life or the fault root causes were not removed by the maintenance.

## 5.3. Ranking

One specific use-case of the health indicator is to derive a daily ordinal ranking of all point machines in the available datasets (or operated in a given region), based on their humpedness in the recent past, here 28 days. An example of such a ranking at a reference date can be seen in [Figure 9](#). Each panel shows the measurement data of one direction of one point machine, ordered and colored by its average health indicator of the past 28 days, from worst (top left) to best (bottom right) in row-major order. As the two directions of a point machine can have very different current curves, they are shown separately here. An overall health indicator for a point machine can be defined as the maximum of the health indicators of the two directions, as one cannot maintain only one direction of a point machine. This ranking answers the question 'Which point machines are currently experiencing the greatest irregular additional mechanical resistance during their throws and therefore need closer remote inspection and prompt maintenance?'. Automating this prioritization helps maintenance operators to oversee a large number of point machines much more efficiently. The ranking can be combined with other metrics, such as the criticality of each switch to train operations.

While an ordinal ranking can directly be derived from the conditional percentile ranks instead of the health indicator, it does not convey any distance information between ranks and can thus result in a distorted perception of how similar panels are. Coloring panels according to their health-indicator value corrects this: panels that share the same color have comparable health values, so even if their exact order changes from one day to the next, their visual grouping remains consistent.

In the absence of labeled data, the most practical way to evaluate the model qualitatively is to have experts visually assess the rankings, color-coded by the health indicator, at various dates. However, it should be noted that each panel displays a different number of motor current curves, potentially measured at different temperatures (see [Figure 1](#)), which means the panels are not necessarily visually comparable in detail.

## 5.4. Discussion

Using the health indicator to rank point machines addresses a specific use-case that maintenance operators encounter. The behavior of the health indicator on the whole dataset over time, however, offers a broader perspective on condition monitoring of point machines with motor current curves.

To perform a precise and comprehensive diagnosis, models must integrate not only a set of interpretable features but also an understanding of how each feature behaves across all time horizons. Models that evaluate only a single motor-current curve are inherently limited in robustness and fail to capture faults that become apparent only over longer observation windows. While extensive research exists on anomaly detection, fault diagnosis, and prediction at the level of individual motor current curves, most studies do not distill their findings into easily digestible information or actionable insights for railway operators. After all, a single switch can be moved over a hundred times a day, with a considerable variance in its motor current curves. Yet the overall health status of that switch may remain essentially unchanged for months, possibly experiencing a nonfatal fault the whole time. The classical pipeline—first performing an anomaly detection and

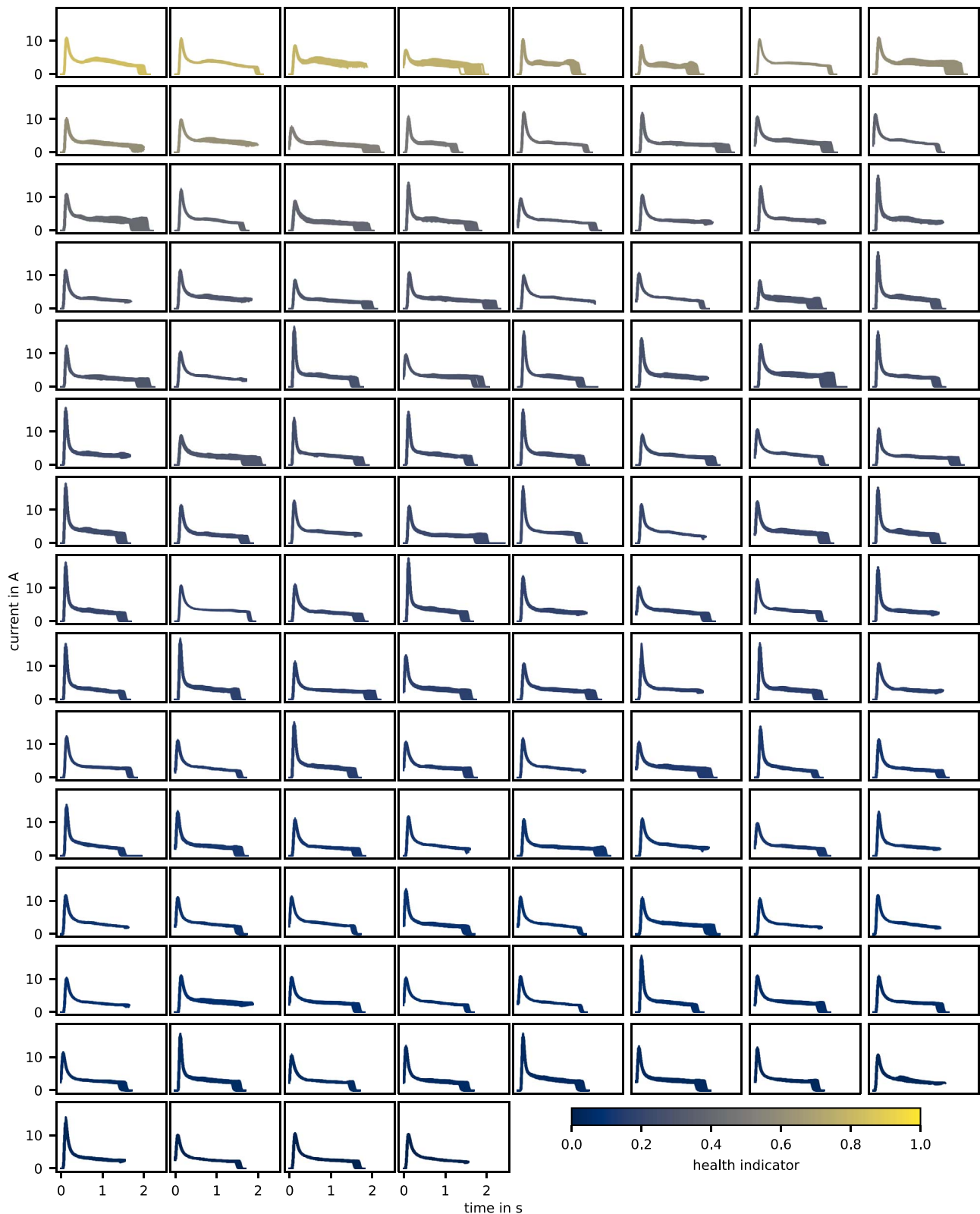


**Figure 9** Example ranking of all point machines and directions for a reference date, based on the average health indicator computed over the past 28 days. Each panel displays the raw measurement data of one direction of one point machine over this 28-day period. Panels are ordered from worst (top left) to best (bottom right) and are colored-coded by the average health indicator value.

secondly a diagnosis on only the abnormal measurements—is not a practical approach in this case.

Instead, based on the results of this study, the workflow shown in [Figure 10](#) is proposed. Central to the workflow is an integrated

diagnostic system that provides an overview of the health state of each switch, using historic and current multi-sensor data, past maintenance records, and any additional available information. Diagnostic models should be deployed at this level, thereby

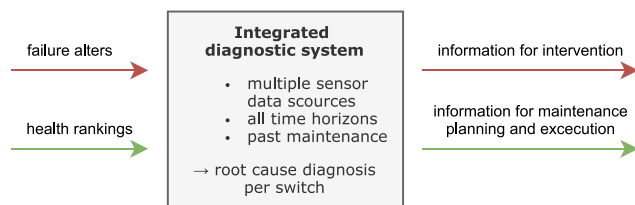


**Figure 10** Proposed workflow for an integrated diagnostic system for railway point machines.

being able to address the key requirements for point machine fault detection systems outlined in Hu et al. (2023). Two complementary workflows are used to access the diagnostic system for decision support. In the first, an event-driven process, failure alerts direct the operator to the diagnostic view of the asset, where its current state and recent history are assessed to decide whether and which intervention is required. In the second, a condition-based process, point machines are continuously evaluated using a health indicator and ranked accordingly; the lowest-ranked point machines are selected for upcoming maintenance cycles and subsequently analyzed in detail via their diagnostic system to define and prioritize appropriate work orders.

Bridging the gap between existing scientific approaches and application requires a condition monitoring system that keeps the operator and their workflow in mind. Further, transparent reasoning processes with interpretable results are essential when communicating results from inherently imperfect models.

## 6. Conclusion and outlook

The health indicator developed in this study allows to compare motor current measurements across varying temperatures and among different point machines with respect to locally increased mechanical resistance during the throw of the switch blades. Using this indicator, an unsupervised ranking that automatically prioritizes point machines for more extensive remote monitoring and upcoming maintenance is generated. Furthermore, the analysis of the indicator on this dataset demonstrates that motor current curves must be examined across different distinct temporal horizons to capture all relevant information. While extensive research exists on condition monitoring at the level of individual motor current curves, these models are inherently limited in robustness and fail to capture faults that become apparent only over longer observation windows.

Looking ahead, future feature engineering should map the detected increased mechanical resistance to the distinct phases of a switch movement and develop additional features that can be compared across point machines. By applying the methodology presented in this paper, more health indicators can be derived, used as input for a diagnostic model and aggregated into a customizable health metric. Any comprehensive diagnostic model must adequately handle information from different time horizons, integrate multi-sensor data, and present its results in an interpretable manner to operators.

## Author contributions

Susanne Reetz (Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Visualization,

Writing—original draft), Thorsten Neumann (Methodology, Software, Writing—review & editing), Douwe Buursma (Resources, Validation, Writing—review & editing), and Wessel Roodenburg (Resources, Validation, Writing—review & editing)

## Conflicts of interest

The authors declare that they have no competing interests.

## Funding

Funded by the European Union. Views and opinion expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Europe's Rail Joint Undertaking. Neither the European Union nor the granting authority can be held responsible for them. The FP3-IAM4Rail project is supported by the Europe's Rail Joint Undertaking and its members. It received funding under Grant Agreement No 101101966.



## References

- Atamuradov, V. et al. (2019) 'Railway Point Machine Prognostics Based on Feature Fusion and Health State Assessment', *IEEE Transactions on Instrumentation and Measurement*, 68: 2691–704.
- Böhm, T. (2017) 'Remaining Useful Life Prediction for Railway Switch Engines Using Classification Techniques', *IJPHM Special Issue on Railway Systems & Mass Transportation*, 8.
- Guo, Z., Wan, Y. and Ye, H. (2020) 'An Unsupervised Fault-Detection Method for Railway Turnouts', *IEEE Transactions on Instrumentation and Measurement*, 69: 8881–901.
- Hamadache, M. et al. (2019) 'On the Fault Detection and Diagnosis of Railway Switch and Crossing Systems: An Overview', *Applied Sciences*, 9: 5129.
- Hu, X. et al. (2023) 'Fault Detection for Point Machines: A Review, Challenges, and Perspectives', *Actuators*, 12.
- Huang, S. et al. (2018) 'Two-Stage Turnout Fault Diagnosis Based on Similarity Function and Fuzzy c-Means', *Advances in Mechanical Engineering*, 10: 391.
- Kornbrot, D. (2014) 'Spearman's Rho', in N. Balakrishnan, T. Colton, B. Everitt, W. Piegorisch, F. Ruggeri and J.L. Teugels (eds.) *Wiley StatsRef: Statistics Reference Online*. John Wiley & Sons, Ltd.
- Narezo, Guzman, D., Hadzic, E. and Groos, J. C. (2019) 'Anomaly Detection Model Based on Switch Monitoring Data for Condition-Based Maintenance', In: *12th World Congress on Railway Research*, Tokyo, Japan.
- Narezo, Guzman, D. et al. (2022) 'Towards the Automation of Anomaly Detection and Integrated Fault Identification for Railway Switches in a Real Operational Environment', In: *World Congress on Railway Research*, Birmingham, United Kingdom.
- Narges, K. H., Ahmad, M. and Fereydoun, G. M. (2021) 'A Hybrid Fault Diagnosis Scheme for Railway Point Machines by Motor Current Signal Analysis', *Proceedings of the Institution of*

- Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 236: 1026–34.
- Neumann, T. *et al.* (2023) ‘Modellierung Und Simulation Atypischer Merkmale Bei Weichenstellstromkurven/Modelling and Simulating the Atypical Features of Switch Engine Current Curves’, *Signal + DRAHT*, 115: 53–63.
- Reetz, S. *et al.* (2024) ‘Expert System Based Fault Diagnosis for Railway Point Machines’, *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 238: 214–24.
- Reetz, S., Buursma, D. and Lawate, A. J. (2026) ‘Impact of Supply Voltage Fluctuations on Condition Monitoring of Railway Point Machines’, In: Kour R, Karim R, Kumar U, Galar D and Jägare V (eds.) *International Congress and Workshop on Industrial AI and eMaintenance 2025*, pp. 693–704. Springer Nature Switzerland: Cham.
- Strukton Rail (2026) ‘Remote Condition Monitoring. <https://strukton.com/innovation/poss-monitoring/>’, accessed 26 Mar. 2026.
- Swier, J. (2021) ‘Ways to Improve the Performance (and Cost) of Switches’, In: Galeazzi R, Kjartansson Danielsen H, Kjær Ersbøll B, Juul Jensen D and Santos I (eds.) *Intelligent Quality Assessment of Railway Switches and Crossings*, pp. 261–74. Springer International Publishing: Cham.
- VRS Railway Industry (2022) ‘Nse Point Machine’. [https://www.vrsrail.nl/sites/vw\\_vrs/files/webcomposer/5815-Point%20Machine%20NSE%20%28VRS%20Infosheet%202022%29.pdf](https://www.vrsrail.nl/sites/vw_vrs/files/webcomposer/5815-Point%20Machine%20NSE%20%28VRS%20Infosheet%202022%29.pdf), accessed 15 Dec. 2025.
- Zhang, Y. *et al.* (2021) ‘A Hierarchical Method Based on Improved Deep Forest and Case-Based Reasoning for Railway Turnout Fault Diagnosis’, *Engineering Failure Analysis*, 127: 105446.