



Machine tool process monitoring by segmented timeseries anomaly detection using subprocess-specific thresholds

Markus Netzer¹ · Yannic Palenga¹ · Jürgen Fleischer¹

Received: 1 October 2021 / Accepted: 9 February 2022
© The Author(s) 2022

Abstract

Time series data generated by manufacturing machines during processing is widely used in mass part production to assess if processes run without errors. Systems that make use of this data use machine learning approaches for flagging a time series as a deviation from normal behaviour. In single part production, the amount of data generated is not sufficient for learning-based classification. Here, methods often focus on global signal variance but have trouble finding anomalies that present as local signal deviations. The referencing of the process states of the machine is usually performed by state indexing which, however, is not sufficient in highly flexible production plants. In this paper, a system that learns granular patterns in time series based on mean shift clustering is used for detecting processing segments in varying machine conditions. An anomaly detection then finds deviating patterns based on the previously identified processing segments. The anomalies can then be labeled by a human-in-the-loop approach for enabling future anomaly classification using a combination of machine learning algorithms. The method of anomaly detection is validated using an industrial machine tool and multiple test series.

Keywords Anomaly detection · Pattern recognition · Condition monitoring

1 Introduction

Machine failures represent an important cost factor in the current production landscape. One approach to avoiding unplanned downtime is to monitor the condition of these machines using computer-aided approaches [1, 2]. Since heterogeneous machines of different ages are often used, it is not possible to use standardized data acquisition methods and even the retrofitting of sensors is often too expensive. This leads to the necessity of machine-specific approaches to condition monitoring. However, since relevant machines are controlled numerically (NC) or by programming logic (PLC) and generate time series data during the execution of processes, it is possible to access this data across machines. In the machining context, a machine workflow usually consists of several sub-processes, which can also be repeated across different processes and machines. As an example, the forward movement of the cutter head of a milling machine is

given, which will appear in a wide variety of work processes and different parts.

The use of these recurring patterns makes it possible to introduce generic condition monitoring across machines and processes by identifying anomalies in the associated time series segments. In the following sections, an anomaly detection system is presented which identifies deviations in time series based on threshold values. However, since the calculations are not performed on complete signals but on a priori calculated, simpler sub-sequences of the superordinate time series, such a slightly modified, threshold-based approach clearly distinguishes itself from standard methods which commonly use global threshold crossings or state-based indexing to identify anomalies.

This distinction is twofold: first, in the reduced intra-cluster variance for sub-process clusters which helps to detect anomalies at much higher granularity, and secondly in the form of easy adaptability to different machine types as data on sub-processes can be shared easier than full process time-series. This aids in generating value in a brownfield setting, where only the set of signals available across all different machines can be considered. Additionally, the system is designed to be expandable, allowing for scalability with

✉ Markus Netzer
markus.netzer@kit.edu

¹ wbk Institute for Production Science, Karlsruhe Institute of Technology (KIT), Kaiserstraße 12, 76131 Karlsruhe, Germany

larger sets of sensors shared across the machines the system is attuned to.

2 State of the art

A variety of different approaches to anomaly detection in time series exist in the literature and can be roughly classified into different categories. A comprehensive overview is presented by Cook et al. [3]. Tree-based methods can be used for tool wear detection [4] and Liu, et al. [5] use an ensemble of binary isolation trees for computationally effective detection of anomalies without explicit computation of densities or distances. Sliding window-based approaches [6], as well as deep learning [7] and neural networks in the form of autoencoders [8] and long-short-term-memory-based approaches [9] [10], are also applicable to the problem. In general, approaches deal with a wide range of use cases, which is why a direct transfer to modified problems is often not straightforward. In the domain of monitoring machining operations, several reviews of the state of the art and future directions in the field exist [2, 11]. Graß et al. presented the GADPL algorithm [12], which clusters and segments time series and machine parameters and applies dissimilarity measures to then detect anomalies. This allows for unsupervised detection of deviations in machine behaviour but relies on domain knowledge to segment the time series. Further unsupervised approaches include the UFAN architecture introduced by [13], who used data from different machines to be able to monitor an entire fleet, which is achieved using a generative adversarial network for low-dimensionality feature generation in combination with a one-class extreme learning machine (ELM) classifier. Jove et al. [14] implement an anomaly detection for industrial control loops in an unsupervised manner by mapping data to two-dimensional space and then setting limits in this space. Among other approaches, principal component analysis, beta Hebbian learning and curvilinear component analysis are applied, with Beta Hebbian learning performing the best out of these approaches. Other approaches in predictive maintenance also focus on feature extraction [15] (FRESH), with the features then being possible to use as high-quality input to supervised machine learning models, provide high accuracy but the need for labeled training data, which might not always be available. Theumer et al. [16] detect point and collective anomalies using sliding windows and autoencoders respectively. Machine learning based algorithms are also popular in the context of wear detection of machine tools [17]. Axinte and Gindy [18] show how spindle power can act as an often readily available data source for identifying anomalies in different machining processes. They showed that spindle power

is sensitive enough for anomaly detection in continuous processes such as drilling and turning, but not sensitive enough to detect deviations in milling processes. However, spindle power is not assessed on a sub-sequence level, which could lead to higher sensitivity regarding the detection of deviations. Indeed, Wang et al. [19] showed that there is a correlation between tool wear and spindle power when specific frequencies are considered. Additionally, some methods work in a sub-process-based context like the approach presented in the following but are not directly transferable due to partially different baseline conditions. A potential approach for such sub-process-based anomaly detection was presented by Zhou et al. [20], who use supervised learning to detect different skills of a robot based on multimodal time series data for the identification of anomalous events. To recognize individual robot skills, the sticky hierarchical Dirichlet process (sHDP) by Fox et al. [21] is used in combination with hidden Markov models. The HMM learned in this way can then be used to compute the probability for a given sequence of observations (time series values), which, when combined with the membership of particular robot skills, allows a conclusion to be drawn about the presence of an anomaly. However, in the present case, since many different subprocesses within the time series often occur at low frequencies, supervised learning of parameters of a model is difficult. In addition, due to machine wear, anomalies occur as temporally extended slow changes from expected values of the time series, which means that consistent assignments to processes are not always possible. Another approach based on time series segmentation followed by anomaly detection was presented by Zhang et al. [22]. Here, anomalies in high-frequency signals are detected by first segmenting them and then examining them for anomalies using a combination architecture of a Recurrent Neural Network for feature extraction and a convolutional neural network-based autoencoder. Signals are considered utilizing two different window sizes in different granularity to enable fast classifications. For the present application, the approach is only partly transferable, because the window-based segmentation of the time series does not provide a clear division into actual processes, which makes the clear assignment of anomalies to subprocesses of a work process on machine tools potentially difficult.

In Table 1, a subset of these papers and respective approaches are rated according to the focus on the following prerequisites for the use case at hand:

- Cyclicity (Cyc.): The method assumes recurring patterns in the data and integrates this property.
- Predictive (Pred.): Model focuses on predicting anomalies in the future. Points are also given if the model focuses on predictive maintenance.

Table 1 Rating of different approaches regarding properties necessary for the use case at hand

Method	Cyc.	Pred.	Onl.	RT.	Unsup.	Mb.
Axinte et al. 2004	--	+	unk.	unk.	++	++
GADPL	+	++	unk.	--	++	--
FRESH	0	++	-	unk.	0	++
UFAN	--	0	-	+	++	--
Jove et al. [14]	--	-	-	+	++	--
Zhou et al. [20]	++	++	--	-	--	0

- On-line training (onl.): Model can be trained on-line.
- Real-time (RT.): The model can identify anomalies in streaming data.
- Unsupervised (Unsup.): The model does not rely on labeled training data
- Model-based (Mb.): The system needs a model of the machinery.

Points are awarded based on whether the approach focuses on implementing the individual properties. The comparison shows that without adaption, the individual approaches are not fully transferable to the use case and no single approach covers all the targeted properties at once.

3 Approach

A machining process of a machine can also nearly be described by the variables defined in the controller. For example, in the case of a milling machine, these include values for the position of the milling head or tool center point (TCP), the spindle current, or the feed rate. Figure 1 displays such a signal, extracted from a machining center performing different milling tasks.

By comparing the target values for these parameters with the actual measured values, it is possible to detect anomalies. This can be achieved, for example, by comparing measured values at certain time indices using a definition of the range of actual values allowed for a process. However, this has the disadvantage that permitted thresholds must be defined on a process-specific basis. It is often the case, however, that a single machining process consists of different sub-processes, some of which may occur again in other processes or other parts. The presented approach uses this transferability and therefore presents itself in two steps:

1. Extraction of process-describing patterns from available time series and retrieval during machine operation.
2. Anomaly detection for indirect tool condition monitoring.

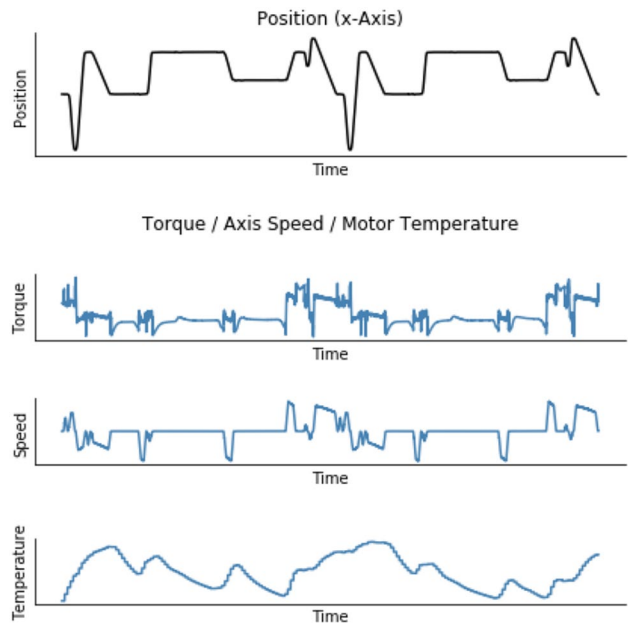


Fig. 1 Example for different signals generated by machine tool: position (black) can act as a reference while other signals (blue) can be compared for anomaly detection

This has the advantage that patterns can be extracted automatically from available time-series data, thus bypassing the process-specific manual matching of target and actual values. In this way, the system can detect anomalies based on the machining sequence being run and not based on generally applicable intervention limits. It should be noted however that process differentiation does not need to group all processes of a similar nature into one cluster. As an example, when feed velocities differ, different sub-processes can be grouped into different process clusters based on the different feed velocities. This leads to a much more granular distinction of processes which helps anomaly detection due to less intra-cluster variance but comes at the disadvantage of needing more process cycles for training data collection.

Based on these properties, the approach is heterogeneously applicable to different machines. With the additional availability of sensor data, this can also be integrated to enable more precise detections.

For the application of the approach, it is necessary to enable standardized access to the data in the controller, which can be achieved by implementing an OPC UA server in combination with intelligent parameter identification [23].

3.1 Pattern recognition

To recognize recurring sub-processes, it is necessary to divide existing time series into sub-sequences. Utilizing these sub-sequences, reference sub-signals can be generated, which can later be searched for in the on-line data. The splitting of the machine tool position time series into sub-sequences works by segmenting the provided input time series based on the detection of local minima. Using this criterion, sub-sequences as shown in Fig. 2 are generated.

These sub-sequences do not necessarily represent entire processes, but rather processing segments (sub-processes) that can reappear across different processing procedures. Afterward, the time series sub-sequences based on the position signal are grouped with the respective signals across other channels (such as currents, torque, etc.), based on the timestamp the splits in the position signal appear. If the position segment appears again, clusters can be generated across these other channels. An approach for clustering these signals was previously presented using Mean-Shift Clustering in [24], subsequently, the approach was extended by extensive data pre-processing, involving smoothing of the positional time series and offset corrections to further improve the matching with previously appearing positional signals of the same type.

To enable re-detection of these positional signals in on-line data, a sliding buffer on the on-line data is used, matching positional signals in the offline database to signals appearing in the data stream. To reduce matching time, the patterns in the offline database are matched at different positions in the signal buffer using the mean absolute error [23]. Using iterative calculations of the distance between the offline patterns and the signal in the buffer in combination with stopping distance calculations for individual offline

patterns early if the distance increases above a specified threshold leads to pattern matching that can be applied for streaming data. Once a pattern is re-detected, the respective clustering of the other channels (torque, current, ...) can then be used to make a comparison to the respective signal that appears in parallel to the on-line position signal.

3.2 Anomaly detection

The anomaly detection now employed is based on a comparison of the identified pattern references (e.g. torque) with the signals during the operation of the machine.

The reference patterns identified in the pattern recognition are generated by a modification of the arithmetic mean adapted to the different pattern lengths.

$$\sigma_{C_k}^2 = \frac{1}{T_k} \sum_{t=1}^{T_k} \sum_{x_{i,t} \in C_{k,t}} \|x_{i,k,t} - \bar{x}_{k,t}\|_2^2 \tag{1}$$

where C_k is the cluster for pattern k , $\bar{x}_{k,t}$ is the cluster mean of Cluster k at timestep t , $x_{i,k,t}$ is the data point i of Cluster k at timestep t , T_k is the number of timesteps available for cluster k .

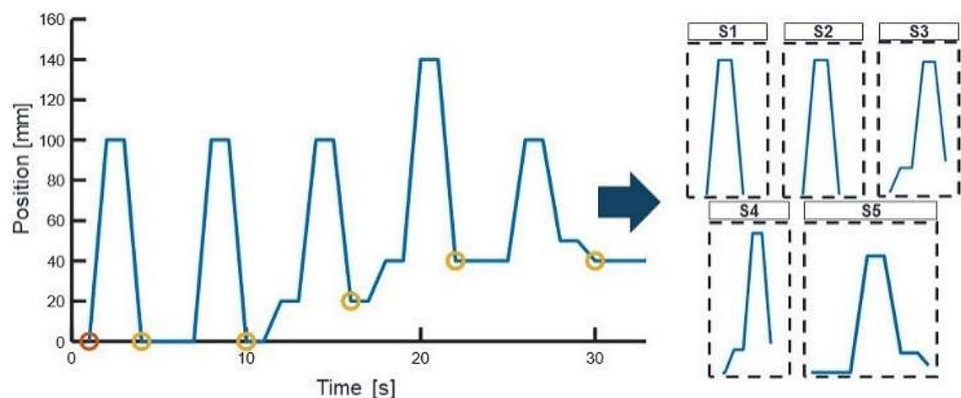
Thus the within-cluster-variance (Eq. 1) can be calculated for each individual data point of the respective cluster and then be used to calculate tolerance limits for this reference pattern. The limits can for example be calculated by using the standard deviation of the cluster members (Eq. 2), which acts as an easy but extendable approach for identifying deviations:

$$\begin{aligned} Tr_k^U &= \{ \bar{x}_{k,t} + \epsilon \sqrt{\sigma_{C_k}^2}, \quad 0 \leq t \leq T_k \} \\ Tr_k^L &= \{ \bar{x}_{k,t} - \epsilon \sqrt{\sigma_{C_k}^2}, \quad 0 \leq t \leq T_k \} \end{aligned} \tag{2}$$

The parameter ϵ can be used to manually weight the standard deviation of a sample.

The disadvantage of this generic approach is that in areas of low signal variance, threshold values are very close to

Fig. 2 Segmentation of positional time series into sub-sequences using local minima. The figure shows the position of the tool head over time, local minima that mark the segmentation into sub-processes are marked. The resulting sub-sequences representing process segments are shown to the right



actually observed (and thus clustered) signal values. An example of this problem is demonstrated in the visualization in Fig. 2, where low and high variance regions exist within a time series. This complicates the choice of a parameter ϵ . In addition, it is visible that in areas with large gradients, signal trajectories are laterally close to the boundaries, making anomaly detection susceptible to shifts.

To overcome these problems, a modified approach for the calculation of the tolerance range is used.

Increasing the spacing of the tolerance range in areas with high signal gradients in lateral direction, the calculation of the upper and lower threshold values is implemented by moving maxima and minima. Two parameters t_b and t_f are used to define the window size for the moving calculation, which results in a hull curve of the signal as follows:

$$\begin{aligned}
 x_{k,t}^U &= \max_{t-t_b \leq t \leq t+t_f, i \in I_k} x_{i,k,t}, & 0 \leq t \leq T_k, \\
 x_{k,t}^L &= \min_{t-t_b \leq t \leq t+t_f, i \in I_k} x_{i,k,t}, & 0 \leq t \leq T_k
 \end{aligned}
 \tag{3}$$

The calculation of the adjusted tolerance range taking into account the waveform is carried out based on these upper and lower thresholds according to the following rule:

$$\begin{aligned}
 Tr_k^U &= \{x_{k,t}^U + \epsilon_1(P_{99}^k) + \epsilon_2(x_{k,t}^U - x_{k,t}^L) + \epsilon_3\sigma_{C_k}^2, & 0 \leq t \leq T_k\} \\
 Tr_k^L &= \{x_{k,t}^L - \epsilon_1(P_{99}^k) - \epsilon_2(x_{k,t}^U - x_{k,t}^L) - \epsilon_3\sigma_{C_k}^2, & 0 \leq t \leq T_k\}
 \end{aligned}
 \tag{4}$$

$x_{k,t}^U$ and $x_{k,t}^L$ denote the previously formed hull curves, which result from the sliding maxima and minima over the respective clustered signal values. $\epsilon_1, \epsilon_2, \epsilon_3$ denote hyperparameters which determine the influence of the corresponding variables with ϵ_1 weighted value P_{99}^k corresponds to the 99th percentile of the difference between the hull curve and the representative for cluster k . This is used to form a wider tolerance range in the case of noisy signals since the value of P_{99}^k is larger here. Using ϵ_2 , the distance of the upper value of the envelope from the lower value can be weighted and included in the calculation. Finally, the intra-cluster variance

$\sigma_{C_k}^2$ is included in the calculation of the tolerance band via weighting parameter ϵ_3 , which can be used alternatively or in combination with P_{99}^k for the inclusion of the noise.

Figure 3 shows the formation of the tolerance band according to these rules for two different parameter combinations, while the tolerance range for a full signal, which has not been divided into individual sub-sequences, is shown to emphasize local changes due to stronger noise.

If the system detects a previously occurring pattern as described at the beginning of the chapter, a comparison is made with the previously stored reference. It is therefore checked whether the observed value lies outside the tolerance range formed for the associated cluster (Fig. 4).

For certain anomalies, movement away from normal values begins to appear over a prolonged period, which requires additional care, as extracted sub-sequences are usually short and such anomalies may therefore not be detected straight away. To solve this issue, time series are marked specifically during training, which helps with detecting such transient anomalies later on.

3.3 Human-in-the-loop

After an anomaly has been detected, labeling is carried out by integrating a human component so that, upon re-occurrence, anomalies can be classified cluster-dependently, and thus across processes, since the anomalies are considered on a sub-sequence basis independently of the higher-level signal and the associated processing procedure, as we assume that anomalies such as blowholes that appear in specific sub-processes present similarly independent of the individual machining process. This labeling extends the unsupervised system so that classical supervised machine learning approaches can be used. Additionally, such labeling in combination with the prior partitioning of signals into sequences independent of the originally performed machining process circumvents a common problem in the context of error detection in production scenarios: Since similar

Fig. 3 Global pattern variance causes problems in specific regions. simply relying on a single calculated intra-cluster variance does not leave room for deviations in regions where larger deviations would occur more frequently

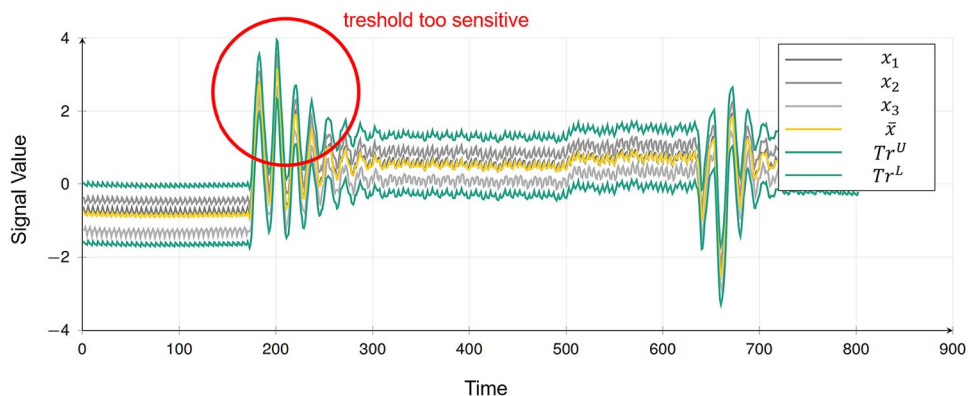
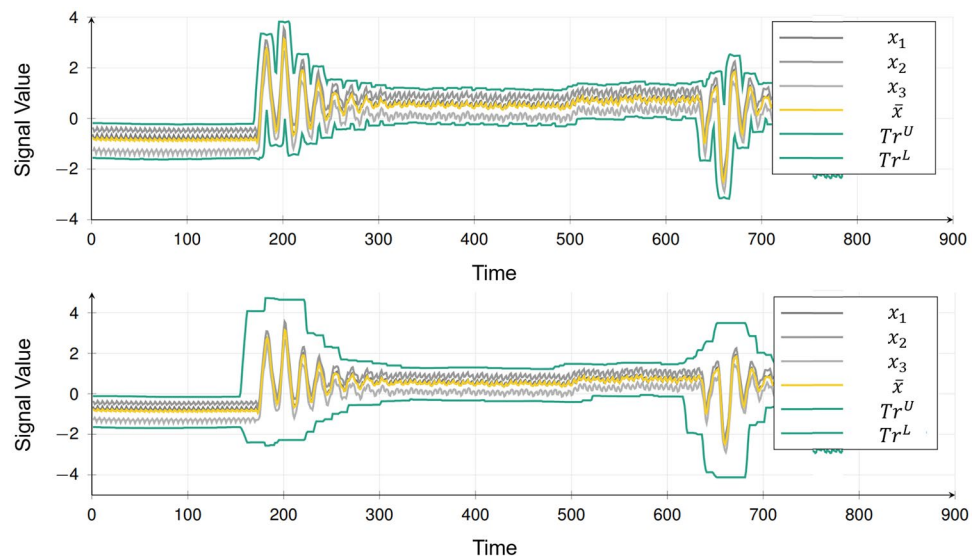


Fig. 4 Visualization of custom thresholds Tr_k^U and Tr_k^L for different moving maximum window sizes and coefficient combinations $\epsilon_1, \epsilon_2, \epsilon_3$ for a complete time series: $t_b = t_f = 5, \epsilon_1 = 0.05, \epsilon_2 = \epsilon_3 = 0$ (Top), $t_b = t_f = 20, \epsilon_1 = 0.05, \epsilon_2 = 0.02, \epsilon_3 = 0$ (Bottom). Depending on the chosen parameters, the sensitivity of the anomaly detection can be controlled



training data is usually required to train a classifier, it is necessary to generate such data at a high cost. In addition, a classifier trained in this way can only be applied to new data if it has a similar structure to the training data. This ensures that many conventional systems can only be used in a very process-specific way. However, by separating the signal from the actual process, it is now possible to reuse similar patterns across processes for anomaly detection, which simplifies data acquisition.

4 Evaluation

For Validation of the previously described system in Brown Field applications, milling tests were carried out at the wbk Institute of Production Science, Karlsruhe, Germany on a DMG 6-axis milling center. The 13 individual signals extracted are shown in Table 2.

The pattern recognition system detected multiple different machining segments during the process. The subsequent anomaly detection should then be able to detect anomalies occurring in the milling process based on these segments. As part of the test series, 10 individual milling tests were carried out for fine machining. Three repetitive machining segments (circular pocket, oblong hole, rectangular pocket, see Fig. 5) were performed for each test. As anomalies, artificially introduced blowholes/pores, realized via additional holes (1–3 mm diameter) were introduced into the workpiece. Furthermore, wear is to be recognized and detected in case of increased tool life. S235JR in 200×20 mm was used as workpiece material. The lubrication was realized by a minimum quantity lubrication (MQL). A short overview on the validation methodology is shown in Fig. 6.

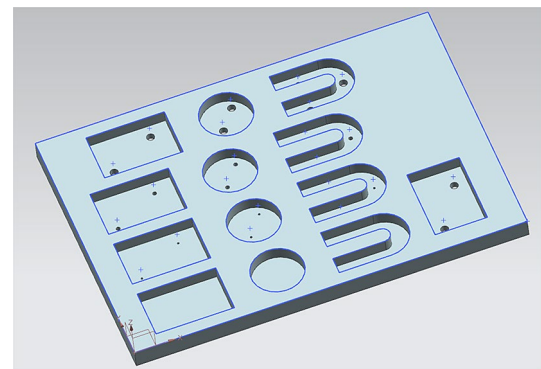


Fig. 5 CAD model of the validation workpiece. The individual shapes were milled repeatedly while blowholes were artificially realized through drilled pores in some of the pockets for the later processes

Table 2 Extracted signals from DMG 6-axis milling machine

Signal name	Axes	Use for algorithm
Milling head position	X, Y, Z	Subsequence clustering
Milling head position	Rotation	–
Spindle position	–	–
Current	X, Y, Z, rotation	Anomaly detection
Spindle current	-	Anomaly detection
Torque	X, Y, Z, rotation	Anomaly detection

By pattern recognition, 40 sub-sequences were detected as part of the milling process. These present as specific patterns in the X, Y and Z milling head position signals. With the respective clustered other signals, the thresholds are calculated as described in chapter 3. Different Anomalies are detected when these thresholds are crossed.

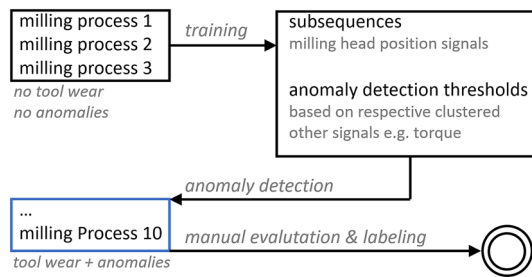


Fig. 6 Validation of anomaly detection. Training using three non-anomalous processes, followed by different tests on anomalous processes. Results were then analyzed manually

Figure 7 shows the time series segment belonging to the detection of tool wear, occurring in milling process 5. As the live data comes from one of the later milling processes, some tool wear is expected, which can result in shifts in the signal. Tool wear presents as a large series of data points crossing the thresholds. When this happens, the offset due to extended tool life and wear is too large and no longer falls inside the acceptance limits. Here, tool maintenance is recommended to avoid anomaly detection from finding such anomalies too often. The anomalous event is detected within the spindle current signal and presents as the crossing of the threshold that was calculated from the training data. The detection of tool wear depends on the training data used. If tool wear is already present during training, it is more difficult to detect later on, as additional variance is introduced. Note that to give more insight into the training data, the thresholds are shown calculated for each individual cluster member—the true threshold is therefore calculated by taking their mean. It is important to note that the splitting of the time series into these segments allows for these thresholds to be calculated on a local level.

This presents a significant advantage compared to approaches that define such thresholds globally, as even slight variations that do disguise on a global level can be identified, as is the case in Fig. 7. Due to the way the thresholds are calculated, when fast changes in the signal occur, the thresholds still have sufficient distance to the signals.

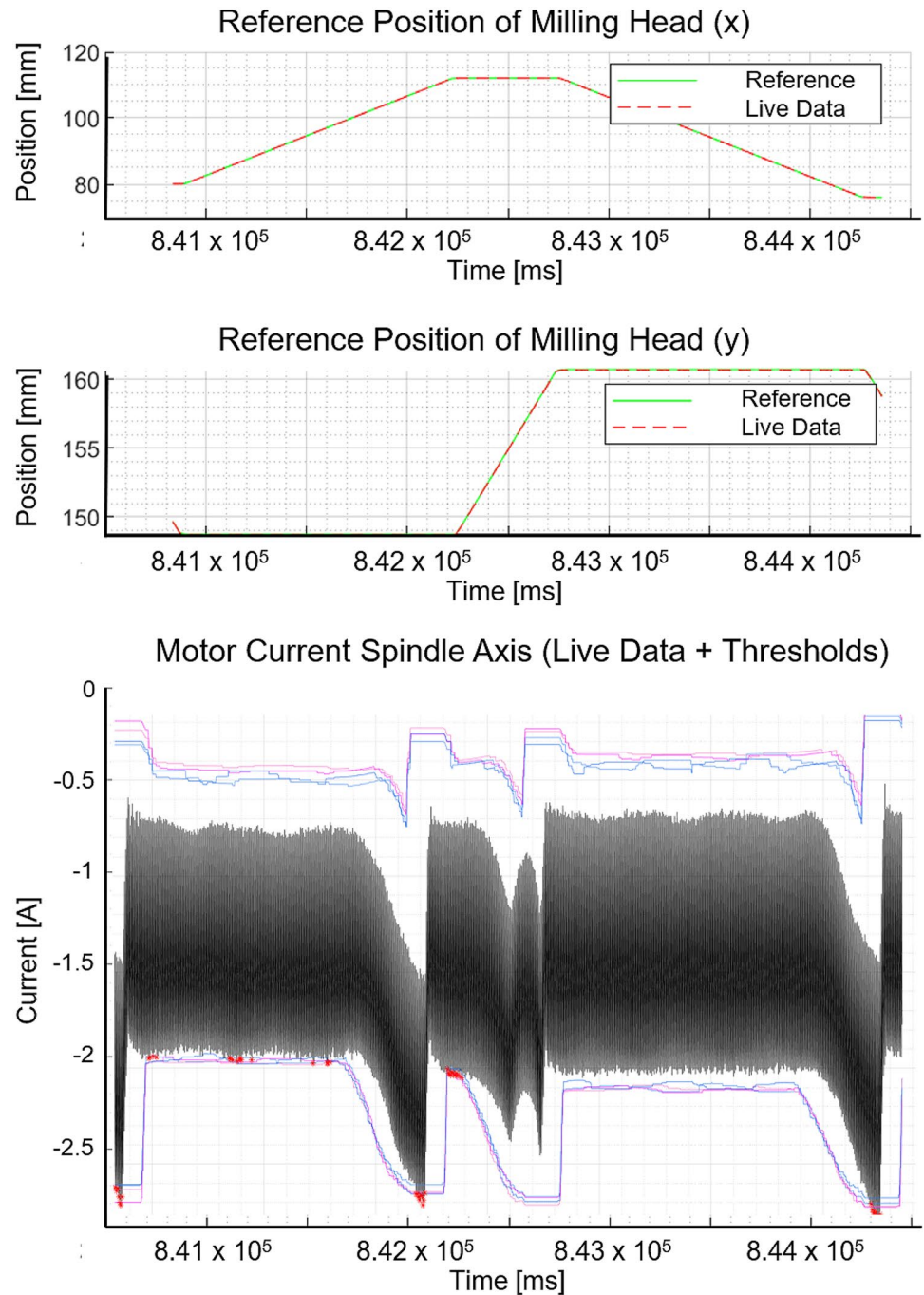
The detected tool wear anomaly is shown in further detail in Fig. 8. In addition to the overall offset, multiple crossings are detected especially at the start of the sequence. As later in the sequence the offset itself does not cause the detection of an anomaly but keeps very close to the thresholds, the previous crossings are likely caused by the offset itself, and unlikely due to point anomalies such as a blowhole.

Other anomalies tend to appear differently. Figure 9 displays a detected blowhole anomaly. Here, the threshold is crossed more sharply and for a shorter periods of time. Even though the offset from the tool wear is still present (and another tool wear anomaly is detected before the blowhole anomaly), the threshold is violated in the upward direction by the sharp rise in current. The local threshold calculation allowed for the fitting of tighter thresholds, leading to the detection of this anomaly, that would have likely been lost when only using global signal variance for threshold calculation.

5 Conclusion

By splitting time series data into simple sub-sequences based on the tracked position of the machine over time and connecting respective signal sequences that appear across other channels to this sub-sequences, the presented system allows for the detection of anomalies independently of global signal variance and across different processes. With the use of custom thresholds, the effect of signals varying in noise and magnitude can be compensated. To further enhance the system, a human-in-the-loop approach is taken after anomalies are detected, which allows for the classification of an anomaly should it appear again in the same form and enables a process independent collection of training data. The system was validated in an experimental setting using milling processes that presented with and without anomalies. Two types of anomalies were detected by the system. Blowholes/pores and tool wear were identified in a milling test carried out after training on three repeated identical milling processes. Anomalies were validated manually. However, a larger validation experiment is needed to test the robustness to the varying presentation of similar anomalies. This would also confirm that the detection of the blowholes/pores would be reliable in cases where no tool wear is present at the same time. Further investigations are planned to apply the approach to other production machines, such as joining units in assembly stations or highly flexible handling robots. Larger validation series could also aid in identifying if other methods of calculating signal thresholds should be explored, especially after assessing the sensitivity of the current calculation to the introduced parameters. Especially in flexible brownfield applications, the presented intelligent anomaly detection approach based on existing data streams from the control unit is shown to provide significant added value for sustainable overall equipment effectiveness improvement.

Fig. 7 One of the detected anomalies (tool wear). The top two plots show the respective reference sub-sequences for the x and y-axis in the position signal that are used to cluster the other signals. The bottom plot shows the live signal at the timestamps belonging to the detected reference signals. The thresholds are plotted for each cluster member without taking a mean to show how much variance is present in the training data. As the thresholds are crossed at multiple points, an anomaly is detected



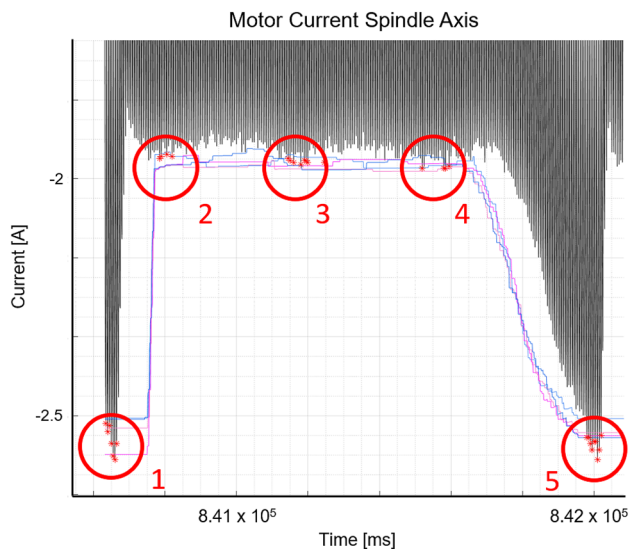


Fig. 8 The anomaly detected above in further detail. There are multiple crossings of the thresholds. Crossing 1 and 5 appear at lower spindle currents than the other anomalies (2–4)

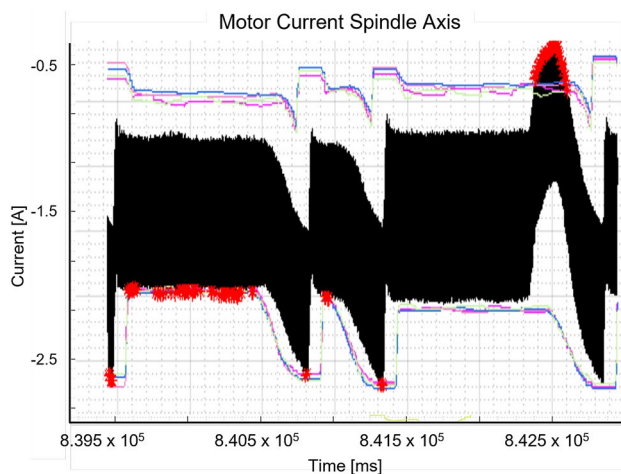


Fig. 9 Detected tool wear and blowhole anomalies. Continuous violation of the thresholds is characteristic of tool wear and may happen after extended periods of machine tool use. The blowhole presents as a sharp crossing in the upwards direction

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will

need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Roy R, Stark R, Tracht K, Takata S, Mori M (2016) Continuous maintenance and the future—foundations and technological challenges. *CIRP Ann* 65(2):667–688. <https://doi.org/10.1016/j.cirp.2016.06.006>
- Teti R, Jemielniak K, O'Donnell G, Dornfeld D (2010) Advanced monitoring of machining operations. *CIRP Ann* 59(2):717–739. <https://doi.org/10.1016/j.cirp.2010.05.010>
- Cook AA, Mısırlı G, Fan Z (2020) Anomaly detection for IoT time-series data: a survey. *IEEE Internet Things J* 7(7):6481–6494. <https://doi.org/10.1109/JIOT.2019.2958185>
- Wu D, Jennings C, Terpenney J, Gao RX, Kumara S (2017) A comparative study on machine learning algorithms for smart manufacturing: tool wear prediction using random forests. *ASME J Manuf Sci Eng* 139(7):071018. <https://doi.org/10.1115/1.4036350>
- Liu F, Ting K, Zhou Z (2012) Isolation-based anomaly detection. *ACM Trans Knowl Discov Data* 6:3:1-3:39. <https://doi.org/10.1145/2133360.2133363>
- Wu G, Zhao Z, Fu G, Wang H, Wang Y, Wang Z, Hou J, Huang L (2019) A fast kNN-based approach for time sensitive anomaly detection over data streams. *ICCS*. https://doi.org/10.1007/978-3-030-22741-8_5
- Wang J, Ma Y, Zhang L, Gao RX, Wu D (2018) Deep learning for smart manufacturing: methods and applications. *J Manuf Syst*. <https://doi.org/10.1016/j.jmsy.2018.01.003>
- Luo T, Nagarajan SG (2018) Distributed anomaly detection using autoencoder neural networks in WSN for IoT. In: 2018 IEEE international conference on communications (ICC), pp 1–6. <https://doi.org/10.1109/ICC.2018.8422402>
- Malhotra P, Vig L, Shroff GM, Agarwal P (2015) Long short term memory networks for anomaly detection in time series. *ESANN*
- Malhotra P, Ramakrishnan A, Anand G, Vig L, Agarwal P, Shroff GM (2016) LSTM-based encoder-decoder for multi-sensor anomaly detection. [arXiv:1607.00148](https://arxiv.org/abs/1607.00148)
- Lauro C, Brandão L, Baldo D, Reis RA, Davim J (2014) Monitoring and processing signal applied in machining processes—a review. *Measurement* 58:73–86. <https://doi.org/10.1016/j.measurement.2014.08.035>
- Graß A, Beecks C, Carvajal S, Jose A (2018) Unsupervised anomaly detection in production lines. In: Conference on machine learning for cyber-physical-systems and Industry 4.0 (ML4CPS). https://doi.org/10.1007/978-3-662-58485-9_3
- Michau G, Fink O (2019) Unsupervised fault detection in varying operating conditions. In: 2019 IEEE international conference on prognostics and health management (ICPHM). <https://doi.org/10.1109/ICPHM.2019.8819383>
- Jove E, Casteleiro-Roca J-L, Quintián H, Méndez-Pérez JA, Calvo-Rolle JL (2019) A fault detection system based on unsupervised techniques for industrial control loops. *Expert Syst* 13(6):e12395. <https://doi.org/10.1111/exsy.12395>
- Christ M, Kempa-Liehr A, Feindt M (2016) Distributed and parallel time series feature extraction for industrial big data applications. [arXiv:1610.07717](https://arxiv.org/abs/1610.07717)
- Theumer P, Zeiser R, Trauner L, Reinhart G (2021) Anomaly detection on industrial time series for retaining energy efficiency. *Proc CIRP* 99:33–38. <https://doi.org/10.1016/j.procir.2021.03.006>
- Zhou Y, Xue W (2018) Review of tool condition monitoring methods in milling processes. *Int J Adv Manuf Technol* 96:2509–2523. <https://doi.org/10.1007/s00170-018-1768-5>

18. Axinte D, Gindy N (2004) Assessment of the effectiveness of a spindle power signal for tool condition monitoring in machining processes. *Int J Prod Res* 42(13):2679–2691. <https://doi.org/10.1080/00207540410001671642>
19. Wang X, Williams RE, Sealy MP, Rao PK, Guo Y (2018) Stochastic modeling and analysis of spindle power during hard milling with a focus on tool wear. *ASME J Manuf Sci Eng* 140(11):111011. <https://doi.org/10.1115/1.4040728>
20. Zhou X, Wu H, Rojas J, Xu Z, Li S (2020) Nonparametric Bayesian learning for collaborative robot multimodal introspection. Springer. <https://doi.org/10.1007/978-981-15-6263-1>
21. Fox E, Sudderth E, Jordan M, Willsky A (2009) A sticky HDP-HMM with application to speaker diarization. *Ann Appl Stat*. <https://doi.org/10.1214/10-AOAS395>
22. Zhang S, Chen X, Chen J, Jiang Q, Huang H (2020) Anomaly detection of periodic multivariate time series under high acquisition frequency scene in IoT. In: International conference on data mining workshops (ICDMW), pp 543–552. <https://doi.org/10.1109/ICDMW51313.2020.00078>
23. Netzer M, Palenga Y, Gönneheimer P, Fleischer J (2021) Offline-online pattern recognition for enabling time series anomaly detection on older NC machine tools. *J Mach Eng*. <https://doi.org/10.36897/jme/132248>
24. Netzer M, Michelberger J, Fleischer J (2020) Intelligent anomaly detection of machine tools based on mean shift clustering. *Proc CIRP* 93:1448–1453. <https://doi.org/10.1016/j.procir.2020.03.043>