Towards predictive part quality and predictive maintenance in industrial machining - a data-driven approach

Laura Rudolph
Karlsruhe Institute of Technology
la.s.rudolph@googlemail.com

Jennifer Schoch
Daimler AG
jennifer.schoch@daimler.com

Hansjoerg Fromm
Karlsruhe Institute of Technology
hansjoerg.fromm@kit.edu

Abstract

Programs such as Industry 4.0 and Internet of Things contain the promise of "intelligent production" with "smart services". In fact, great advances have already been made in sensor technology and machine connectivity. Production plants continuously generate and communicate large amounts of data and have become "cyber-physical systems". However, the task of gaining knowledge from these large amounts of data is still challenging. Data generated by numerical control (NC) and programmable logic controllers (NC) comes in a raw format that doesnt allow the application of analytical methods directly. Extensive preprocessing and feature engineering has to be applied to structure this data for further analysis. An important application is the timely detection of deviations in the production process which allows immediate reactions and adjustments of production parameters or indicates the necessity of a predictive maintenance action. In our research, we aimed at the identification of special deviant behavior of a grinding machine based on NC data. One finding wast the distinguishing the warm-up program from regular production and the other to recognize imprecise identification of the grinding process window. Both tasks could be solved with extensive preprocessing of the raw data, appropriate feature extraction and feature reduction, and the subsequent application of a clustering algorithm.

1. Introduction

Programs such as Industry 4.0 and Internet of Things contain the promise of "intelligent production" with "smart services". In fact, great advances have already been made in sensor technology and machine connectivity. Production plants continuously generate and communicate large amounts of data. These "digital data streams" [1] are considered as invaluable resources for more intelligent industrial services [2]. The extraction of knowledge and insight from raw data, however, is not a trivial task and requires considerable efforts in signal processing and advanced data analytics. Only the "marriage" of machines with analytics can provide benefits that are "multiple and significant" [3]. Lade et al. [4] highlight some of these benefits for manufacturing: reduction of test time and calibration, quality improvement, reduction of warranty cost, yield improvement, the possibility the predictive maintenance.

In this research, the fundamental steps for processing digital data streams in a manufacturing environment have been investigated exemplarily for a profile and gear grinding machine. Compared to standard business analytics, the challenges of applying analytics in manufacturing are largely in the steps that precede the application of classical data mining or machine learning algorithms. Data sources are typically the numerical control units (NC) and programmable logic controllers (PLC) of the machinery. The possibilities to get closer to the visions of intelligent production and smart services with the help of NC and PLC data are manifold.

About 80 % of the effort of a data mining project is spent on explorative data analysis and data cleaning [5, S. ix]. Especially since NC and PLC data are time-dependent data, i.e. time series, complex feature engineering is necessary before intelligent approaches can gain hidden information from the data. The data on which this work is based originates from a profile and gear grinding machine. The process considered is the hard fine machining of double gears after hardening. The task of hard finishing is to generate the macro geometry (shape) and micro geometry (surface roughness) in the specified quality [6]. The tool for continuous generating grinding is a grinding worm with rack and pinion profile. During machining, the worm and part rotate synchronously. The tool is engaged in the part by feed in X-direction. [7] The interaction of tool and part can be seen in Figure 1.

The machining process consists of five steps, two roughing strokes, one finishing stroke and two polishing strokes. The machine has twelve axes, of which the data

URI: https://hdl.handle.net/10125/63933
978-0-9981331-3-3

Figure 1. Tool (left) and part (right) in generating grinding [8]

of eight axes are taken into account. Designation and orientation of all axes can be found in Table 1.

With the help of advanced analytics, deviations in the production process can be detected in order to be able to react adequately if and when required.

2. Related Work

Some studies dealt with the prediction of part quality, especially surface quality, on the basis of temporally constant production parameters as e.g. the depth of cut, feed rate and cutting speed. Various methods of machine learning were used. Thus, Azouzi and Guillot [10], Zain et al. [11], Benardos and Vosniakos[12] and Asiltürk and Çunkaş [13] used artificial neural networks (ANN) for prediction. In addition, Asiltürk and Çunkaş [13] compared the result of the ANN with the result of a regression. Correa et al. [14] classified the surface condition of a component in a milling process using Bayesian networks. Tseng et al. [15] used Fuzzy Logic to develop a formula for predicting the surface finish in a machining process.

In addition to fixed input parameters, Abouelatta and Madl [16] and Axinte et al. [17] also used time series of tool vibrations and acoustic signals. The interpretation of the results was done manually and not with the help of a learning approach.

Chen and Savage [18] and Tsai et al. [19] considered both fixed and time-dependent input variables and used Fuzzy Net approaches for modelling. In Salgado et al. [20] a Least Squares Support Vector Machine was used to predict the surface condition of a part in a turning process. Plaza and López [21] exclusively considered vibration signals and modelled the part surface condition using multiple linear regression. Keak and Song [22] considered acoustic emissions and an ANN for fault diagnosis in a grinding process.

No work could be found in the literature which takes into account NC data such as displacements, drive currents and active drive power of the axes as input variables. Beyond the idea of using NC and PLC data as basis for smart analytics and services Lee et al. [23] point out that data is not useful unless it is processed so that it provides a context. This applies especially for time dependent data.

3. Approach

Our approach is inspired by the CRISP-DM model, but has been further refined based on the data available and the fine-grained steps in data processing.

In step 1, the source of the data and reading process is described briefly. Additionally, the structure and characteristics of the data set are explained.

In step 2, the data is preprocessed. The necessary steps of the data cleaning as well as the elimination of quality problems in the data set are carried out.

In step 3, the extensive feature extraction is performed. For this purpose, the information available in the time series is compressed in several features.

In step 4, the information content of the feature data set is improved by identifying irrelevant and redundant features and removing them from the data set.

In step 5, the production processes of individual components are grouped using k-means cluster analysis.

Figure 2 visualizes the approach.

Figure 2. Five-step approach for the advanced analysis of NC data

The processing takes place in R and RStudio. R is a programming language that is suitable for the statistical calculations to be carried out. A large number of online packages contain ready-made functions for transforming and analyzing the data with regard to different questions.

3.1. Data Source

With the help of an external device, the data of the NC and PLC control are read out and stored in a backend. The read-out data contains high-frequency sampled NC data (sampling rate 100 Hz) as well as low-frequency sampled PLC data.

The NC data record contains an ID which is incremented for each data point read, a time stamp indicating the date and time of the data point, as well as the values of the axis positions (actual and nominal), drive currents and active drive power for the 8 axes under consideration. The values of the axis positions

Designation	Movement	Orientation	Description
A1	rotatory	horizontal	swivel movement tool
B1	rotatory	horizontal	rotary movement tool
C2	rotatory	vertical	rotary movement workpiece
U1	translational	horizontal	axial movement lubricant tube
V1	translational	horizontal	tangential movement tool
X1	translational	horizontal	radial movement column slide
Z1	translational	vertical	axial movement tool
Z4	Translational	Vertical	vertical movement counterholder

Table 1. Axes in profile and gear grinding machine [9]

are given for all translational axes (U1, V1, X1, Z1, Z4) in millimeters [mm] and for all rotary axes (A1, B1, C2) in degrees []. Drive currents are given in amperes [A] and active drive power in watts [W]. Some PLC data are also relevant for the evaluation of the NC data. The low-frequency sampled PLC data again contains an ID which is incremented for each read data point and a time stamp indicating the date and time of the data point. For each of these data points, a bit indicates whether production is running (1) or not (0). The part identifier of the currently produced part is also stored.

3.2. Data Preprocessing

The explorative data analysis and data cleansing causes about 80 % of the effort for a data mining project [5, S. ix]. The data preprocessing and feature extraction step is used to prepare the data for modelling. During preprocessing, data quality problems are solved and the data is prepared in such a way that characteristic features from the raw data time series can be calculated in the feature engineering step.

Regarding data quality, two aspects are particularly relevant. On the one hand the handling of missing values in the data set, on the other hand the equidistant distribution of data points.

The following steps were made during data preprocessing:

1. removal of irrelevant information from the record
2. conversion to correct data formats
3. correction of fragmented data series
4. sample rate conversion by interpolation
5. identification of production periods
6. removal of random signal jumps
7. calculation of relative machining times

3.3. Feature Extraction

In order to train a model, a structured data set must be created. This should correspond to a $n \times m$ matrix. The rows correspond to the n observations, meaning the parts produced, and the columns to the m features. The feature engineering serves to break down the information, which is in the different time series, into meaningful features. It is possible to generate features directly from the preprocessed original data. In addition, signal transformations are also necessary for the characterization of the signal.

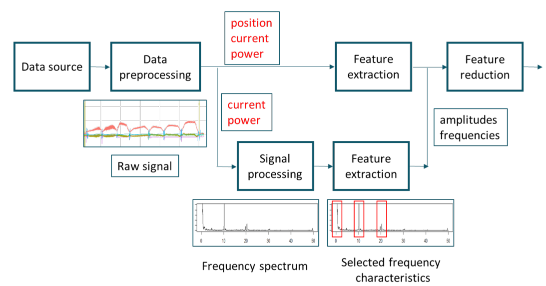

Figure 3. Signal processing steps before the application of machine learning

Raw signal

To characterize a time series, the raw signal can be described using basic statistical location parameters such as minimum, maximum, mean, median, and standard deviation. In addition, the length of the time series, i.e. the duration of the production, is highly characterizing. Position deviations can be determined on the basis of the nominal and actual values of the axis positions. These deviations over time are compressed to individual features using maximum, minimum, median and the mean squared error. In addition, for all other signal curves, deviations from a standard curve are calculated with the help of the mean squared error. For this purpose, the signal curves are adapted to the norm signal using cross correlation.

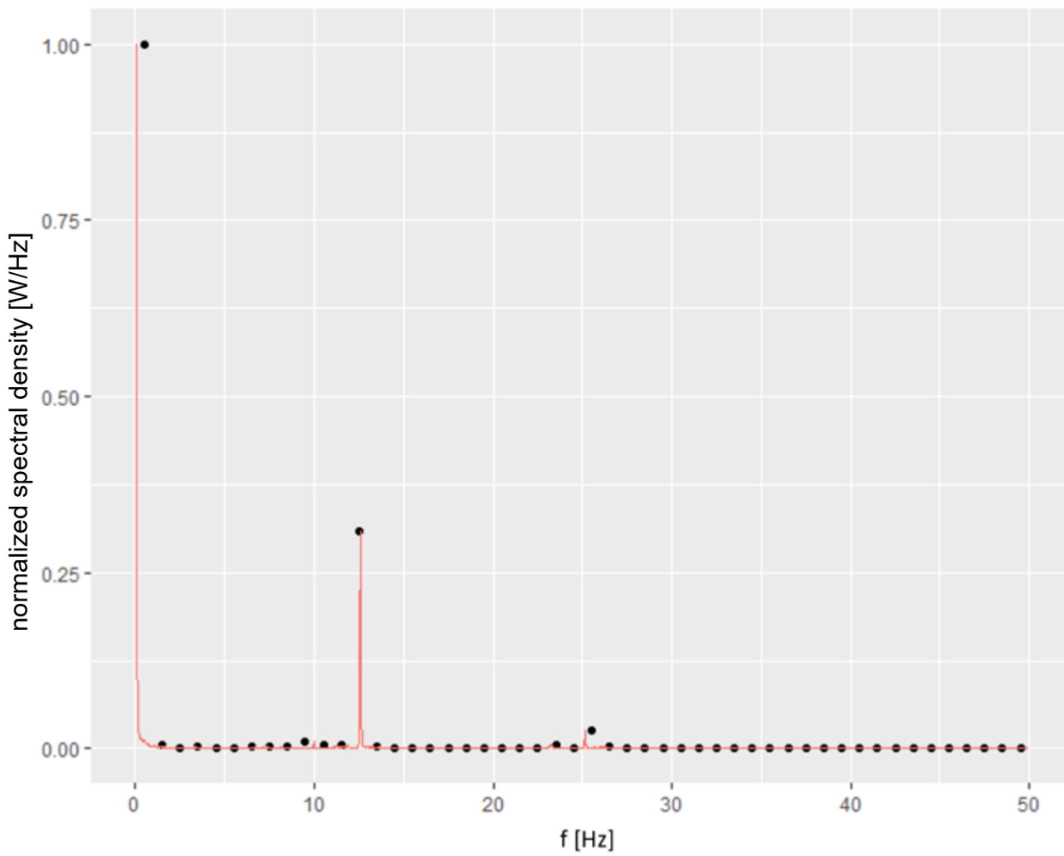

Figure 4. Frequency spectrum of active drive power C2-axis and associated bin information

In addition, further advanced attributes for time series description are calculated: several features based on the autocorrelation function, number of mean crossings, standard deviation of the first derivative, stability, flat spots, percentage of random walker-crossings and the maximum mean and variance differences of consecutive time windows.

Spectral Analysis

In addition to the time domain, signals can also be analyzed in the frequency domain [24, p.10]. The description of the frequency and time domain is equivalent. By a Fourier transformation the signal can be converted from time to frequency domain. [25, p.25]. In Fourier analysis, signals are decomposed into individual sinusoidal components with different wavelengths [26, p.31].

According to Nyquist's sampling theorem, an upper cutoff frequency f_g must apply to the sampling rate $f_A \geq 2f_g$ [25, p.68f]. Thus, the sampling rate of $f_A = 100$ Hz results in an upper frequency limit of $f_g = 50$ Hz. This means that all functions whose frequencies lie between 0 Hz and 50 Hz are uniquely determined by a sampling rate of 100 Hz. The result of the Fourier transformation is the spectral density of 0 Hz - 50 Hz in steps of 0,01 Hz. Every time dependant raw signal was transformed in a frequency dependant signal with 5000 data points. It would be a possibility to consider these 5000 frequency values per spectrum as 5000 individual features and the corresponding spectral density of each part produced as characteristic of those features.

However, this would result in a high number of features, many of which contain little information. This would lead to the so-called *Curse of Dimensionality*. This basic principle in the analysis of high dimensional data states that adding features that have no real meaning for the response results in a deterioration of the model [27, p.242f]. Therefore, the information is summarized in so-called bins. This means that the spectrum is divided into 50 bins with a width of 1 Hz and the maximum value within the bin is recorded as a characteristic. Thus, the signal was broken down to 50 features. This is visualized in Figure 4.

The solid line shows the signal obtained by the Fourier transformation with 5,000 data points; the points are the 50 values obtained by the division into bins. Thus, the number of features was significantly reduced (1 % of the original set), deviations in the spectrum can still be detected. Since the frequency range is the same for all signals (0-50 Hz), the signals are also compared by calculating the mean correlation of the spectra to the spectra of all other parts. Deviations in the spectrum should be noticeable by a low mean correlation.

Envelope and Amplitude

Considering the original signals, it shows certain fluctuations, i.e. the amplitude of the signal displays deviations. This offers the possibility to gain interesting insights through the amplitude curve of the signal. To compute the amplitude, the difference between the upper and lower envelope is calculated.

The course of the amplitude is characterized by statistical position parameters such as the minimum, maximum, mean and median as well as the mean squared error to an amplitude standard course.

Table 2 summarizes all calculated features.

Figure 5 gives an overview of the feature extraction steps and the corresponding number of features generated. In total, we constructed a structured data set with 633 features.

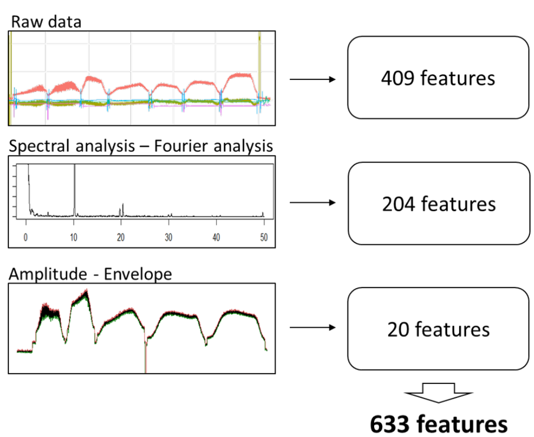

Figure 5. Number of calculated features

3.4. Feature Reduction

Many data sets contain irrelevant and redundant data that negatively affects the performance of the learning algorithm. This explains the need for dimensionality reduction in high-dimensional data [28]. The first step removes all features that could not be calculated for

Feature	Description
Time	Duration of part processing
Min	Minimum value
Max	Maximum value
Mean	Mean value
Median	Median value
Sd	Standard deviation
x_acf1	First autocorrelation coefficient of time series
x_acf10	Sum of squares of the first 10 autocorrelation coefficients of time series
diff1_acf1	First autocorrelation coefficient of first derivation
diff1_acf10	Sum of squares of the first 10 autocorrelation coefficients of first derivation
diff2_acf1	First autocorrelation coefficient of second derivation
diff2_acf10	Sum of squares of the first 10 autocorrelation coefficients of second derivation
walker_propcross	Percentage of walker-crossings
crossing_points	Number of times the mean line is crossed
stability	Variance of means of tiled windows
max_level_shift	Largest mean shift between two consecutive sliding windows
time_level_shift	Associated time index
max_var_shift	Largest variance shift between two consecutive sliding windows
time_var_shift	Associated time index
flat_spots	Flat spots using discretization
std1st_der	Standard deviation of the first derivative
mse	Mean squared error
Pos_dis_min	Minimum position deviation
Pos_dis_max	Maximum position deviation
Pos_dis_mean	Mean position deviation
Pos_dis_median	Median position deviation
freq_x	Maximum spectral density at frequency bin x
mean_corr	Mean pairwise correlation of part's spectrum
ampl_min	Minimum amplitude
ampl_max	Maximum amplitude
ampl_mean	Mean amplitude
ampl_median	Median amplitude
mse_ampl	Mean squared error amplitude

Table 2. Summary of calculated features

at least one part, meaning all features in the matrix where at least one entry is NA. Then all features with a variance of $var = 0$ are removed. These features do not offer any information, and, in addition, a scaling of the data cannot take place which would be necessary for a principal component analysis. In the final step, a filter approach is used to identify and remove redundant features. All features with a pairwise correlation of $> 0,95$ are identified. To decide which of the correlated features will be removed the absolute values of the correlations are considered and the variable with a higher average absolute correlation is removed. These activities remove 44 of the 633 existing features in the first step, 11 in the second step, and 206 in the last step. This reduces the feature record from 633 features to 373 features. This corresponds to 58.9% of the originally calculated number of features.

3.5. Clustering/Modelling

In the previous steps, we created a structured data set that can be used as input data for machine learning algorithms. The data available should further be analyzed in order to get valuable insights. Clustering methods can be used to identify groupings in the data set and thus, support a better data understanding. K-means is one of the most basic partitioning clustering techniques, which is why it was chosen as the approach for initial modeling. To maintain cluster interpretability, no principal component analysis (PCA) is applied before the data is modelled. Since the result of k-means depends on the initial choice of cluster centers, each model is calculated 20 times and the best solution is chosen. The selection of the optimal number of clusters is based on the silhouette width. This criterion was published by [29] and evaluates the positioning of a data point in relation to the cluster assigned. On the one hand, the proximity to the other points in the same cluster ($a(i)$) and the distance to the points of the nearest cluster ($b(i)$) are calculated. The silhouette $s(i)$ for each data point is calculated as follows

$$s(i) = \frac{b(i) - a(i)}{max\{a(i), b(i)\}}$$

and is in the range of [-1,1], where 1 is the assignment to an appropriate cluster and -1 is the assignment to an inappropriate cluster. If the silhouette value of the object i of cluster A is close to 0, this means that the object can also be in the nearest neighboring cluster of A. A clustering is characterized using the silhouette width by calculating the average value of all objects to be classified. In a plot, the silhouette value can be visualized for each cluster number. The maximum corresponds to the appropriate number of clusters.

4. Results

After the steps of preprocessing, feature extraction and feature reduction a cluster analysis is performed in order to gain further knowledge. For this purpose, the number of cluster is determined and cluster analysis is run and its result is described. This is followed by a root cause analysis for the cluster separation and a final interpretation of clusters.// In order to determine the optimum number of clusters, the silhouette width is calculated. Figure 6 shows the result of the mean silhouette width for the training data set and $k = 2, ..., 50$.

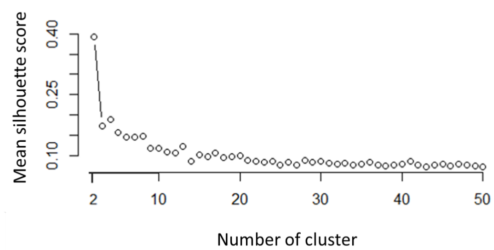

Figure 6. Silhouette width for different numbers of clusters k

The maximum can be found at k = 2, which is why this number of clusters is chosen. The model is then generated using the k-means algorithm. Each observation is assigned to one of the two clusters.

There are 695 parts (92.54 %) in cluster 1 (spheres) and 56 parts (7.46 %) in cluster 2 (triangles). Thus, it can be said, that time series assigned to cluster 1 represent a "normal" production process. Time series in cluster 2 show might deviate from the normal production process. In order to gain insights in which axis occurances of deviations to a normal signal can be found, the root causes for this cluster separation are further investigated. This means features that are leading in the cluster separation should be identified in order to draw conclusions which raw signal is affected. For a better visualization a principal component analysis is conducted (see figure 7). Figure 7 shows that principal component 1 in particular clearly separates the clusters. In order to draw a conclusion from the principal component to the feature and thus to the original signal, the loadings of principal component 1 are considered. Loadings are the weights of the original variables used in the calculation of the principal components [30].

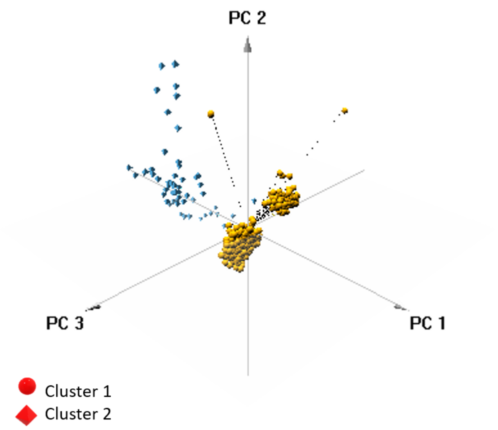

Figure 7. Visualization of the data set using the first three principal components

$$Y_1 = a_{11}X_1 + a_{12}X_2 + ... + a_{1p}X_p$$

where X_p is an original variable and $a_{i,j}$ are the associated loadings. This means that the higher the loading, the greater the influence of the original variable on the principal component.

Table 3 lists the first five features that have the greatest influence on the principal component.

Feature	Abs.loading PC1
mean_corr_ drive.power _B1.axis	0.1128158
mean_corr_ drive.power _C2.axis	0.1114471
drive.power _B1.axis_sd	0.1094315
drive.current _B1.axis_x_acf1	0.1047473
drive.power_Z1.axis_min	0.1044609

Table 3. Loadings of first principal component and associated feature, sorted in decreasing order

The feature *mean_corr_drive.power_B1.axis* describes the mean correlation of the active drive power spectrum of the B1 axis to the spectra of the other analyzed parts. Figure 8 shows the distribution of the values of this feature over all parts in the dataset. The assignment of the respective parts to the clusters is also shown in this figure. Another feature (*drive.power_B1.axis_sd*), which is based on the active drive power of the B1 axis, has the third largest loading amount for principal component 1. Figure 9 shows the distribution of this feature over all parts including the marking of cluster membership.

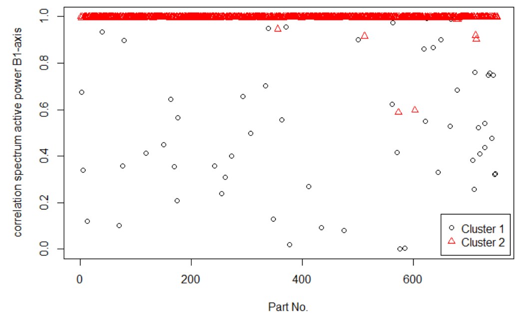

Figure 8. Distribution of mean_corr_drive.power_B1.axis feature

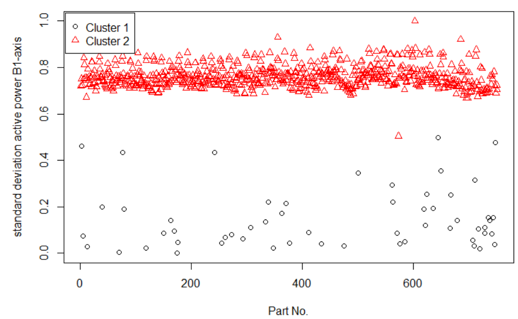

Figure 9. Distribution of drive.power_B1.axis_sd feature

The figures clearly show that these two features ensure good separation between the two clusters. Figures 10 and 11 show the original signal of the active drive power of the B1 axis, as well as the median and mean values. The signals are divided into clusters, Figure 10 are raw signals from parts assigned to cluster 1 and Figure 11 are raw signals from parts assigned to cluster 2.

The signal curve of the signals from cluster 1 (Figure 10) shows the characteristic five peaks of the five processing steps. The signals from cluster 2 (Figure 11) do not have this pattern. In order to obtain a qualitative evaluation of the signals, a manual interpretation of the signals and assignment to real production processes is necessary. Experts with detailed knowledge of the machine's processes and the movement of the axes stated that the corresponding data could not be the machining of a workpiece.

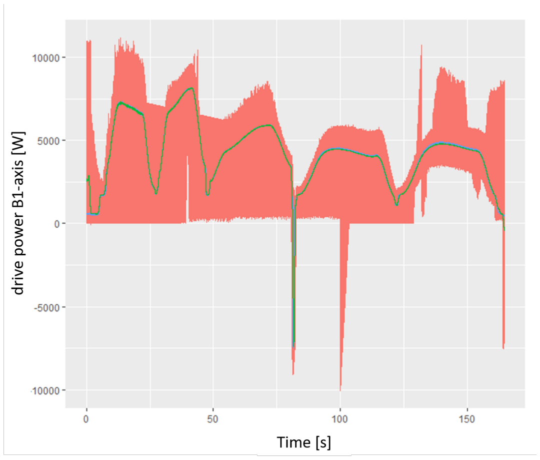

Figure 10. Raw signals of active drive power B1-axis of signals assigned to cluster 1. The line describes the median and mean value

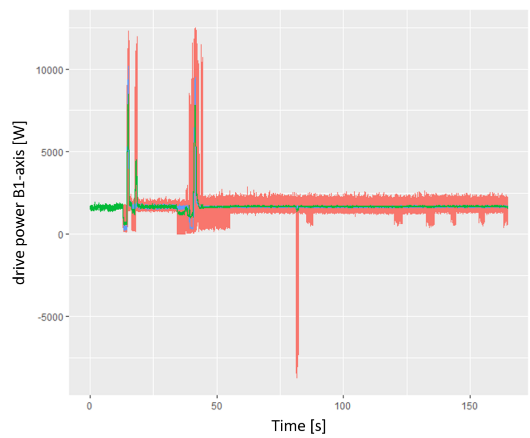

Figure 11. Raw signals of active drive power B1-axis of signals assigned to cluster 2. The line describes the median and mean value

5. Conclusion

An extensive literature review was conducted to determine the coverage of NC and PLC data analysis in the area of service analytics. In addition, the search was further specialized in the use case of part classification based on production data to determine part quality. In that research no work could be identified using NC and PLC data for prediction. Only fixed machine parameters as well as acoustic emissions and vibrations were used for the analysis. Compared to other applications of big data analysis, the analysis of machine data generated by numerical control (NC) and programmable logic controllers (NC) requires significant efforts of preprocessing, before traditional data mining and machine learning methods can be used. We have exemplarily demonstrated this for a profile and gear grinding machine.

Subsequently, the raw data was qualitatively processed and an extensive feature engineering was carried out in order to reduce the existing multivariate time series to meaningful features. Thereafter, the generated data set was cleaned of irrelevant and redundant features. A cluster analysis was performed with the generated data set of 373 features and 751 observations. The silhouette width was used to determine the optimal number of clusters and the k-means algorithm was applied to identify two clusters. By interpreting the partitioning principal component, it was possible to identify raw signals that cannot correspond to a productive period. This means, using cluster analysis, it was found that the section of PLC data that indicates active production is also set for other machine movements that do not correspond to component production. By assigning further data to the identified clusters, a separation into production data and other machine data can be made. Therefore, By means on k-means clustering as well as principal component analysis, our results show that the warm-up program as well as imprecise identification of the grinding process window can be clearly identified. Therefore, we conclude that for NC-data analysis of industrial processes a precise identification of relevant samples, data and process understanding is crucial.

References

[1] G. Piccoli and F. Pigni, "Harvesting external data: The potential of digital data streams.," *MIS Quarterly Executive*, vol. 12, no. 1, 2013.

[2] M. M. Herterich, F. Uebernickel, and W. Brenner, "The impact of cyber-physical systems on industrial services in manufacturing," *Procedia Cirp*, vol. 30, pp. 323–328, 2015.

[3] P. C. Evans and M. Annunziata, "Industrial internet: Pushing the boundaries," *General Electric Reports*, 2012.

[4] P. Lade, R. Ghosh, and S. Srinivasan, "Manufacturing analytics and industrial internet of things," *IEEE Intelligent Systems*, vol. 32, no. 3, pp. 74–79, 2017.

[5] T. Dasu and T. Johnson, *Exploratory data mining and data cleaning*, vol. 479. John Wiley & Sons, 2003.

[6] F.-W. Bach and K. Kerber, *Prozesskette Präzisionsschmieden*. Berlin Heidelberg New York: Springer-Verlag, 2014. aufl. ed., 2013.

[7] 3M, "Kontinuierliches Wälzschleifen von Verzahnungen." Online available under: `https://www.schleifprofi.com/schleifverfahren/zahnflankenschleifen/kontinuierliches-waelzschleifen-von-verzahnungen/#kontinuierliches-walzschleifen`, accessed 2019/02/19, o.J.

[8] Liebherr, “Verzerrungsfreies Wälzschleifen Präzisiert Zahnflanken.” https://www.liebherr.com/de/che/aktuelles/news-pressemitteilungen/detail/verzerrungsfreies-w%C3%A4lzschleifen-pr%C3%A4zisiert-zahnflanken.html#lightbox, accessed 2019/02/19, 2016.

[9] Liebherr, “Die CNC-Wälz- and Profilschleifmaschinen LCS 200 - 500,” 2016. Online available under: https://www.liebherr.com/shared/media/gear-technology-and-automation/brochures/gear-cutting-machines/liebherr-w%C3%A4lz-und-profilschleifmaschinen-lcs200-500.pdf, accessed 2019/01/15.

[10] R. Azouzi and M. Guillot, “On-line prediction of surface finish and dimensional deviation in turning using neural network based sensor fusion,” *International Journal of Machine Tools and Manufacture*, vol. 37, no. 9, pp. 1201–1217, 1997.

[11] A. M. Zain, H. Haron, and S. Sharif, “Prediction of surface roughness in the end milling machining using artificial neural network,” *Expert Systems with Applications*, vol. 37, no. 2, pp. 1755–1768, 2010.

[12] P. Benardos and G. C. Vosniakos, “Prediction of surface roughness in cnc face milling using neural networks and taguchi’s design of experiments,” *Robotics and Computer-Integrated Manufacturing*, vol. 18, no. 5-6, pp. 343–354, 2002.

[13] I. Asiltürk and M. Çunkaş, “Modeling and prediction of surface roughness in turning operations using artificial neural network and multiple regression method,” *Expert systems with applications*, vol. 38, no. 5, pp. 5826–5832, 2011.

[14] M. Correa, C. Bielza, M. d. J. Ramirez, and J. R. Alique, “A bayesian network model for surface roughness prediction in the machining process,” *International Journal of Systems Science*, vol. 39, no. 12, pp. 1181–1192, 2008.

[15] T.-L. B. Tseng, U. Konada, and Y. J. Kwon, “A novel approach to predict surface roughness in machining operations using fuzzy set theory,” *Journal of Computational Design and Engineering*, vol. 3, no. 1, pp. 1–13, 2016.

[16] O. Abouelatta and J. Madl, “Surface roughness prediction based on cutting parameters and tool vibrations in turning operations,” *Journal of materials processing technology*, vol. 118, no. 1-3, pp. 269–277, 2001.

[17] D. A. Axinte, N. Gindy, K. Fox, and I. Unanue, “Process monitoring to assist the workpiece surface quality in machining,” *International Journal of Machine Tools and Manufacture*, vol. 44, no. 10, pp. 1091–1108, 2004.

[18] J. Chen and M. Savage, “A fuzzy-net-based multilevel in-process surface roughness recognition system in milling operations,” *The International Journal of Advanced Manufacturing Technology*, vol. 17, no. 9, pp. 670–676, 2001.

[19] Y.-H. Tsai, J. C. Chen, and S.-J. Lou, “An in-process surface recognition system based on neural networks in end milling cutting operations,” *International Journal of Machine Tools and Manufacture*, vol. 39, no. 4, pp. 583–605, 1999.

[20] D. R. Salgado, F. Alonso, I. Cambero, and A. Marcelo, “In-process surface roughness prediction system using cutting vibrations in turning,” *The International Journal of Advanced Manufacturing Technology*, vol. 43, no. 1-2, pp. 40–51, 2009.

[21] E. G. Plaza and P. N. López, “Surface roughness monitoring by singular spectrum analysis of vibration signals,” *Mechanical Systems and Signal Processing*, vol. 84, pp. 516–530, 2017.

[22] J.-S. Kwak and J.-B. Song, “Trouble diagnosis of the grinding process by using acoustic emission signals,” *International Journal of Machine Tools and Manufacture*, vol. 41, no. 6, pp. 899–913, 2001.

[23] J. Lee, E. Lapira, B. Bagheri, and H.-a. Kao, “Recent advances and trends in predictive manufacturing systems in big data environment,” *Manufacturing letters*, vol. 1, no. 1, pp. 38–41, 2013.

[24] Z. M. Hussain, A. Z. Sadik, and P. O’Shea, *Digital Signal Processing - An Introduction with MATLAB and Applications*. Berlin Heidelberg: Springer Science & Business Media, 2011.

[25] C. Roppel, *Grundlagen der Nachrichtentechnik-Übertragungstechnik und Signalverarbeitung*. München: Carl Hanser Verlag GmbH Co KG, 2018.

[26] D. Sundararajan, *Fourier Analysis–a Signal Processing Approach*. SpringerLink: Bücher, Singapore: Springer, 2018.

[27] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An introduction to statistical learning*, vol. 112. Springer, 2013.

[28] L. Yu and H. Liu, “Feature selection for high-dimensional data: A fast correlation-based filter solution,” in *Proceedings of the 20th international conference on machine learning (ICML-03)*, pp. 856–863, 2003.

[29] P. J. Rousseeuw, “Silhouettes: a graphical aid to the interpretation and validation of cluster analysis,” *Journal of computational and applied mathematics*, vol. 20, pp. 53–65, 1987.

[30] S. M. Holland, “Principal components analysis (pca),” *Department of Geology, University of Georgia, Athens, GA*, pp. 30602–2501, 2008.