Innovations in
Manufacturing Technology

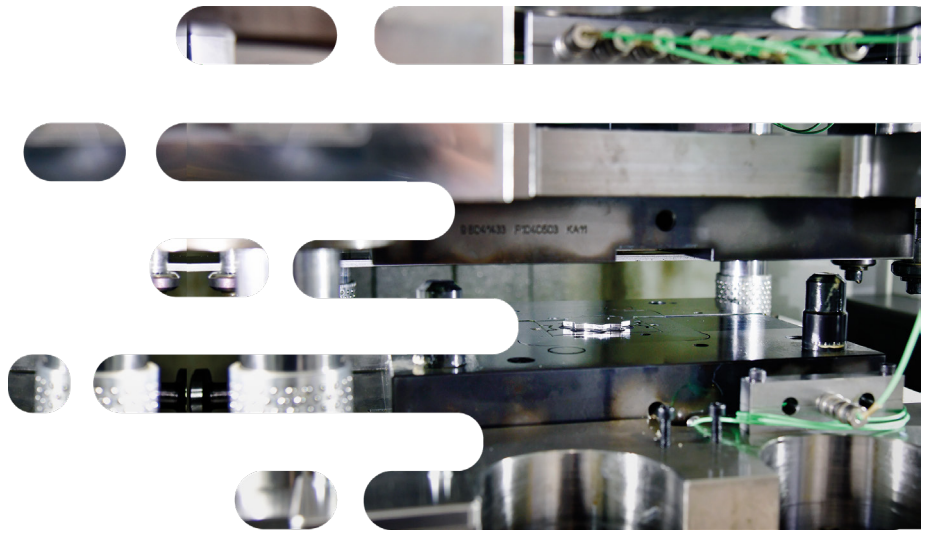
The increasing scarcity of energy and the resulting rise in energy prices significantly impact operating costs in production facilities. This trend is particularly pronounced in sheet-metal stamping, where the use of higher-strength steels for lightweight construction leads to higher process friction, which increases the cost pressure on already tight margins due to unpredictably high tool wear. Due to the closed tool design, direct condition monitoring of sheet-metal stamping tools is not feasible.
This dissertation addresses these challenges by presenting a systematic approach for data-driven indirect condition monitoring of sheet-metal stamping tools using acoustic emission signals. Acoustic emission sensors are suitable signal sources for tool condition monitoring due to their cost efficiency and high information content of the signals. However, acoustic emission signals are difficult to process and interpret due to their complexity.
For this dissertation, 398,047 fineblanking process cycles from industrial and laboratory settings were acquired, preprocessed, and analyzed using a combination of machine learning and signal processing methods. Through domain-specific preprocessing, a process image representing the friction conditions between sheet metal and punch was isolated and exploratively analyzed. This data analysis revealed systematic patterns within and between the collected datasets that indicate a correlation with punch wear. To model the wear of the punch, a critical tool component, the quality of the sheared surface of the scrap web was observed and evaluated as a proxy variable for the punch wear in the laboratory setting. This proxy variable was used to generate labels for supervised machine learning models and to predict these labels from the acoustic emission data of the shearing and stripping phases of the fineblanking process. Evaluation of the models from a model performance and an explainable AI perspective, as well as analyses of the punches, revealed that adhesive wear of the shearing edges of the punches is a likely root cause of sheared surface quality degradation of the scrap web. These results were used to develop a scalar indicator for the tool condition, and the evolution of this indicator was evaluated for the datasets from the industrial setting. The analyses showed that the developed indicator allows the identification of problematic sequences of fineblanking process cycles based on the strength of its dispersion.

ISBN 978-3-98555-259-7

Data-driven Indirect Tool Condition Monitoring in Sheet-metal Stamping

Martin Unterberg

Innovations in
Manufacturing Technology



# Data-driven Indirect Tool Condition Monitoring in Sheet-metal Stamping

Martin Unterberg

02/2025

MTI | RWTH AACHEN UNIVERSITY

Fraunhofer IPT

Apprimus

# Data-driven Indirect Tool Condition Monitoring in Sheet-metal Stamping

# Datengetriebene indirekte Zustandsüberwachung von Stanzwerkzeugen

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines
D o k t o r s d e r I n g e n i e u r w i s s e n s c h a f t e n
genehmigte Dissertation

vorgelegt von

Martin Unterberg

aus

Worms

**Berichter/in:**

Univ.-Prof. Wolfgang Prinz, PhD
Univ.-Prof. Dr.-Ing. Thomas Bergs

Tag der mündlichen Prüfung: 27. November 2024

Diese Dissertation ist auf den Internetseiten der Universitätsbibliothek online verfügbar.

INNOVATIONS IN
MANUFACTURING TECHNOLOGY

**Martin Unterberg**

# Data-driven Indirect Tool Condition Monitoring in Sheet-metal Stamping

**Herausgeber:**
Prof. Dr.-Ing. T. Bergs

Band 2/2025






**Bibliografische Information der Deutschen Nationalbibliothek**
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über https://portal.dnb.de abrufbar.

Martin Unterberg:

Data-driven Indirect Tool Condition Monitoring in Sheet-metal Stamping

1. Auflage, 2025

Gedruckt auf holz- und säurefreiem Papier, 100% chlorfrei gebleicht.

Copyright Apprimus Verlag, Aachen, 2025
Wissenschaftsverlag des Instituts für Industriekommunikation und Fachmedien
an der RWTH Aachen
Steinbachstr. 25, 52074 Aachen
Internet: www.apprimus-verlag.de, E-Mail: info@apprimus-verlag.de

Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, vorbehalten.

Printed in Germany

ISBN 978-3-98555-259-7

D 82 (Diss. RWTH Aachen University, 2024)

# Preamble and Acknowledgement

*Vorwort und Danksagung*

> *Und wenn der Sandsturm vorüber ist, wirst du kaum begreifen können, wie du ihn durchquert und überlebt hast. Du wirst auch nicht sicher sein, ob er wirklich vorüber ist. Nur eins ist sicher. Derjenige, der aus dem Sandsturm kommt, ist nicht mehr derjenige, der durch ihn hindurchgegangen ist.*
>
> DER JUNGE NAMENS KRÄHE
> *2002, somewhere*

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Werkzeugmaschinenlabor WZL der RWTH Aachen University. Die mündliche Prüfung absolvierte ich als wissenschaftlicher Mitarbeiter des Manufacturing Technology Institutes MTI der RWTH Aachen University, vormals Lehrstuhl für Technologie der Fertigungsverfahren des WZL. Herrn Prof. Dr.-Ing. Thomas Bergs, Leiter des MTI, danke ich herzlich für sein entgegengebrachtes Vertrauen sowie die fachliche und persönliche Förderung, die ich durch ihn erfahren habe. Herrn Prof. Wolfgang Prinz, PhD gilt mein Dank für die Betreuung und eingehende Durchsicht dieser Arbeit. Herrn Prof. Dr.-Ing. Ulrik Schroeder (Fachprüfer) und Herrn Prof. Dr. rer. nat. Torsten Kuhlen (Prüfungsvorsitz) danke ich für ihre Mitwirkung in der Prüfungskommission.

Das Bundesministerium für Wirtschaft und Klimaschutz BMWK trug durch die Förderung des Forschungsprojektes SPAICER substanziell zu meiner persönlichen und fachlichen Entwicklung bei und ermöglichte die dieser Arbeit zugrundeliegenden Forschungsarbeiten. Im Zusammenhang mit dem Projekt SPAICER möchte ich mich herzlich bei Herrn Jens Gerhard (im Projektzeitraum tätig bei Feintool System Parts Jena GmbH) und Herrn Dr.-Ing. Andreas Peters (Mendritzki Holding GmbH) bedanken. Die zahlreichen Gespräche und die gute Zusammenarbeit in den Themenfeldern Feinschneiden und Blechwerkstoff trugen wesentlich zu den gewonnenen Erkenntnissen in dieser Arbeit bei. Die für diese Arbeit nötigen Rechenressourcen wurden dankenswerterweise von der RWTH Aachen University im Rahmen des Projekts mit der Projekt-ID RWTH0840 bereitgestellt.

Herrn Dr.-Ing. Daniel Trauth gilt mein besonderer Dank für sein Vertrauen und die persönliche und fachliche Förderung während seiner Zeit als Oberingenieur am WZL, sowie in seiner Rolle als CEO der Senseering GmbH, in der er wesentlich zur Sammlung der industriellen Feinschneiddaten beitrug, die in dieser Arbeit verarbeitet wurden. Meinem Oberingenieur und Kollegen Philipp Niemietz, M. Sc. danke ich herzlich für die gute Zusammenarbeit und für die fachliche und persönliche Förderung während meiner Zeit als wissenschaftlicher Mitarbeiter sowie für die kritische Durchsicht des Manuskripts. Weiterhin möchte ich mich bei meinem Gruppenleiter Herrn Marco Becker, M. Sc. für die gute Führung der Forschungsgruppe Angewandtes Maschinelles Lernen und die kritische Durchsicht des Manuskripts danken. Frau Dr. rer. nat. Mia Kornely möchte ich herzlich für ihre kritische Durchsicht und die folgenden fachlichen Diskussionen danken, mit denen sie insbesondere zum Stand der

Technik dieser Arbeit wesentlich beitrug. Weiterhin danke ich Herrn Dr.-Ing. Herman Voigts herzlich für die aufmerksame Durchsicht des Manuskripts und die hilfreichen Anmerkungen. Herrn Dr.-Ing. Herman Voigts sowie Herrn Frank Schweinshaupt, M. Sc. möchte ich für ihre kollegiale, geduldige, und präzise Art danken, die mir ermöglichte, mich in ingenieurswissenschaftliche Fragestellungen im Bereich der Fertigungstechnik, insbesondere beim Feinschneiden, einzufinden.

Bei meinen Abteilungskolleginnen und -kollegen Herrn Tobias Kaufmann, M. Sc., Herrn Johannes Mayer, M. Sc. und Frau Lucia Ortjohann, M. Sc. möchte ich mich für die kritischen Diskussionen im Verlauf der Jahre zu fachlichen Themen sowie die Unterstützung während Projekten und die zahlreichen Gespräche und Aktivitäten außerhalb unserer beruflichen Tätigkeit bedanken.

Bei Herrn Carlos Millán Torres bedanke ich mich für seine Bereitschaft mich bei den experimentellen Untersuchungen für diese Arbeit zu unterstützen. Herrn Marcel Botz danke ich für seine Unterstützung in messtechnischen Fragestellungen. Herrn Thomas Fischer danke ich für seine Unterstützung bei der Durchführung der Rauheitsmessungen.

Meinen Eltern Angelika und Hans-Jürgen Unterberg möchte ich meine Dankbarkeit dafür aussprechen, dass sie mich in jeder Phase meines Lebens bedingungslos unterstützt und damit meinen persönlichen und beruflichen Werdegang ermöglicht haben.

Bei meiner Frau Lisa Ina Unterberg möchte ich mich für ihr Vertrauen und ihr entgegengebrachtes Verständnis während der Anfertigung dieses Manuskripts von ganzem Herzen bedanken. Dass du immer schon so sehr an mich geglaubt hast, hat mir diesen Weg überhaupt erst ermöglicht. Danke.

Meinen beiden Töchtern Ella und Fine Unterberg danke ich für ihre Neugier und Liebe, die mich jeden Tag aufs Neue inspiriert und antreibt. Ich bin unendlich dankbar dafür, euer Papa sein zu dürfen.

Troisdorf, Juni 2024 Martin Unterberg

*Für Ella und Fine.*

*Für Lisa, Angelika und Jürgen.*

# Table of contents

*Inhaltsverzeichnis*

**1 Introduction**
*Einleitung* **1**

**2 State of the art for data-driven tool condition monitoring in sheet-metal stamping**
*Stand der Technik zur datengetriebenen Zustandsüberwachung von Stanzwerkzeugen* **7**
2.1 Foundations of fineblanking
*Grundlagen des Feinschneidens* . . . . . 7
2.1.1 Process principles of fineblanking
*Verfahrensprinzipien des Feinschneidens* . . . . . 7
2.1.2 Tool wear mechanisms in sheet-metal stamping processes
*Werkzeugverschleißmechanismen bei Stanzprozessen* . . . . . 10
2.1.3 The necessity of indirect tool condition monitoring in sheet-metal stamping
*Die Notwendigkeit der indirekten Zustandsüberwachung von Stanzwerkzeugen* . . . . . 12
2.2 Foundations of statistics and machine learning
*Grundlagen der Statistik und des maschinellen Lernens* . . . . . 14
2.2.1 Key concepts and paradigms
*Wichtige Konzepte und Paradigmen* . . . . . 14
2.2.2 Introduction to selected algorithms and methods
*Einführung ausgewählter Algorithmen und Methoden* . . . . . 25
2.3 Foundations of time series analysis
*Grundlagen der Zeitreihenanalyse* . . . . . 32
2.3.1 Introduction to and key definitions in time series analysis
*Einführung in und wichtige Definitionen der Zeitreihenanalyse* . . . . . 32
2.3.2 Selected concepts and methods in time series analysis
*Ausgewählte Konzepte und Methoden in der Zeitreihenanalyse* . . . . . 33
2.4 State of the art for indirect tool condition monitoring of sheet-metal stamping tools
*Stand der Technik zur indirekten Zustandsüberwachung von Stanzwerkzeugen* 40
2.5 Identified research deficits
*Identifizierte Forschungsdefizite* . . . . . 42

**3 Objective and research methodology**
*Zielsetzung und Forschungsmethodik* **45**

**4 Generation, description, and preprocessing of the acoustic emission datasets**
*Generierung, Beschreibung, und Vorverarbeitung der Körperschall-Datensätze* **49**
4.1 Dataset overview
*Überblick über die Datensätze* . . . . . 49
4.2 Laboratory datasets
*Labordatensätze* . . . . . 51
4.2.1 Experimental setup
*Experimenteller Versuchsaufbau* . . . . . 51

4.2.2 Time series segmentation
*Zeitreihensegmentierung* . . . . . 53
4.2.3 Exploratory feature extraction
*Explorative Extraktion von Features* . . . . . 56
4.3 Industrial datasets
*Industriedatensätze* . . . . . 56
4.3.1 Experimental setup
*Experimenteller Versuchsaufbau* . . . . . 56
4.3.2 Time series segmentation
*Zeitreihensegmentierung* . . . . . 58
4.4 Feature extraction
*Feature Extraktion* . . . . . 61
4.5 Intermediate conclusion and unresolved challenges
*Zwischenfazit und offene Herausforderungen* . . . . . 62

**5 Exploratory analysis of industrial acoustic emission datasets**
*Explorative Analyse industrieller Körperschall-Datensätze* **65**
5.1 Structural evolution of low-dimensional acoustic emission signal embeddings
*Strukturelle Entwicklung niedrig-dimensionaler Körperschallsignal-Embeddings* . . . . . 65
5.1.1 Methodology
*Methodik* . . . . . 65
5.1.2 Results
*Ergebnisse* . . . . . 66
5.1.3 Intermediate conclusion and unresolved challenges
*Zwischenfazit und offene Herausforderungen* . . . . . 73
5.2 Structural evolution of domain-specific acoustic emission signal features
*Strukturelle Entwicklung domänen-spezifischer Feature von Körperschall-signalen* . . . . . 74
5.2.1 Methodology
*Methodik* . . . . . 74
5.2.2 Results
*Ergebnisse* . . . . . 75
5.3 Intermediate conclusion and unresolved challenges
*Zwischenfazit und offene Herausforderungen* . . . . . 85

**6 Observation and evaluation of a punch wear proxy variable**
*Beobachtung und Evaluierung einer Proxy-Variablen für den Stempel-verschleiß* **87**
6.1 Motivation and physical background
*Motivation und physikalischer Hintergrund* . . . . . 87
6.2 Experimental setup and methodology
*Experimenteller Aufbau und Methodik* . . . . . 88
6.3 Results
*Ergebnisse* . . . . . 91
6.3.1 Results of the preparative experiment L1
*Ergebnisse der Voruntersuchungen L1* . . . . . 91
6.3.2 Results of the preparative experiment L2
*Ergebnisse der Voruntersuchungen L2* . . . . . 95
6.3.3 Results of the main experiment L3
*Ergebnisse der Hauptuntersuchungen L3* . . . . . 98

6.4 Intermediate conclusion and unresolved challenges
*Zwischenfazit und offene Herausforderungen* . . . . . 103

**7 Modeling the punch wear proxy variable from acoustic emission data**
*Modellierung der Proxy-Variablen für den Stempelverscheiß anhand von Körperschallsignalen* **105**
7.1 Data experiment setup
*Experimenteller Aufbau des Datenexperiments* . . . . . 105
7.2 Results
*Ergebnisse* . . . . . 108
7.2.1 Results of the model performance evaluation
*Ergebnisse der Evaluierung der Modellgüten* . . . . . 109
7.2.2 Results of the feature importance evaluation
*Ergebnisse der Evaluierung der Feature Importances* . . . . . 110
7.3 Intermediate conclusion and unresolved challenges
*Zwischenfazit und offene Herausforderungen* . . . . . 116

**8 Transfer of experimental results to industrial datasets**
*Transfer der Ergebnisse auf die Industriedatensätze* **119**
8.1 Motivation and methodology
*Motivation und Methodik* . . . . . 119
8.2 Results
*Ergebnisse* . . . . . 121
8.3 Intermediate conclusion and unresolved challenges
*Zwischenfazit und offene Herausforderungen* . . . . . 124

**9 Summary and outlook**
*Zusammenfassung und Ausblick* **127**

**Bibliography**
*Literaturverzeichnis* **XIII**

**A Appendix**
*Anhang* **XXIII**
A.1 Appendix for Chapter 4 . . . . . XXIII
A.2 Appendix for Chapter 6 . . . . . XXIV
A.3 Appendix for Chapter 8 . . . . . XXV

**B Abstract**
*Kurzzusammenfassung* **XXVII**

# Symbols and abbreviations

*Formel- und Abkürzungsverzeichnis*

### Acronyms

| | |
|---|---|
| ABP | Absolute bandpower |
| AE | Acoustic emission |
| AECs | Autoencoders |
| AG | Aktiengesellschaft (stock corporation) |
| ANNs | Artificial neural networks |
| Cr | Chemical elements: Chromium |
| DC | Distance correlation |
| DFT | Discrete Fourier transform |
| DIN | Deutsches Institut für Normung |
| DSP | Digital signal processing |
| e.g. | Exempli gratia |
| EDX | Energy dispersive X-ray |
| etc. | Et cetera |
| FES | Finite element simulation |
| FFT | Fast Fourier transform |
| GBTs | Gradient-boosted trees |
| GPR | Gaussian process regression |
| i.e. | Id est |
| I1 | Industrial stroke series 1 |
| I2 | Industrial stroke series 2 |
| I3 | Industrial stroke series 3 |
| I4 | Industrial stroke series 4 |

| | |
|---|---|
| I5 | Industrial stroke series 5 |
| IFFT | Inverse fast Fourier transform |
| IIR | Infinite impulse response filter |
| KDDCup | Data mining and knowledge discovery competition |
| L1 | Laboratory stroke series 1 |
| L2 | Laboratory stroke series 2 |
| L3 | Laboratory stroke series 3 |
| Lasso | Least absolute shrinkage and selection operator |
| LR | Linear regression |
| MAD | Median absolute difference |
| MASS | Mueen's algorithm for similarity search |
| ML | Machine learning |
| MSE | Mean squared error |
| NDA | Non-disclosure agreement |
| PC | Principal component |
| PCA | Principal component analysis |
| PCC | Pearson's correlation coefficient |
| PS | Power spectrum |
| PSD | Power spectral density |
| RBP | Relative bandpower |
| RG | Research gap |
| RQ | Research question |
| RSS | Residual sum of squares |
| SEM | Scanning electron microscope |
| SHAP | Shapley additive explanation |
| SNR | Signal-to-noise ratio |
| SVD | Singular value decomposition |
| t-SNE | t-distributed stochastic neighbor embedding |

| | |
|---|---|
| TBP | Total bandpower |
| TCM | Tool condition monitoring |
| TDC | Top dead center |
| TSFEL | Time series feature extraction library |
| UMAP | Uniform manifold approximation and projection |
| USD | US Dollar |
| V | Chemical element: Vanadium |
| XAI | Explainable artificial intelligence |
| XGBoost | Extreme gradient boosting |

**Physical quantities and units**

| | | |
|---|---|---|
| $f_{\max}$ | Hz | Maximum frequency |
| $f_{\mathrm{Nyquist}}$ | Hz | Nyquist frequency |
| $f_{\mathrm{s}}$ | Hz | Sampling frequency |
| $F_{\mathrm{B}}$ | N | Blank holder force |
| $F_{\mathrm{C}}$ | N | Counter holder force |
| $F_{\mathrm{E}}$ | N | Ejector force |
| $F_{\mathrm{N}}$ | N | Normal force |
| $F_{\mathrm{P}}$ | N | Punch force |
| $F_{\mathrm{S}}$ | N | Stripping force |
| $F_{\mathrm{T}}$ | N | Tangential force |
| $P_{XX}(k)$ | $\mathrm{V}^2$ | Power spectrum at the $k$-th frequency bin |
| $R_{\mathrm{a}}$ | µm | Arithmetical mean roughness |
| $R_{\mathrm{m}}$ | $\mathrm{N/mm}^2$ | Tensile strength |
| $R_{\mathrm{p0.2}}$ | $\mathrm{N/mm}^2$ | Yield strength |
| $S_{XX}(k)$ | $\mathrm{V}^2/\mathrm{Hz}$ | Power spectral density at the $k$-th frequency bin |
| $v$ | mm/s | Velocity |

| | | |
|---|---|---|
| $\omega_c$ | Hz | Cutoff frequency of a digital filter |

**Mathematical symbols**

| | |
|---|---|
| $\mathbf{a}$ | A vector |
| $A$ | A set of all pairs of data points |
| $A_{j,k}, B_{j,k}$ | Centered distance matrices |
| $b$ | A scalar |
| $\mathbb{C}$ | Complex numbers |
| $\mathbf{C}$ | A covariance matrix |
| $\mathbf{D}$ | An orthogonal matrix |
| $D$ | A distance profile of a target and a query sequence |
| $\mathrm{dCor}$ | Population distance correlation |
| $\mathrm{dCor}_n$ | Sample distance correlation |
| $\mathrm{dCov}$ | Population distance covariance |
| $\mathrm{dCov}_n$ | Sample distance covariance |
| $\mathrm{dVar}$ | Population distance variance |
| $\mathrm{dVar}_n$ | Sample distance variance |
| $\mathbb{E}$ | Expectation operator |
| $E$ | Definition ML: Experience a computer program learns from |
| $e$ | EULER's number |
| $f$ | A function that describes a relationship between $\mathbf{X}$ and $\mathbf{y}$ |
| $\hat{f}$ | An estimate of $f$ |
| $f_\theta$ | A family of functions with different model parameters $\theta$ |
| $f_{\theta_0}$ | Theoretical optimal model from the family $f_\theta$ |
| $\mathbf{H}$ | Latent embedding layer (code) of an autoencoder |
| $i$ | In the context of Fourier transform (Eq. 41): The imaginary unit |
| $L(\theta)$ | Loss function $L$ as a function of model parameters $\theta$ |

| | |
|---|---|
| $M(\mathrm{MSE})$ | Denotes the median MSE across a fixed number of model runs |
| $M_{x_T}$ | Moving mean of the sequence $x_T(t)$ |
| $\mathbb{N}$ | Natural numbers |
| $N$ | Length of a time series |
| $n$ | Usually: Number of observations in a dataset |
| $O(g(n))$ | Asymptotic upper bound of algorithm time complexity (input $n$) |
| $\mathrm{obj}(\theta)$ | Objective function $\mathrm{obj}$ as a function of model parameters $\theta$ |
| $P$ | Definition ML: A performance measure for a computer program |
| $p$ | Usually: Number of features in a dataset |
| $\mathbb{R}$ | Real numbers |
| $\mathbf{R}$ | Reconstruction of the input of an autoencoder |
| $r$ | Sample Pearson correlation coefficient |
| $s$ | Number of randomly drawn synthetic labels |
| $s_k$ | Sample standard deviation of the $k$-th feature |
| $T$ | Definition ML: A class of tasks |
| $T$ | Decision trees: Number of leaves in a tree |
| $\mathbf{U}$ | An orthogonal matrix containing left singular vectors of $\tilde{\mathbf{X}}$ |
| $\mathbf{V}$ | An orthogonal matrix containing right singular vectors of $\tilde{\mathbf{X}}$ |
| $\mathbf{V}_k$ | A diagonal matrix containing the first $k$ columns of $\mathbf{V}$ |
| $w_j$ | Score on the $j$-th leaf of a decision tree |
| $\mathbf{X}$ | A model matrix |
| $\tilde{\mathbf{X}}$ | The centered model matrix $\mathbf{X}$ |
| $x(t)$ | A time series |
| $x_Q(t)$ | A query sequence |
| $x_T(t)$ | A target sequence |
| $\hat{x}(k)$ | A series of Fourier coefficients |
| $\bar{x}$ | Arithmetic mean of $x$ |
| $\check{\mathbf{X}}$ | Lower-dimensional embedding of $\mathbf{X}$ |

| | |
|---|---|
| $x_i$ | The $i$-th observation of a dataset |
| $X^{(i)}$ | A random variable |
| $\mathbf{X}^{(i)}$ | Realization of a multivariate random variable $X^{(i)}$ |
| $\mathbf{x}^{(i)}$ | Realization of a univariate random variable $X^{(i)}$ |
| $\mathbf{x}_k$ | The $k$-th feature across all observations of a dataset |
| $y(t)$ | Filtered output signal of a digital filter |
| $\mathbf{y}$ | Label vector of a data set |
| $\hat{\mathbf{y}}$ | Predicted labels |
| $y_i$ | The label corresponding to the $i$-th observation |
| $\mathbf{Z}$ | PCA embedding of $\mathbf{X}$ in an $n$-dimensional space |
| $\mathbf{Z}_k$ | PCA embedding of $\mathbf{X}$ in a $k$-dimensional space |
| $\mathbf{z}_k$ | Standardized $k$-th feature of a dataset |
| $z(t)$ | Result of the time-domain convolution of two sequences |
| $\beta_0$ | Intercerpt (Linear regression) |
| $\beta_1$ | Slope (Linear regression) |
| $\hat{\beta}_\lambda^L$ | Lasso coefficients |
| $\varepsilon$ | An error term |
| $\gamma$ | Penalty for the number of terminal nodes in a decision tree |
| $\lambda$ | A regularization parameter |
| $\lambda_i$ | $i$-th largest eigenvalue of a covariance matrix $\mathbf{C}$ |
| $\mu_{X^{(i)}}$ | Expected value of a random variable $X^{(i)}$ |
| $\nu$ | Similarity of data points (low-dimensional representation) |
| $\pi$ | The number Pi |
| $\rho$ | Pearson correlation coefficient for random variables |
| $\mathbf{\Sigma}$ | Diagonal matrix with singular values of $\tilde{\mathbf{X}}$ in descending order |
| $\Sigma_{x_T}$ | Moving standard deviation of the sequence $x_T(t)$ |
| $\sigma_i$ | $i$-th largest singular value of a centered model matrix $\tilde{\mathbf{X}}$ |

| | |
|---|---|
| $\sigma_{X^{(i)}}$ | Standard deviation of a random vector $X^{(i)}$ |
| $\hat{\theta}$ | Estimate of theoretical optimal $\theta_0$ based on the observed data |
| $\zeta$ | Similarity of data points (high-dimensional representation) |

# 1 Introduction

*Einleitung*

**Content** This dissertation focuses on indirect condition monitoring of sheet-metal stamping tools using acoustic emission data. The primary objective is to develop a novel robust framework for extracting, preprocessing, and analyzing acoustic emission signals during fineblanking, a sheet-metal stamping process. Integral to this framework are data mining and machine learning algorithms, which are indispensable for efficiently handling the complexity and high dimensionality of acoustic emission data. These tools are crucial for deriving accurate, predictive insights that enable operational improvements. In light of stringent sustainability targets and increasing energy and material costs, this research paves the way to advanced wear monitoring and thus optimized resource utilization in sheet-metal stamping processes, serving both economic and environmental objectives.

**Motivation** The EUROPEAN UNION intends to reduce greenhouse gas emissions to at least 55 % below 1990 levels until the year 2030 [EURO21a] and to be climate-neutral by 2050 [EURO21b]. Lightweight design in automotive represents an important building block to reducing $CO_2$ emissions by weight reduction, e.g., with the increasing usage of higher-strength steels [MCKI12] with a high strength-to-weight ratio in sheet-metal forming processes, such as sheet-metal stamping, which represents one of the most common manufacturing techniques for vehicle components [KLOC16]. The transition towards electric vehicles further accentuates this development, given the additional weight of batteries [JOOS12].

The increasing usage of higher-strength steels, in turn, leads to higher process forces and stresses on sheet-metal stamping tools [SHAN19], which result in excessive wear of critical tool components, such as the punch [VOIG21]. Excessive wear impairs the economic viability of sheet-metal stamping processes such as fineblanking since it leads to reject production and unplanned downtimes [UNTE21]. With already tight margins in sheet-metal stamping [KUBI22], the recent development in energy [STAT23a] and steel markets [STAT23b], especially in Germany, has further exacerbated the cost pressure on companies through increased operating and material costs.

To avoid reject production and to reduce unplanned downtimes and therefore optimize resource utilization in sheet-metal stamping, the tool wear must be controlled through timely maintenance. Understanding the severity of wear sufficiently fine-granularly is essential for this control. However, due to the closed tool design in sheet-metal stamping, the wear of critical tool components is not observable directly during the process.

**Research gap** Past approaches often relied on knowledge-based methods such as finite element simulations to design and operate sheet-metal stamping processes and to estimate suitable maintenance intervals. While these computational tools are crucial, they are limited by oversimplified assumptions that do not account for the variable conditions encountered in industrial environments [LIEW22]. Hence, numerous efforts were made recently by industry and research to monitor suitable quantities during sheet-metal stamping to infer the wear state (condition monitoring) and derive optimal maintenance intervals (predictive maintenance) with traditional signal processing methods and machine learning algorithms.

However, to this day, condition monitoring (and thus predictive maintenance) remains an open problem in sheet-metal stamping [UNTE23]. The reasons for this are diverse. MCKINSEY & COMPANY distilled main roadblocks that generally stand in the way of successful predictive

DEEPL (`https://www.deepl.com/translator/`) was used for translation purposes in this chapter.

maintenance programs [McKi21]. The roadblocks 'Insufficient data', 'Lack of capabilities', and 'Low economic return' (see Figure 1.1) will be addressed within this work as follows.

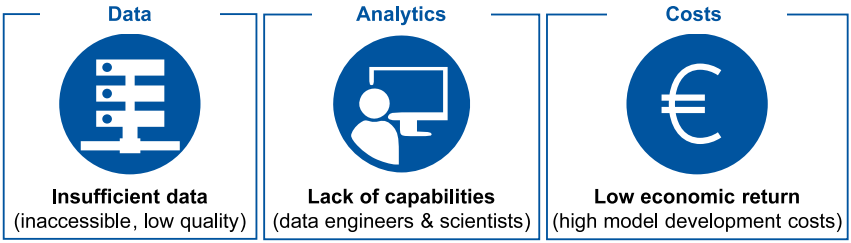



**Figure 1.1:** Roadblocks for successful predictive maintenance programs ([McKi21]) addressed within this dissertation.
*Hindernisse für erfolgreiche Projekte im Bereich der vorrausschauenden Wartung, die in dieser Dissertation behandelt werden.*

**Research approach** Within this dissertation, blueprints for data preprocessing and data analysis pipelines are developed to monitor the wear state of the punch during fineblanking in a laboratory (yet industrial-grade) and in an industrial fineblanking setting with acoustic emission signals. Acoustic emission sensors are particularly suitable for industrial sheet-metal stamping, due to their low cost and straightforward installation process. However, in contrast to force sensors, the signals are harder to interpret [Koll18] and preprocess due to the complexity of the resulting signals.

These challenges are managed by leveraging and combining suitable algorithms. A pipeline that leverages data mining techniques to systematically extract signal segments that contain information about the punch wear and subsequently uncover underlying patterns within the data is developed to address the problem of 'Insufficient data'. Furthermore, a proxy for the punch wear is proposed, observed, and evaluated in the laboratory setting and used to label a comprehensive dataset. This dataset is used to train and evaluate supervised machine learning models from both a performance and an explainable AI perspective. The results of the analyses are applied to the industrial datasets by developing a scalar tool condition monitoring indicator and evaluating its evolution throughout the datasets. It is shown that the indicator contains information to identify stroke series that were ultimately discontinued earlier than expected, thereby providing critical insights into the underlying dynamics of industrial fineblanking processes. These insights facilitate improved decision-making, enabling proactive maintenance strategies that can preempt costly failures and optimize production efficiency. The work documented in this dissertation thus directly addresses the challenges of 'Lack of capabilities' and 'Low economic return' in predictive maintenance programs by proposing and evaluating data-driven modeling approaches and pipelines and thus contributing to reducing the costs associated with model development.

**Contribution** Through the aforementioned work documented within this dissertation it contributes to the knowledge of condition monitoring in sheet-metal stamping processes by providing a novel data-driven framework to extract, preprocess, and analyze acoustic emission data for punch wear monitoring in fineblanking. Furthermore, a set of important signal features for punch wear prediction is derived from extensive data collection and analysis.

# Einleitung

**Inhalt** Diese Dissertation widmet sich der indirekten Zustandsüberwachung von Stanzwerkzeugen anhand von Körperschallsignalen. Das Hauptziel ist die Entwicklung eines neuartigen, robusten Frameworks für die Extraktion, Vorverarbeitung und Analyse dieser Signale während des Feinschneidens, einem Stanzverfahren. Integraler Bestandteil dieses Frameworks sind Data-Mining- und maschinelle Lernalgorithmen, die für die effiziente und effektive Analyse komplexer und hochdimensionaler Daten unerlässlich sind. Diese Werkzeuge sind entscheidend für die Gewinnung präziser, prädiktiver Erkenntnisse, die Verbesserungen im Fertigungsbetrieb ermöglichen. Angesichts strenger Nachhaltigkeitsziele und steigender Energie- und Materialkosten ebnet diese Forschung den Weg zu einer fortschrittlichen Verschleißüberwachung und damit zu einer optimierten Ressourcennutzung in Stanzprozessen, die sowohl wirtschaftlichen als auch ökologischen Zielen dient.

**Motivation** Die EUROPÄISCHE UNION beabsichtigt, die Treibhausgasemissionen bis zum Jahr 2030 auf mindestens 55 % unter das Niveau von 1990 zu senken [EURO21a] und bis 2050 klimaneutral zu sein [EURO21b]. Der Leichtbau im Automobilbereich ist ein wichtiger Baustein zur Reduzierung der $CO_2$-Emissionen durch Gewichtsreduzierung, z.B. durch den zunehmenden Einsatz von höherfesten Stählen mit einem hohen Verhältnis von Festigkeit zu Gewicht in Blechumformungsverfahren wie dem Stanzen, das eines der gängigsten Fertigungsverfahren für Fahrzeugkomponenten darstellt [KLOC16]. Der Übergang zu Elektrofahrzeugen verstärkt diese Entwicklung noch aufgrund des zusätzlichen Gewichts der Batterien [JOOS12].

Der zunehmende Einsatz von höherfesten Stählen wiederum führt zu höheren Prozesskräften und Spannungen bei Stanzwerkzeugen [SHAN19], die zu übermäßigem Verschleiß von kritischen Werkzeugkomponenten wie dem Stempel [VOIG21] führen. Übermäßiger Verschleiß beeinträchtigt die Wirtschaftlichkeit von Stanzprozessen wie dem Feinschneiden, da er zu Ausschussproduktion und ungeplanten Stillstandszeiten führt [UNTE21]. Bei ohnehin knappen Margen beim Stanzen [KUBI22] hat die jüngste Entwicklung auf den Energie- [STAT23a] und Stahlmärkten [STAT23b], insbesondere in Deutschland, den Kostendruck auf die Unternehmen durch gestiegene Betriebs- und Materialkosten weiter verschärft.

Um die Produktion von Ausschuss zu vermeiden und ungeplante Stillstandszeiten zu reduzieren und damit die Ressourcenausnutzung bei Stanzprozessen zu verbessern, muss der Werkzeugverschleiß durch rechtzeitige Wartung kontrolliert werden. Ein hinreichend feingranulares Verständnis der Schwere des Verschleißes ist für diese Kontrolle unerlässlich. Aufgrund der geschlossenen Werkzeugkonstruktion beim Stanzen ist der Verschleiß kritischer Werkzeugkomponenten jedoch nicht direkt während des Prozesses beobachtbar.

**Forschungslücke** Frühere Ansätze stützten sich häufig auf wissensbasierte Methoden wie Finite-Elemente-Simulationen, um Stanzprozesse auszulegen und um geeignete Wartungsintervalle im Voraus abzuschätzen. Diese Methoden sind zwar weiterhin von entscheidender Bedeutung, werden durch stark vereinfachende Annahmen den variablen Bedingungen in der industriellen Umgebung allerdings nicht gerecht [LIEW22]. Daher wurden in letzter Zeit von Industrie und Forschung zahlreiche Anstrengungen unternommen, um geeignete Prozesssignale während des Stanzens zu überwachen und daraus mit Hilfe traditioneller Signalverarbeitungsmethoden und Algorithmen des maschinellen Lernens den Verschleißzustand (Zustandsüberwachung) und in Folge dieses optimale Wartungsintervalle (vorausschauende Wartung) abzuleiten.

Bis heute stellt die Zustandsüberwachung (und damit die vorausschauende Wartung) beim Stanzen allerdings ein offenes Problem dar [UNTE23]. Die Gründe dafür sind vielfältig. MCKINSEY & COMPANY hat die wichtigsten Hindernisse herausgearbeitet, die erfolgreichen

Projekten zur vorausschauenden Wartung im Allgemeinen im Wege stehen [McKi21]. Die Hindernisse 'Unzureichende Daten', 'Mangel an Fähigkeiten' und 'Geringe Wirtschaftlichkeit' (siehe Abbildung 1.2) werden in dieser Arbeit wie folgt adressiert.

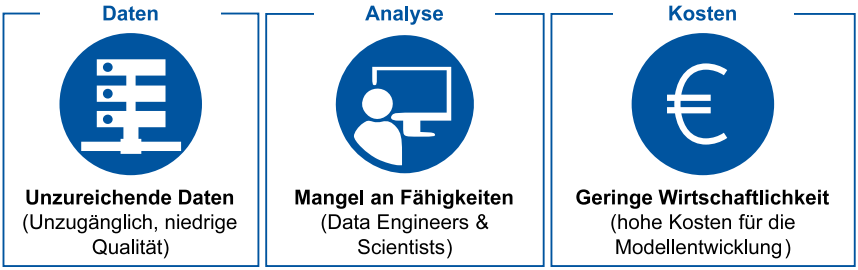



**Abbildung 1.2:** Hindernisse für erfolgreiche Projekte im Bereich der vorrausschauenden Wartung ([McKi21]), die in dieser Dissertation behandelt werden.

**Forschungsansatz** In dieser Dissertation werden Konzepte für die Datenvorverarbeitung und Datenanalyse entwickelt, um den Verschleißzustand des Stempels während des Feinschneidens in einer Laborumgebung mit einer industriellen Feinschneidanlage und in einer industriellen Feinschneidumgebung mit Körperschallsignalen zu überwachen. Körperschallsignale eignen sich aufgrund ihrer geringen Kosten und der einfachen Installation besonders gut für industrielle Umgebungen. Im Gegensatz zu Kraftsensoren sind die Signale jedoch aufgrund der Komplexität der resultierenden Signale schwieriger zu interpretieren [Koll18] und vorzuverarbeiten.

Diese Herausforderungen werden durch den Einsatz und die Kombination geeigneter Algorithmen adressiert. Es wurde eine Pipeline entwickelt, die Data-Mining-Techniken einsetzt, um systematisch Signalsegmente zu extrahieren, die Informationen über den Stempelverschleiß enthalten, und anschließend zugrundeliegende Muster in den Daten aufzudecken, um das Problem 'Unzureichende Daten' zu lösen. Darüber hinaus wird ein Proxy für den Stempelverschleiß vorgeschlagen, in der Laborumgebung beobachtet und bewertet und zum Labeling eines umfassenden Datensatzes verwendet. Dieser Datensatz dient zum Training und zur Evaluierung überwachter maschineller Lernmodelle, wobei sowohl die Modellgüte als auch Aspekte der erklärbaren KI (Explainable AI) untersucht werden. Die Ergebnisse der Analysen werden auf die industriellen Datensätze angewandt, indem ein skalarer Indikator für die Überwachung des Werkzeugzustands entwickelt und seine Entwicklung über die Datensätze hinweg bewertet wird. Es zeigt sich, dass der Indikator Informationen zur Identifizierung von Hubserien enthält, die letztlich früher als erwartet eingestellt wurden, und damit entscheidende Einblicke in die zugrunde liegende Dynamik industrieller Feinschneidprozesse liefert. Diese Erkenntnisse erleichtern eine bessere Entscheidungsfindung und ermöglichen proaktive Wartungsstrategien, die kostspielige Ausfälle verhindern und die Produktionseffizienz optimieren können. Die in dieser Dissertation dokumentierte Arbeit befasst sich somit direkt mit den Herausforderungen 'Mangel an Fähigkeiten' und 'Geringer wirtschaftlicher Ertrag' bei vorausschauenden Wartungsprogrammen, indem sie datengesteuerte Modellierungsansätze und Pipelines vorschlägt und bewertet und so dazu beiträgt, die mit der Modellentwicklung verbundenen Kosten zu senken.

**Beitrag** Die in dieser Dissertation dokumentierte Arbeit leistet einen Beitrag zum Wissen über die Zustandsüberwachung in der Blechumformung, indem sie einen neuartigen datengesteuerten Rahmen zur Extraktion, Vorverarbeitung und Analyse von Schallemissionsdaten

für die Überwachung des Stempelverschleißes beim Feinschneiden bereitstellt. Darüber hinaus wird eine Reihe wichtiger Signalmerkmale für die Vorhersage des Stempelverschleißes aus einer umfangreichen Datenerfassung und -analyse abgeleitet.

# 2 State of the art for data-driven tool condition monitoring in sheet-metal stamping

*Stand der Technik zur datengetriebenen Zustandsüberwachung von Stanzwerkzeugen*

Chapter 2 lays out the basics of data-driven tool condition monitoring in sheet-metal stamping and presents a summary of the state of the art. This work resides at the intersection of machine learning – which itself is an amalgam of computer science and mathematics – and engineering. Therefore, important concepts from these disciplines are also introduced.

Introducing these concepts enables to subsequently analyze the state of the art for data-driven tool condition monitoring in sheet-metal stamping processes. Consequently, it paves the way to identify the research gaps this work intends to bridge.

The experiments in this study focus on fineblanking, a specialized form of sheet-metal stamping. Hence, Section 2.1 introduces the *technological foundations of fineblanking*. Section 2.2 presents *important concepts, definitions, and methods within machine learning*. Section 2.3 lays out the *foundations of time series analysis in a signal processing context*. To broaden the overview of the current state of the art, Section 2.4 presents the *state of the art of indirect tool condition monitoring in sheet-metal stamping processes*. Based on this, the *research deficit* is derived within Section 2.5.

## 2.1 Foundations of fineblanking

*Grundlagen des Feinschneidens*

Section 2.1 presents the *technological foundations of fineblanking*. Specifically, the *process principles of fineblanking* are laid out in Section 2.1.1. Subsequently, *wear mechanisms* in tribological systems and specifically *tool wear mechanisms in sheet-metal stamping processes* such as fineblanking are presented in Section 2.1.2. Finally, Section 2.1.3 provides an *introduction to the general theory and motivation of tool condition monitoring* in sheet-metal stamping processes.

### 2.1.1 Process principles of fineblanking

*Verfahrensprinzipien des Feinschneidens*

Fineblanking is a highly efficient sheet-metal forming technology for the mass production of identical functional metal components. It is a separation process under the group of shearing processes according to DIN 8588 [DEUT13]. Fineblanked parts are characterized by high geometrical accuracy and smooth sheared surfaces. Hence, unlike conventional sheet-metal stamping processes, no further edge or surface processing is required for fineblanked parts [KLOC14].

Fineblanking tools can be designed to combine a shearing operation with other forming operations, such as coining or bending (*progression die*) [SIEG15]. With production rates of up to 200 parts per minute, fineblanking is particularly relevant for the mass production of safety-relevant components for the automotive industry, such as brake pad carriers or seat belt tongues [DJAV10], and for the aerospace and medical industries [ZHEN19]. Figure 2.1 illustrates four exemplary safety-critical fineblanked components for an automobile with a

The content of Section 2.4 was largely published in [UNTE23]. CHATGPT (GPT-4, https://chatgpt.com/) was used in this chapter to proofread texts generated by the author and for support with LATEX formatting. DEEPL (https://www.deepl.com/translator/) was used for translation purposes.

demand for high geometric accuracy and mechanical strength. As a result, fineblanking is chosen over conventional sheet-metal stamping for applications such as the displayed.

In contrast to conventional sheet-metal stamping processes, fineblanking uses a counterforce (applied to the sheet metal by the *counter holder*) and a *blank holder* force with a V-ring to induce compressive stresses in the shearing zone. The *die clearance* (the length of the gap between punch and die) is usually chosen as approximately 0.5 % of the sheet metal thickness in fineblanking, which is 10 times lower than the usual die clearance in conventional sheet-metal stamping processes [KLOC14]. The high requirements for the design and production of fineblanking tools result in substantially higher costs for fineblanking tools compared to conventional sheet-metal stamping tools [LEE95].

A fineblanking plant generally consists of a *decoiler*, a *leveler*, and the actual *fineblanking press*, see Figure 2.2. After decoiling, the material is fed to the leveler to straighten it and reduce residual stresses. Subsequently, the material is fed into the press, where an appropriate lubricant is applied to the sheet metal, and the fineblanking operation is carried out. Finally, the manufactured part is ejected through mechanical or pneumatic means.

Figure 2.3 visualizes the fineblanking operation carried out by the tool (referenced to as a *stroke*), which can be separated into four consecutive process phases [KLOC14]. In the initial stage, (a) the sheet-metal material is fed to the press while the tool remains open. Followingly, (b) the V-ring is indented into the material with the blank holder force $F_{\mathrm{B}}$ to induce compressive stresses in the shearing zone and prevent lateral movement of the material.

The counter holder applies a counter force $F_{\mathrm{C}}$. After the V-ring indentation, (c) the die is shifted towards the punch, and a reaction force emerges as the punch has sheet metal contact, measured as the punch force $F_{\mathrm{P}}$. The material is sheared when the flow stress of the material is surpassed.

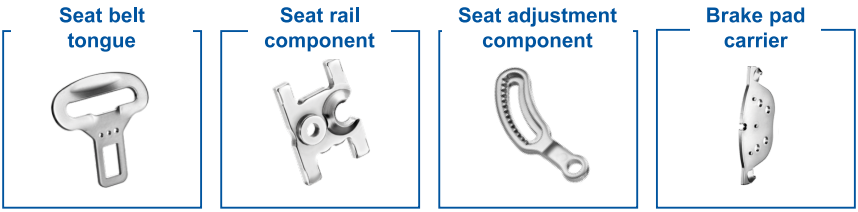



**Figure 2.1:** Exemplary fineblanked and safety-critical components in automobiles (FEINTOOL AG).
*Beispielhafte Feinschneid- und sicherheitskritische Bauteile in Automobilen.*

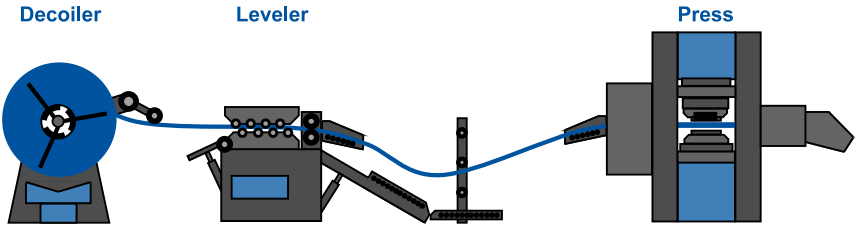
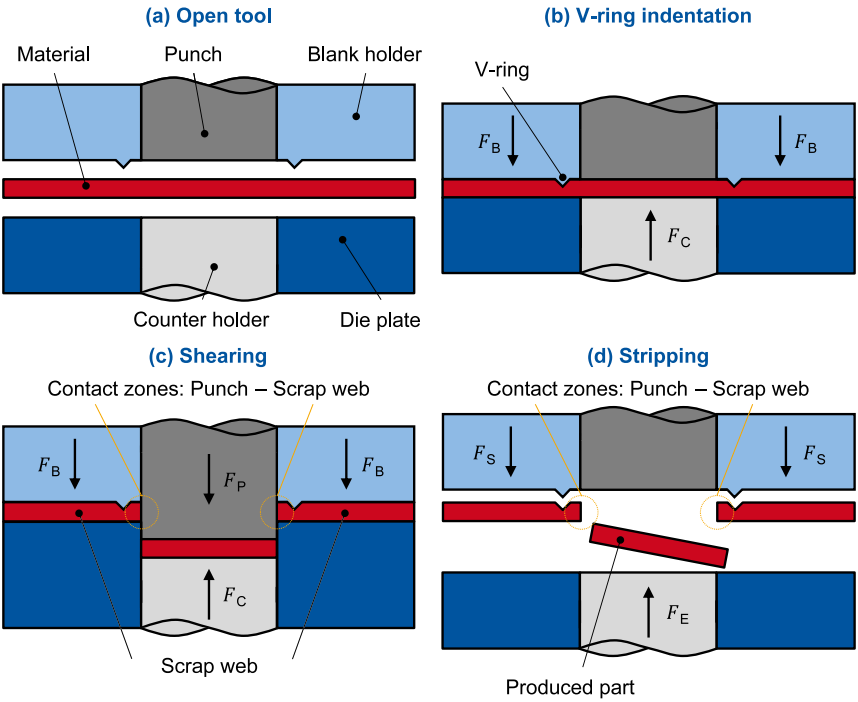



**Figure 2.2:** Schematics of a typical fineblanking plant.
*Schematische Darstellung einer typischen Feinschneidanlage.*




Legend: $F_B$: Blank holder force; $F_C$: Counter holder force; $F_P$: Punch force; $F_S$: Stripping force; $F_E$: Ejection force

**Figure 2.3:** Depiction of consecutive process phases in fineblanking ([UNTE23]).
*Darstellung der aufeinanderfolgenden Prozessphasen beim Feinschneiden.*

Over the whole process, the material is tightly clamped and controlled through the blank and counter holder forces. The blank holder and counter holder force remain approximately constant during the shearing phase, creating a hydrostatic stress state within the shearing zone. Due to the required punch force, the punch is subject to compressive stresses during the shearing phase [KLOC14]. After the shearing phase, (d) the stripping force $F_S$ is applied to the blank holder to overcome the friction between the punch and the resulting scrap part (the *scrap web*) is stripped off of the punch, and the tool opens.

During the stripping phase, the punch is subject to a combination of bending and tensile stresses since the clamping force between the punch and the scrap web needs to be surpassed [VOIG21]. Finally, the counter holder ejects the part into the tool room with the ejector force $F_E$. After ejection, the material is fed forward, and the scrap web (see Figure 2.4) is continuously dissected into *scrap web segments* for disposal.

The usage of higher-strength steels in sheet-metal stamping processes, such as fineblanking, increased within recent years due to the trend of lightweight design [ZHEN19]. This increased usage of higher-strength steels leads to increased process forces during mass production [SHAN18]. For fineblanking, these increased process forces, in turn, increase the punch load and accelerate the wear development of the punch, specifically during the shearing and

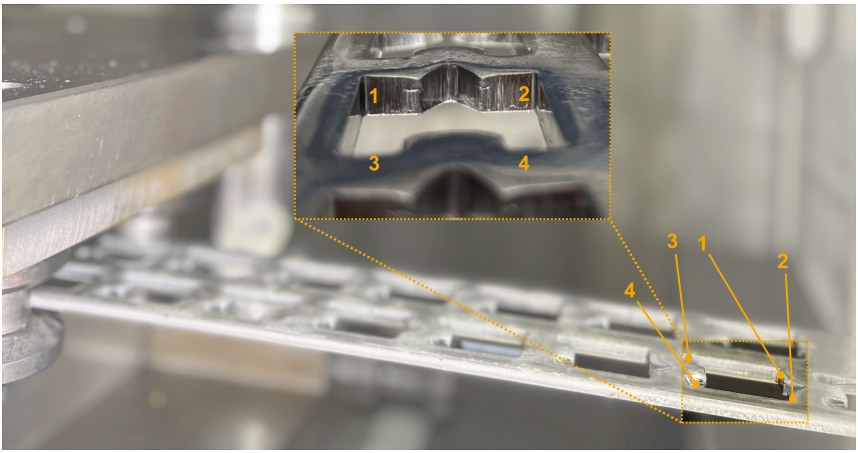



**Figure 2.4:** A resulting scrap web before dissection into scrap web segments and a close-up of selected sheared surfaces.
*Resultierendes Stanzgitter vor der Zerteilung und Detailansicht ausgewählter Schnittflächen.*

stripping phases, when the punch is subject to alternating compressive and tensile stresses [VOIG21]. The following section will lay out the known wear mechanisms and describe tool wear mechanisms in sheet-metal stamping.

### 2.1.2 Tool wear mechanisms in sheet-metal stamping processes

*Werkzeugverschleißmechanismen bei Stanzprozessen*

Wear is characterized as the gradual loss and smearing of material resulting from mechanical factors, such as contact and relative motion between solid, liquid, or gaseous mating bodies [CZIC10]. In manufacturing processes, wear of the tool (following: *tool wear*) is generally undesirable, as it diminishes the value and functionality of produced parts since tolerances can not be met and potentially leads to tool failure, which subsequently leads to the production of reject parts as well or causes machine standstills due to the need to maintain or change the tool.

According to the DEUTSCHE GESELLSCHAFT FÜR TRIBOLOGIE E.V., the four principal wear mechanisms are *adhesion*, *abrasion*, *tribochemical reaction* (tribo-oxidation), and *surface degradation* (fatigue) [GESE02]. Figure 2.5 illustrates these mechanisms, further elucidated in the subsequent paragraph.

Adhesion, as depicted in Figure 2.5 (a), involves the formation of local interfacial bonds, referred to as *cold welds*. Adhesive wear transpires due to the subsequent separation of bonded joints beyond the original interfaces, specifically in the adjacent material of the contact partner with lower strength. This leads to material transfer from the lower-strength material to the higher-strength material. Adhesive wear is one of the most common wear mechanisms and is characterized by "*high wear rates and early initiation*" [TROE17]. In contrast, as visualized in Figure 2.5 (b), abrasion arises from hard particles removing material through furrowing and scoring stresses. The third principal wear mechanism is the tribochemical reaction, illustrated in Figure 2.5 (c). This mechanism encompasses chemical

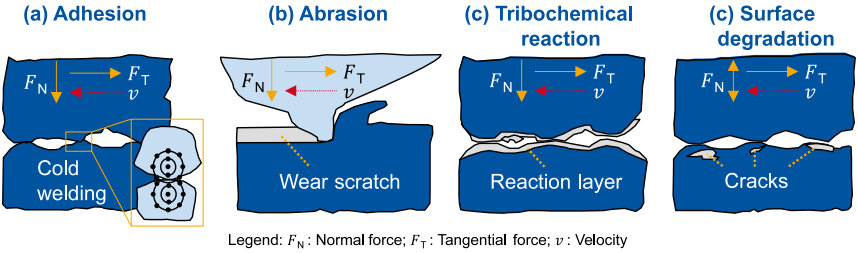



**Figure 2.5:** Visualization of fundamental wear mechanisms (adapted from [CZIC10]).
*Visualisierung der grundlegenden Verschleißmechanismen.*

reactions between the materials involved in the wear partners and their surroundings, leading to alterations in material characteristics. The chemically modified zones often exhibit different wear properties than the original material. Lastly, surface degradation, the fourth fundamental wear mechanism, is depicted in Figure 2.5 (d). Surface degradation occurs due to fatigue and cracking in near-surface regions, causing material separation and degradation. [CZIC10]

As production time progresses and the number of accumulated process cycles increases, the observable wear of two sliding bodies also increases. Traditionally, the accumulation of tool wear is modeled as non-linear with three different stages that differ in their rate of change (see Figure 2.6). Specifically:

(i) A *break-in phase*, which denotes an initial phase of rapid wear accumulation,

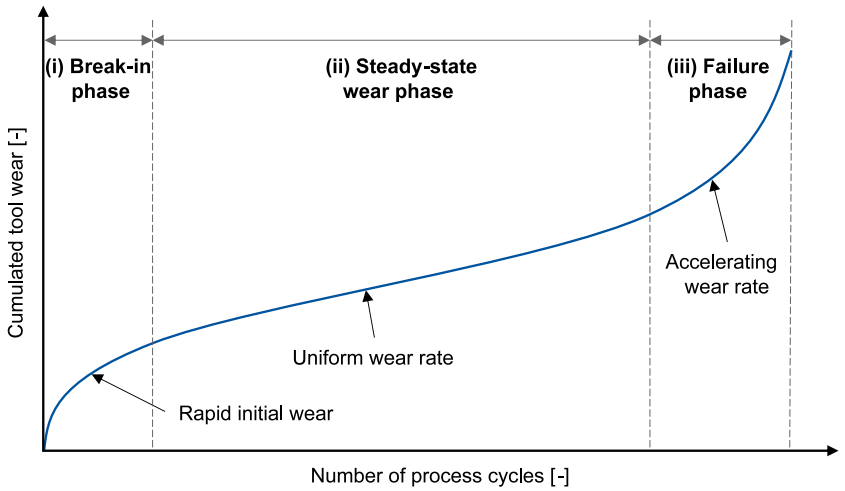



**Figure 2.6:** Schematic visualization of the traditional three-stage model of tool wear accumulation (adapted from [GROO10]).
*Schematische Visualisierung des traditionellen 3-Phasen Modells zur Verschleißprogression.*

(ii) a *steady-state wear phase*, which denotes the second phase with nearly constant wear accumulation, and finally

(iii) a *failure phase*, which denotes the last phase of the tool life cycle with an accelerated accumulation that ultimately results in tool failure.

Sheet-metal stamping processes, such as fineblanking, are considered stable processes once they are set up, i.e., no further adjustments to the parametrization (e.g., lubrication intervals, lubrication amount, forces, stroke frequency, etc.) have to be made to achieve target product quality. Still, both research and industry report fluctuations in product quality and tool life due to unpredictable wear accumulation, impeding the economic efficiency of the process. These fluctuations reveal a stochastic behavior that can not be accounted for with the traditional approach that mainly relies on experience and finite element simulations (FES) to estimate the process behavior prior to production time [LIEW22].

PEREIRA et al. found that the accumulation of wear is not evenly distributed over the die radius in sheet-metal stamping and emphasize the criticality of the *galling* wear mechanism for the overall tool wear response and tool lifetime [PERE12]. Galling denotes multi-stage surface damage in sheet-metal forming due to adhesion with an initial stage, where sheet-metal material is transferred to the tool surface. Due to subsequent sliding, the transfer layer continuously grows, leading to microscopic local scratching of the sheet-metal material, which may be attributed to strain hardening and oxidation [GAAR09].

VOIGTS stated that in fineblanking, the shearing edges of the punch and the sheared surface of the scrap web represent a tribological system during the *shearing* (see Figure 2.3 (a)) and *stripping* phase (see Figure 2.3 (b)) with the lubricant as an intermediate [VOIG21].

KLOCKE et al. argued that due to adhesion on the punch in fineblanking processes, the friction increases. Besides quality impairment of the produced part, the punch life is decreased through galling as well, since the stresses on the punch surface are altered, which in turn leads to tool fatigue near the surface and crack development and crack coalescence [KLOC01].

The cost for individual sheet-metal stamping tools within the automotive industry can range up to 1 Million USD, according to industry experts, depending on the part complexity. Hence, even small improvements to tool maintenance intervals (reducing unnecessary downtimes and determining when maintenance is necessary) can lead to a substantial cost reduction [SHAN17]. Since the tool wear is unobservable during production due to the closed tool design in sheet-metal stamping, a *direct tool condition monitoring* is not economically viable. It would require a machine standstill and subsequent disassembling of the tool to examine the wear state of the active surfaces. Thus, a substantial amount of research has been conducted to indirectly infer the wear state of the tool through the analysis of suitable process signals [NIEM22]. This approach is called *indirect tool condition monitoring* and is described and motivated in more detail in the next subsection.

### 2.1.3 The necessity of indirect tool condition monitoring in sheet-metal stamping

*Die Notwendigkeit der indirekten Zustandsüberwachung von Stanzwerkzeugen*

Ideally, the tool wear state would be known at every stroke in industrial sheet-metal stamping processes to infer optimal maintenance intervals. However, due to the closed design of sheet-metal stamping tools, the wear state of critical tool elements such as the die (i.e., the state of the die chamfer) and the punch (i.e., the state of the shearing edges of the punch)

can not be observed during production. Thus, industrial sheet-metal stamping processes typically rely on *preventive* and *reactive maintenance* approaches.

Preventive maintenance involves performing maintenance at predefined intervals based on experience, which can lead to inefficiencies due to early or late maintenance relative to the actual wear state [MARW12]. Reactive maintenance, or run-to-failure management, occurs when damage has already taken place, which can be economically inefficient and risky for other tool and machine components [MARW12]. As this approach puts other tool and machine components at risk (where the repair and substitution costs may surpass the replacement cost of wear-prone elements [MOBL02]), it is less common in industrial practice.

Since both preventive and reactive maintenance have inherent drawbacks, research, and industry have been trying to find ways to *predict* when maintenance is needed (*predictive maintenance* (PM)) in sheet-metal stamping for decades. Predictive maintenance (PM) can be viewed as a *condition-based preventive maintenance* strategy [MOBL02]. To this end, the tool condition needs to be monitored during production time (*tool condition monitoring* (TCM)). However, no TCM strategy has yet prevailed in industrial sheet-metal stamping processes that preserves the economic efficiency of the process and is reliable enough to make cost-intensive decisions based upon [NIEM22].

Tool condition monitoring can be done *directly* and *indirectly* [REHO05]. *Direct TCM* requires a direct assessment of the tool state, e.g., by disassembling the tool and subsequent visual inspection of the tool elements with a suitable procedure. Since machine standstills directly affect the already low margins per part [KUBI22], this approach is not viable for industrial production. *Indirect TCM* involves inferring the tool wear state from auxiliary quantities, i.e., observable phenomena that do not require stopping the process, such as forces or acoustic emissions (AE), so-called *process signals* [NIEM22]. Indirect TCM approaches are based on the assumption that wear state changes are detectable through changes in suitable process signals [UNTE21].

Process signals (and therefore observed changes within said process signals) in industrial reality are subject to *process noise* [LIEW22], which encompasses both hard-to-model physical effects due to incomplete information or complex interactions, as well as stochastic effects that can not be modeled with deterministic modeling approaches. Traditional deterministic modeling approaches (e.g., FES in forming technologies) make simplifying assumptions and do not allow modeling process noise [LIEW22]. In contrast, machine learning (ML) algorithms are suitable for learning from data containing stochastic effects because they can, conceptually, be understood as algorithms that "*search through a large space of candidate programs, guided by training experience, to find a program that optimizes the performance metric*" [JORD15].

The growth of available computational resources, combined with the development and accessibility of sophisticated ML frameworks in recent decades, has resulted in the widespread use of ML in PM and TCM research across various domains. This is visualized in Figure 2.7, which shows the number of publications (indexed by CLARIVATE WEB OF SCIENCE) with keywords *Machine learning AND (Predictive maintenance OR Condition monitoring)*).

The following subsection introduces fundamental ML definitions, concepts, and models needed to grasp the state of the art for TCM in sheet-metal stamping and the approach taken by this work.

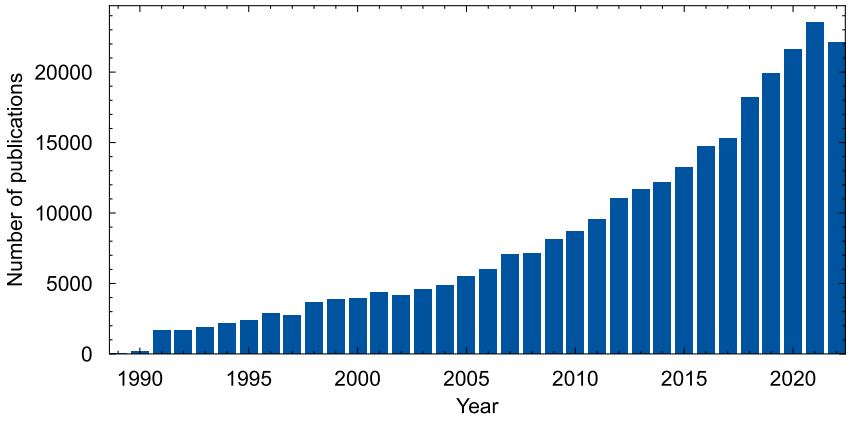



**Figure 2.7:** Number of publications over time featuring keywords 'Machine Learning' and either 'Predictive Maintenance' or 'Condition Monitoring'.
*Anzahl der Veröffentlichungen im Laufe der Zeit mit den Stichwörtern 'Machine Learning' und entweder 'Predictive Maintenance' oder 'Condition Monitoring'.*

## 2.2 Foundations of statistics and machine learning

*Grundlagen der Statistik und des maschinellen Lernens*

Section 2.2 introduces the *conceptual foundations of machine learning*. To this end, Section 2.2.1 presents *important concepts, definitions, and paradigms* within machine learning that are needed to grasp the current state of the art of indirect TCM in sheet-metal stamping processes. Section 2.2.2 introduces *selected machine learning algorithms and methods* that are used within this work.

### 2.2.1 Key concepts and paradigms

*Wichtige Konzepte und Paradigmen*

A commonly used definition of ML has been proposed by MITCHELL in 1997:

"*Machine Learning is the study of computer algorithms that improve automatically through experience. A computer program is said to learn from experience $E$ with respect to some class of tasks $T$ and performance measure $P$, if its performance at tasks in $T$, as measured by $P$, improves with experience $E$.*" [MITC97]

The term $T$ in this definition can stand for a broad spectrum of tasks, ranging from predicting the quality of manufactured parts to grouping data points in a reasonable way. To increase the experience $E$ of a computer program, the program is *trained* on specific data, executing task $T$ on this data to enhance its performance, measured by $P$. In the realm of machine learning, the data used to train and evaluate a computer program is known as a *dataset*.

From a probabilistic perspective, a dataset is a collection of realizations from underlying *random variables*. A random variable can be informally understood as a variable whose values depend on the outcomes of a random phenomenon. More formally, a real-valued

random variable $X$ is defined as a function that maps the outcomes of a process that depends on chance (e.g., unplanned machine standstills) to a real number. In practical scenarios, the properties of a random variable $X$ are mostly unknown and need to be *estimated* by collecting and analyzing datasets that contain *samples* from *realizations of random variables*. Interest often lies in specific properties of the distribution of the random variable, such as the expected value. Understanding these properties can enable informed predictions about future similar events and offer insights into the theoretical population from which the random variables are drawn.

A dataset consists of $n \in \mathbb{N}$ *observations*. An observation contains a set of $p \in \mathbb{N}$ *features*. These features together form a *predictor* $x_i$ with

$$x_i = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{ip} \end{pmatrix}, \tag{1}$$

where $i \in \{1, \ldots, n\}$ denotes the $i$-th observation. The $k$-th feature across all observations will be refered to as $\mathbf{x}_k$ with

$$\mathbf{x}_k = \begin{pmatrix} x_{1k} \\ x_{2k} \\ \vdots \\ x_{nk} \end{pmatrix} \tag{2}$$

Additionally to the predictor, an observation $i$ can contain a *label* $y_i \in \mathbb{R}^n$ (also called *response*). If, for every observation, both a predictor and a label are available, the dataset is represented as a set of pairs $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$.

However, datasets from industrial use cases seldom contain labels for each $i \in \{1, \ldots, n\}$. This is because the predictor is usually easily accessible (e.g., measurement data from a sensor during a manufacturing process), but the label is expensive to generate (e.g., resulting quality measurements for the manufacturing process) or represents the occurrence of rare events (e.g., machine anomalies).

All predictor observations of a measurement can be represented with an $n \times p$ matrix $\mathbf{X}$, often referenced to as the *model matrix* [HAST09] or *design matrix* [EVER10], with

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1p} \\ x_{21} & x_{22} & \ldots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \ldots & x_{np} \end{pmatrix}. \tag{3}$$

In this work, the number of features $p$ denotes the *dimensionality* of a dataset. Assuming the presence of labels for all observations, i.e., for each $i \in \{1, \ldots, n\}$ a label $y_i$ exists, the labels of the dataset can be represented with a vector

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}. \tag{4}$$

When features vary in their scales, possibly because they represent different physical quantities or domains, a scaling procedure is often an essential preprocessing step to ensure the numerical stability of numerous machine learning algorithms [KUHN13]. Among several scaling techniques, *standardization* (also known as *z-transform*) is a commonly applied method. This procedure rescales each feature $\mathbf{x}_k$, where $k \in \{1, 2, \ldots, p\}$, to zero mean and unit variance. This is accomplished by subtracting the *arithmetic mean* $\overline{\mathbf{x}}_k$ from feature $\mathbf{x}_k$, where

$$\overline{\mathbf{x}}_k = \frac{1}{n} \sum_{i=1}^{n} x_{ik}, \tag{5}$$

to ensure zero mean and subsequently dividing by the *sample standard deviation* $s_k$ of feature $\mathbf{x}_k$ with

$$s_k = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_{ik} - \overline{\mathbf{x}}_k)^2} \tag{6}$$

to ensure unit variance. This yields the standardized feature $\mathbf{z}_k = \frac{\mathbf{x}_k - \overline{\mathbf{x}}_k}{s_k}$ for each $k \in \{1, 2, \ldots, p\}$, effectuating a scaling transformation of the entire dataset.

Intuitively, the z-score

$$\mathbf{z}_k = \begin{pmatrix} z_{1k} \\ z_{2k} \\ \vdots \\ z_{nk} \end{pmatrix} \tag{7}$$

can be interpreted as the number of standard deviations by which each observation $i$ for the $k$-th feature deviates from the mean of that $k$-th feature. KUHN et al. state that standardizing $\mathbf{X}$ has no downsides besides a loss of interpretability in terms of units since the original units are not preserved [KUHN13]. Standardization will be used as a preprocessing technique within this work.

**Selected correlation measures** *Ausgewählte Korrelationsmaße*

Correlation measures seek to quantify relationships (*dependences*) between random variables. *The Pearson correlation coefficient* (PCC) represents the most well-known (linear) correlation measure. The PCC is defined for two real-valued random variables $X^{(1)}$ and $X^{(2)}$ and given by $\rho(X^{(1)}, X^{(2)})$ with

$$\rho(X^{(1)}, X^{(2)}) = \frac{\mathbb{E}\left[(X^{(1)} - \mu_{X^{(1)}})(X^{(2)} - \mu_{X^{(2)}})\right]}{\sigma_{X^{(1)}} \sigma_{X^{(2)}}}, \tag{8}$$

where $\mathbb{E}[\ldots]$ denotes the *expectation operator*, $\mu_{X^{(i)}} = \mathbb{E}[X^{(i)}]$ the *expected value* or *population mean* of $X^{(i)}$, and $\sigma_{X^{(i)}} = \sqrt{\mathbb{E}[(X^{(i)} - \mu_{X^{(i)}})^2]}$ the *population standard deviation* of $X^{(i)}$. The expected value and the population standard deviation of $X^{(i)}$ remain unknown in most practical settings. Hence, it is reasonable to extend the concept of the PCC to realizations of random variables for inferring information about the relationships between these variables from realizations.

Let $\mathbf{x}^{(1)} \in \mathbb{R}^n$ and $\mathbf{x}^{(2)} \in \mathbb{R}^n$ denote realizations of $X^{(1)}$ and $X^{(2)}$ respectively. The *sample PCC* of $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$ is then denoted as $r(\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ and defined as

$$r(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = \frac{\sum_{i=1}^{n}(x_i^{(1)} - \bar{\mathbf{x}}^{(1)})(x_i^{(2)} - \bar{\mathbf{x}}^{(2)})}{\sqrt{\sum_{i=1}^{n}(x_i^{(1)} - \bar{\mathbf{x}}^{(1)})^2}\sqrt{\sum_{i=1}^{n}(x_i^{(2)} - \bar{\mathbf{x}}^{(2)})^2}}, \tag{9}$$

where $x_j^{(i)}$ denotes the $j$-th entry of the sample $\mathbf{x}^{(i)}$ and $\bar{\mathbf{x}}^{(i)}$ the arithmetic mean of $\mathbf{x}_i$.

The PCC has the following properties:

(i) If $r(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = 0$, there is *no linear relationship* between $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$.

(ii) If $r(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = 1$, there is a *perfect positive linear relationship* between $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$.

(iii) If $r(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = -1$, there is a *perfect negative linear relationship* between $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$.

While the PCC represents an accessible and established measure for correlation, it is limited to quantifying linear dependence between two random variables and realizations of random variables. In physical and biological systems, such as manufacturing systems and gene co-expression, non-linear relationships between random variables and, thus, their realizations are to be expected.

*Population distance correlation* (population DC) $\mathrm{dCor}(X^{(3)}, X^{(4)})$ is a statistical measure to quantify the degree of dependence between two *random vectors*, i.e., multivariate random variables $X^{(3)}$ that takes values in $\mathbb{R}^p$ and $X^{(4)}$ that takes values in $\mathbb{R}^q$ for $p, q \in \mathbb{N}$, introduced by SZÉKELY et al. in 2007 [SZEK07]. In contrast to the PCC, which only detects linear relationships between real-valued random variables, population DC can detect both linear and non-linear relationships and quantify their strength. In contrast to PCC, the nature of the dependency (positive linear and negative linear) can not be inferred from population DC.

The population DC $\mathrm{dCor}(X^{(3)}, X^{(4)})$ of two random vectors $X^{(3)}$ and $X^{(4)}$ is defined as

$$\mathrm{dCor}(X^{(3)}, X^{(4)}) = \frac{\mathrm{dCov}(X^{(3)}, X^{(4)})}{\sqrt{\mathrm{dVar}(X^{(3)})\mathrm{dVar}(X^{(4)})}}, \tag{10}$$

where $\mathrm{dCov}$ denotes the *population distance covariance* between the random vectors and $\mathrm{dVar}$ represents the *population distance variance*, which is the population distance covariance of a random vector with itself:

$$\mathrm{dVar}(X^{(i)}) = \mathrm{dCov}(X^{(i)}, X^{(i)}). \tag{11}$$

The structure of Eq. 10 bears a clear resemblance to that of Eq. 8. In Eq. 8, the numerator contains the population covariance of $X^{(1)}$ and $X^{(2)}$, and the denominator contains the product of their population standard deviations. Similarly, in Eq. 10, the numerator contains the population distance covariance between $X^{(3)}$ and $X^{(4)}$, and the denominator is effectively the geometric mean of their population distance variances, serving as an analog to population standard deviations in the context of Eq. 8.

Random variables are uniquely characterized by their Fourier transform, known as the *characteristic function*. The characteristic function $\phi_{X^{(1)},X^{(2)}}$ of the joint distribution of $X^{(1)}$ and $X^{(2)}$ is equal to the product of their individual characteristic functions $\phi_{X^{(1)}}\phi_{X^{(2)}}$ if and

only if $X^{(1)}$ and $X^{(2)}$ are stochastically independent. The idea behind population distance covariance is to measure the deviation from this independence:

$$\mathrm{dCov}^2(X^{(1)}, X^{(2)}; w) = \int_{\mathbb{R}^{p+q}} \left|\phi_{X^{(1)},X^{(2)}}(t,s) - \phi_{X^{(1)}}(t)\phi_{X^{(2)}}(s)\right|^2 w(t,s)\, dt\, ds, \tag{12}$$

where $w(t,s)$ represents a weighting function. It is an intrinsic property of the measure that $\mathrm{dCov}^2(X^{(1)}, X^{(2)}; w) = 0$ if and only if $X^{(1)}$ and $X^{(2)}$ are stochastically independent.

In practice, the characteristic function is unknown for a given sample and thus must be estimated. It turns out that with a suitable weighting function $w(t,s)$, this estimation can be accomplished through straightforward calculations, leading to the *(sample) DC* and *sample distance covariance* [SZEK07].

The DC $\mathrm{dCor}_n(\mathbf{X}^{(3)}, \mathbf{X}^{(4)})$ of $n$ realizations of two random vectors, where $\mathbf{X}^{(3)} \in \mathbb{R}^{n\times p}$ and $\mathbf{X}^{(4)} \in \mathbb{R}^{n\times q}$, is defined analogously to Eq. 10:

$$\mathrm{dCor}_n(\mathbf{X}^{(3)}, \mathbf{X}^{(4)}) = \frac{\mathrm{dCov}_n(\mathbf{X}^{(3)}, \mathbf{X}^{(4)})}{\sqrt{\mathrm{dVar}_n(\mathbf{X}^{(3)})\mathrm{dVar}_n(\mathbf{X}^{(4)})}}, \tag{13}$$

where $\mathrm{dCov}$ denotes the *sample distance covariance* between the random vectors and $\mathrm{dVar}_n$ the *sample distance variance*. The sample distance covariance is calculated by initially calculating the pairwise distance matrices

$$\begin{aligned} a_{j,k} &= \|x_j^{(3)} - x_k^{(3)}\|,\ j,k \in \{1,2,\ldots,n\} \\ b_{j,k} &= \|x_j^{(4)} - x_k^{(4)}\|,\ j,k \in \{1,2,\ldots,n\}, \end{aligned} \tag{14}$$

with $\|\cdot\|$ representing the Euclidean norm. Following this, the centered distance matrices are calculated as:

$$\begin{aligned} A_{j,k} &= a_{j,k} - \bar{a}_{j,\cdot} - \bar{a}_{\cdot,k} + \bar{a}_{\cdot\cdot} \\ B_{j,k} &= b_{j,k} - \bar{b}_{j,\cdot} - \bar{b}_{\cdot,k} + \bar{b}_{\cdot\cdot}, \end{aligned} \tag{15}$$

where $\bar{a}_{j,\cdot}$ and $\bar{b}_{j,\cdot}$ denote the mean over $k$ for a fixed $j$, $\bar{a}_{\cdot,k}$ and $\bar{b}_{\cdot,k}$ denote the mean over $j$ for a fixed $k$, and $\bar{a}_{\cdot\cdot}$ and $\bar{b}_{\cdot\cdot}$ represent the grand means of $A_{j,k}$ and $B_{j,k}$, respectively. Then, the sample distance covariance can be calculated as

$$\mathrm{dCov}_n(\mathbf{X}^{(3)}, \mathbf{X}^{(4)}) = \sqrt{\frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} A_{j,k}\, B_{j,k}}. \tag{16}$$

The sample distance variance is defined as the distance covariance of a realization with itself, e.g. for $X^{(3)}$:

$$\mathrm{dVar}_n(\mathbf{X}^{(3)}) = \mathrm{dCov}_n(\mathbf{X}^{(3)}, \mathbf{X}^{(3)}) = \sqrt{\frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} A_{j,k}\, A_{j,k}}. \tag{17}$$

The centered distance matrices $A_{j,k}$ and $B_{j,k}$ capture the relative distances between observations, removing the effect of individual observation's general proximity or distance to all other observations. This centering is crucial for ensuring the independence of the distance covariance from the mean distances.

Population DC and DC are normalized measures, i.e.,

$$\begin{aligned} 0 \leq \mathrm{dCor}(X^{(3)}, X^{(4)}) \leq 1, \\ 0 \leq \mathrm{dCor}_n(\mathbf{X}^{(3)}, \mathbf{X}^{(4)}) \leq 1. \end{aligned} \tag{18}$$

Population DC fulfills the desirable properties that

(i) $\mathrm{dCor}(\mathbf{X}^{(3)}, \mathbf{X}^{(4)}) = 0$ if and only if $\mathbf{X}^{(3)}$ and $\mathbf{X}^{(4)}$ are stochastically independent.

(ii) If $\mathrm{dCor}(\mathbf{X}^{(3)}, \mathbf{X}^{(4)}) = 1$ , then a vector $\mathbf{a}$, a scalar $b \in \mathbb{R}$, and an orthogonal matrix $\mathbf{D}$ such that $\mathbf{X}^{(4)} = \mathbf{a} + b\mathbf{X}^{(3)}\mathbf{C}$ exist *almost surely* [SZEK07].

In a probabilistic context, the term almost surely refers to the probability of an event being 1. However, counterintuitively and contrary to deterministic certainty, this does not guarantee the occurrence of the event since uncountably many different outcomes may occur.

SZÉKELY et al. used DC to test for independence in a simulation study. The power of the test was found to perform comparably to more specialized tests when identifying deviations in the case of a multivariate normal distribution [SZEK07]. Unlike specialized tests that are fine-tuned to specific forms of dependencies, DC can identify a broader range of deviations, making it a highly versatile and promising tool, specifically for data with unclear distribution, such as the data in this work.

In Chapter 8, DC is used to

(i) identify features within laboratory generated and labeled datasets with a *strong association with the tool condition*,

(ii) leverage the derived information to *develop a scalar indicator for the tool condition* to exploratively analyze unlabelled datasets.

To this end, this work utilizes the DC implementation from the PYTHON package *dcor* [SÁN20].

**Machine Learning paradigms** *Paradigmen im Maschinellen Lernen*

While ML algorithms can be distinguished arbitrarily fine granular and grouped by their utilized methodology/architecture, a distinction between *learning paradigms* has proven useful. These learning paradigms can be distinguished into *supervised learning*, *unsupervised learning*, and *reinforcement learning*. Reinforcement learning is not utilized in this work and is therefore not included in the following description.

A widely spread use case for ML models across different domains is to *predict* a label based on observed features. A problem, as well as the models developed to solve it, is classified as *supervised* if two conditions are met: i) all observations in the dataset contain a label and ii) these labels are used for *fitting* or *training* a model, i.e., model the relationship of $\mathbf{X}$ and $\mathbf{y}$ or subsets of them by optimizing a set of *model parameters*.

Let an unknown but fixed function $f$ describe the underlying relationship between $\mathbf{X}$ and $\mathbf{y}$, i.e.

$$\mathbf{y} = f(\mathbf{X}) + \varepsilon, \tag{19}$$

where $\varepsilon$ denotes an error term which is considered a realization from a random element that is stochastically independent of $\mathbf{X}$, i.e., $\mathrm{dCor}(\mathbf{X}, \varepsilon) = 0$. Let $f_\theta$ be a family of functions with $\theta$ representing different model parameters. Ideally, $f_\theta$ contains a function $f_{\theta_0}$ with

$$f_{\theta_0} = f. \tag{20}$$

In practice, $\theta_0$ may not be directly identifiable or might not even exist within the model family $f_\theta$. Therefore, the aim of supervised learning can be seen as finding a $\hat{\theta}$ that minimizes the discrepancy between $f_{\hat{\theta}}$ and $f$, typically measured using some form of *loss function*

$$L(\theta) \geq 0, \tag{21}$$

where $L(\theta) = 0$ corresponds to a perfect fit of the data. Supervised learning can be seen as the process of *estimating* the parametrization $\theta_0$ based on the dataset. This involves finding a set of model parameters $\hat{\theta}$ such that the resulting function $f_{\hat{\theta}}$ closely approximates the unknown function $f$ [JAME13]. This approximation can be expressed as:

$$\hat{f} = f_{\hat{\theta}}, \tag{22}$$

where $\hat{f}$ represents the estimate of $f$. When applied to the input data $\mathbf{X}$, this function predicts values $\hat{\mathbf{y}}$:

$$\hat{f}(\mathbf{X}) = \hat{\mathbf{y}}. \tag{23}$$

The process of supervised learning encompasses various techniques and algorithms designed to optimize this loss function, hence refining the estimation of $\hat{\theta}$ and, consequently, the predictions $\hat{\mathbf{y}}$ [BISH06].

Supervised machine learning models can be further distinguished by the type of labels present in the dataset. If the labels are represented by *quantitative variables* that take on continuous values, a supervised modeling task is referred to as a *regression* task. Popular and frequently used algorithms used for regression tasks are *linear regression* (LR) [WEIS05], *artificial neural networks* (ANNs) [BISH06], and *gradient-boosted trees* (GBTs) such as *eXtreme Gradient Boosting* XGBoost [CHEN16] or LightGBM [KE17].

Linear regression models are examples of *restrictive models* since they are limited to learning linear relationships between predictors and labels. They are, however, easily interpretable, and the influence of the predictors on the model predictions can be inferred without great effort. In contrast, ANN and GBT models are *flexible models* in the sense that they can learn arbitrarily complex and non-linear relationships between a predictor and a label but are generally harder to interpret. JAMES et al. state that *"in general, as the flexibility of a method increases, its interpretability decreases"* [JAME13]. Given a suitable metric for tool wear in sheet-metal stamping, an example of a regression task would involve finding a suitable function $\hat{f}$ that models the relationship between force signatures measured during a stroke and the corresponding tool wear metric.

Suppose the labels in the dataset are represented by *qualitative variables*, i.e., variables that take on a finite set of discrete, non-metric values or concepts (e.g., colors, brands, etc.). In that case, a supervised modeling task is called a *classification* task [BISH06]. Examples of widely used algorithms for classification tasks are *logistic regression* [HAST09], ANNs, and GBTs. Thus, certain algorithms, such as ANNs and GBTs, can be used both for regression and classification tasks (with slight architectural adaptions). Examples of classification tasks in a manufacturing context are reject and failure detection [CARV19].

For completeness purposes, it should be noted that *semi-supervised* machine learning models exist, where labels are available and used for a subset of the observations. However, as this work does not employ semi-supervised learning models, further discussion is omitted.

A problem, along with the models developed to address it, is referred to as *unsupervised* when the training process does not utilize labels, either due to their unavailability or a deliberate choice to exclude them. Unsupervised learning aims to uncover the underlying structure of relationships within the observations and to analyze these relationships with regard to different objectives. One objective could be to find *clusters* within the observations, i.e., to group observations into clusters of high similarity. Examples for popular clustering algorithms are *k-means* [HART75], *hierarchical clustering* [JOHN67], and *hierarchical density-based spatial clustering* [MCIN17].

Another objective of unsupervised learning models could be *density estimation*, i.e., estimating the underlying (and unobservable) probability density function based on the observations [BISH06], e.g., with *gaussian mixture models* [REYN09] or *kernel density estimation* [BISH06].

Figure 2.8 visualizes the basic concepts of supervised and unsupervised learning. Specifically, it shows a supervised binary classification and an unsupervised clustering using the same set of features. In the supervised case, the known true labels are used to learn a decision boundary, which is used to classify an observation based on its features to one of the two classes represented by the two labels in the data. In the unsupervised case, only the feature data is used to learn a clustering of the data. The data in Figure 2.8 only has two features (its dimensionality is 2). Generally, data from an industrial context, especially sensor data acquired with a high sampling frequency, has substantially more than two features.

**Explainable AI** *Erklärbare KI*

Past ML research has been heavily focused on further enhancing *model performance* for various performance metrics and benchmarking datasets, e.g., through the development of

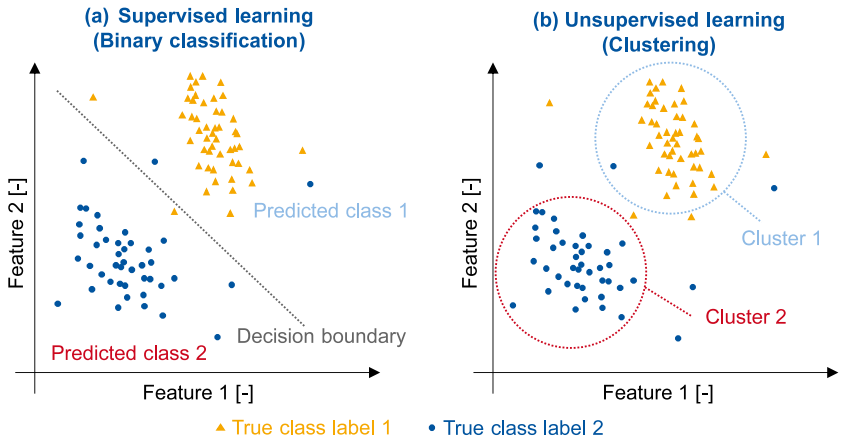



**Figure 2.8:** Schematic visualization of a (a) supervised (binary classification) and an (b) unsupervised (clustering) learning approach.
*Schematische Visualisierung eines (a) überwachten (binäre Klassifizierung) und eines (b) unüberwachten (Clustering) Lernansatzes.*

new state-of-the-art architectures or innovative approaches to optimizing *hyperparameters*, i.e., configuration variables that control the training process and architecture of a model. For real-world applications, where trust is required for acceptance, *model explainability*, i.e., understanding the rationale for model decisions, is of particular importance [BARR20]. Recent research coined the term *eXplainable Artificial Intelligence* (XAI) to systemize approaches towards model explainability and interpretability. DOŠILOVIĆ et al. furthermore point out that, although used interchangeably in a large portion of literature, interpretability and explainability refer to different concepts. While interpretability describes a mapping of an abstract concept into a domain that humans can make sense of (e.g., inferring specific rules from model decisions), explainability in the context of ML refers to identifying a collection of features from an interpretable domain that has contributed substantially to a model decision [DOŠ18]. Model explainability allows for the verification of model decisions in a physical context. If features unexpectedly rank high or exhibit correlations contrary to common knowledge, e.g., physical laws, the model may have learned spurious relationships or is susceptible to biases [RUDI19].

Research distinguishes between *inherently explainable* (*transparent* models such as Lasso or Ridge regression) and *black-box models*, such as ANNs or tree-ensembles, that require the usage of post-hoc methods after the training process to explain model predictions [BARR20]. Generally, model performance and transparency are often conflicting objectives and must be traded off [DOŠ18]. Thus, using post-hoc methods to explain model decisions while retaining high predictive capabilities has become an established approach both in industry and research.

Evaluating *feature importance* is an approach towards model explainability that seeks to quantify the contribution of each input feature to the model's predictions [BREI01]. By identifying and ranking the most influential features, one can gain insights into the key factors driving the model's decision-making process, potentially revealing patterns or relationships that were previously hidden or overlooked [FRIE01]. Understanding feature importance can help guide further analysis, refine the model, and provide actionable insights for practical applications in the industrial context [LOUP13]. Building on this foundation, the following exposition explores a powerful and versatile metric for feature importance analysis, offering a robust framework for model explainability.

**SHAP values** $SHAP\ Werte$

Shapely values, originating from cooperative game theory as the optimal payout for each player in a cooperative game, have been adapted to the machine learning context as a powerful tool for explaining individual model predictions within the concept of *SHAP* (SHapeley Additive exPlanations) values [LUND17]. Each feature of the input data can be regarded as a player (in a game-theoretical sense) that contributes to the overall prediction made by the model. SHAP values quantify the contribution of each feature towards the prediction for a specific instance, considering all possible feature combinations and their marginal contributions. The higher the SHAP value for a specific feature, the greater its contribution to the model prediction. SHAP values have been employed across various healthcare, finance, and manufacturing applications to aid decision-making and model transparency [MOLN20].

The exact computation of SHAP values is computationally intensive due to the vast number of possible feature subsets. LUNDBERG proposed *KernelSHAP* as an approximation method for SHAP values. KernelSHAP samples feature coalitions (combinations of features) and uses a weighted linear regression to estimate the SHAP values. The computational complexity of this method is $O(tm2^m)$, where $t$ is the number of samples and $m$ is the number of features. Later, *TreeSHAP* was introduced as a model-specific method to efficiently approximate SHAP

values for tree-based models, reducing the computational complexity to $O(tld^2)$, where $d$ is the maximum depth of any tree and $l$ is the number of leaves. TreeSHAP leverages the hierarchical structure of tree-based models to achieve this substantial reduction in computational complexity [LUND20].

In this work, SHAP values are used to globally assess the contributions of features within an interpretable feature set to XGBoost model predictions in Chapter 7. The TreeSHAP implementation from the PYTHON library *shap* was employed for this purpose [LUND20].

**The curse of dimensionality and dimensionality reduction** *Der Fluch der Dimensionalität und Dimensionalitätsreduktion*

As the number of dimensions in a dataset increases, the distance of a datapoint "*to the nearest neighbor approaches the distance to the farthest neighbor*", i.e., the *contrast* in distances vanishes [BEYE99]. This dilemma was termed the *curse of dimensionality*. The curse of dimensionality poses challenges for machine learning algorithms to discern patterns and relationships among features. To mitigate the curse of dimensionality, *dimensionality reduction* techniques can be utilized to preprocess high-dimensional data.

Dimensionality reduction algorithms essentially project high-dimensional observations $\mathbf{X} \in \mathbb{R}^{n \times p}$ to a lower-dimensional space $\mathbb{R}^{n \times k}$ with $k < p$. Dimensionality reduction is based on the hypothesis that not all features of a sample $x_i$ carry an equal amount of information, i.e., the $p$-dimensional data does not necessarily occupy the entirety of a $p$-dimensional space. A significant portion of the information might be concentrated in a lower-dimensional space, while the remaining dimensions might be information-sparse but not necessarily devoid of all information. Due to the curse of dimensionality, it is often expedient to focus on this lower-dimensional representation, accepting a marginal loss of information as a trade-off. This notion aligns with the *manifold hypothesis*, suggesting that data lies on a lower-dimensional *manifold* [GHOJ23]. The resulting lower-dimensional projections $\check{\mathbf{X}} \in \mathbb{R}^{n \times k}$ are called *embeddings* and the lower-dimensional space *embedding space*.

While projecting to a lower-dimensional space can potentially omit some information, the key focus of these techniques is to retain the most salient and relevant aspects of the data. In practice, dimensionality reduction is an invaluable step in data analysis, fine-tuned to ensure optimal retention of crucial information for the specific task at hand. Thus, a wide set of dimensionality reduction methods was developed in the past decades, encompassing two primary approaches: *feature extraction* and *feature selection* [JIA22].

Feature selection describes the process of selecting a subset of $k$ features with $k < p$ and dropping the remaining $p - k$ features. Common feature selection methods are *filter methods* (e.g., correlation analysis [HALL99]), *wrapper methods* (e.g., backward elimination [KOHA97]), and *embedded methods* (e.g., *Least absolute shrinkage and selection operator* (Lasso) [TIBS96]). Notably, Lasso is a regression technique that incorporates embedded feature selection through its $L_1$ regularization and will be discussed in Section 2.2.2.

Feature extraction describes the process of generating a new set of $k$ features with $k < p$ through a transformation of the original features within the dataset. This can be done with domain knowledge, e.g., by using domain-specific engineered features or dimensionality reduction algorithms that make implicit assumptions about the manifold, specifically, whether it is linear or non-linear, see Figure 2.9 [GHOJ23].

Widely used algorithms for dimensionality reduction are *principal component analysis* (PCA [HOTE33], supposing a linear manifold), *t-distributed stochastic neighbor embeddings* (t-SNE [VAN 08], supposing a non-linear manifold), *uniform manifold approximation and projection* (UMAP [MCIN18], supposing a non-linear manifold), and *autoencoders* (AECs [BISH06], typically supposing a non-linear manifold).

Both feature extraction and feature selection are employed in a manufacturing and indirect TCM context to mitigate the curse of dimensionality, especially for high-resolution sensor signals, as well as visualization tools to identify patterns in the data [LIEW22].

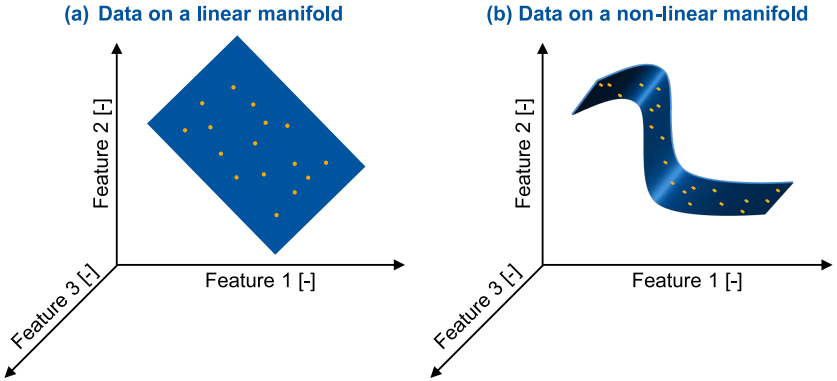



**Figure 2.9:** Schematic visualization of data lying on (a) a linear vs. (b) a non-linear manifold (adapted from [GHOJ23]).
*Schematische Visualisierung von Daten, die auf (a) einer linearen und (b) einer nichtlinearen Mannigfaltigkeit liegen.*

### 2.2.2 Introduction to selected algorithms and methods
*Einführung ausgewählter Algorithmen und Methoden*

Following, the regression methods Lasso and XGBoost are discussed. Moreover, the dimensionality reduction approaches PCA, UMAP, and AECs are explored.

**Lasso regression** *Lasso-Regression*

A widely adopted supervised learning technique is LR which models the linear relationship between predictor variables and a response variable. The relationship can be represented as:

$$f(\mathbf{X}) = \beta_0 + \mathbf{X}\beta_1. \tag{24}$$

The parameters $\beta_0$ (intercept) and $\beta_1$ (slope) that represent the best fit for the observed data are typically estimated by minimizing the loss function *residual sum of squares* (RSS):

$$\mathrm{RSS} = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \|\mathbf{y} - f(\mathbf{X})\|^2 = \|\mathbf{y} - \mathbf{X}\beta_1 - \beta_0\|^2, \tag{25}$$

where $y_i$ denotes the observed $i$-th label (i.e., the $i$-th entry of $\mathbf{y}$) and $\hat{y}_i$ denotes the prediction by the linear regression model based on the features for the $i$-th observation (i.e., the $i$-th row of $\mathbf{X}$).

Lasso regression extends the idea of fitting linear models by minimizing RSS by introducing *regularization* to control model complexity. Lasso is particularly effective for high-dimensional datasets with multicollinearity among predictor variables. Its distinct feature is the use of an $L_1$ penalty term, which combines the RSS term from LR with a penalty on the magnitude of coefficients. The regularization strength is controlled by a regularization parameter $\lambda > 0$, where higher values enforce stronger regularization.

A linear model is fitted using Lasso by minimizing an *objective function* consisting of a loss function and a regularization term. The loss function $L(\theta)$, with $\theta = (\beta_0, \ldots, \beta_p)$, is given by:

$$L(\theta) = \sum_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2. \tag{26}$$

The regularization term, represented by $\Omega_\lambda(\theta)$, is:

$$\Omega_\lambda(\theta) = \lambda \sum_{j=1}^{p} |\beta_j|. \tag{27}$$

Thus, the objective function for Lasso can be expressed as:

$$\mathrm{obj}(\theta) = \min_{\theta} \left( L(\theta) + \Omega_\lambda(\theta) \right), \tag{28}$$

The minimizing quantity of this function is denoted by $\hat{\beta}_\lambda^L$ and is called the *Lasso coefficients* [HAST09]. It is important to note that $\lambda$ is held constant during the minimization with respect to $\theta$. As such, the Lasso coefficients $\hat{\beta}_\lambda^L$ are dependent on the choice of $\lambda$. The $L_1$ penalty specifically targets the coefficient estimates of predictors, shrinking them towards zero. This uniquely forces some coefficients to be exactly zero, effectively performing feature selection as an embedded process. By doing so, Lasso offers a twofold benefit: reduced risk of *overfitting*, i.e., fitting the model too closely to training data, thereby compromising its generalization to new data, and enhanced explainability through the selection of non-zero coefficients.

In the predictive modeling conducted in Chapter 7, the *sklearn* [PEDR11] module *LassoCV* was used. The regularization parameter, $\lambda$, was optimized using *cross-validation*, a technique where the dataset is partitioned into $k$ subsets of observations, e.g., $k = 5$ or $k = 10$. The model is then trained iteratively on each of $k - 1$ of these subsets and validated on the remaining one of length $r$, where $r < n$, for a fixed $\lambda_i$. The *mean squared error* (MSE) of the validation subset is computed for each iteration with

$$\mathrm{MSE} = \frac{1}{r} \sum_{i=1}^{r} (y_i - \hat{y}_i)^2. \tag{29}$$

The optimal $\hat{\lambda}$ is selected based on the lowest average MSE across all iterations. This optimal $\hat{\lambda}$ is then used to fit a Lasso model to obtain the final model parameters $\hat{\theta}$. In the context of model evaluation, a widely accepted practice applicable beyond Lasso and general to supervised machine learning is to partition the dataset into a *training set* (e.g., 75 % of the data) and a *test set* (e.g., the remaining 25 %). The Lasso model is then fit to the training set and evaluated on the test set, using the MSE as a performance metric. For an in-depth discussion of cross-validation, see [JAME13].

**XGBoost regression** *XGBoost-Regression*

XGBoost is a GBT method with regularization, initially released in 2014 by CHEN as an open-source library with support for regression and binary classification problems. Gradient-boosted trees are *tree-based ensemble learning* algorithms that build an ensemble of decision trees to make predictions.

*Decision trees* are series of binary splits on the feature space that recursively partition the data into smaller subsets based on the values of the input features. Each tree leaf node represents a prediction for the response variable, with a corresponding *leaf score* that quantifies the prediction. The path from the root node to the leaf node determines the conditions for making that prediction.

In ensemble methods such as GBTs, these trees' individual predictions (or leaf scores) are aggregated, often by weighted summation, to produce a final prediction, enhancing both the model's robustness and accuracy.

*Boosting* is an approach to supervised machine learning problems that can be described as combining multiple weak learners (e.g., decision trees that perform just slightly better than random guessing) into a single strong learner (e.g., a decision tree ensemble that performs substantially better than random guessing) [FREU99]. Gradient Boosting extends this idea by utilizing gradient descent on an objective function to find a suitable combination of learners [CHEN16].

XGBoost gained attention in the annual *Data Mining and Knowledge Discovery competition* KDDCUP 2015 (organized by *ACM Special Interest Group on Knowledge Discovery and Data Mining*), where it was used by 10 out of 10 winning teams [CHEN16]. Since then, the library XGBoost has been extended to support multiclass classification and has become a widely used machine learning library. SHWARTZ-ZIV and ARMON have shown recently that XGBoost shows on-par performance or even outperforms deep learning models such as TabNet [ARIK21] in both regression and classification tasks for tabular data, with substantially shorter training and hyperparameter optimization time [SHWA22].

As with Lasso regression, the objective function for XGBoost is also formulated as the sum of a loss function and a regularization term:

$$\mathrm{obj}(\theta) = \min_{\theta} \left( L(\theta) + \Omega_{\gamma,\lambda}(\theta) \right). \tag{30}$$

In XGBoost, the regularization term $\Omega_{\gamma,\lambda}(\theta)$ specifically penalizes the complexity of the decision trees. The term can be expressed as:

$$\Omega_{\gamma,\lambda}(\theta) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2, \tag{31}$$

where:

- $T$ is the number of leaves in the tree.
- $w_j$ represents the score on the $j$-th leaf.
- $\gamma$ is a parameter that penalizes the number of terminal nodes, encouraging simpler trees.
- $\lambda$ is a parameter that controls the magnitude of the leaf scores.

The first term $\gamma T$ encourages sparsity in the resulting model by penalizing the number of leaves, thereby rewarding simpler models and reducing the risk of overfitting. The second term $\frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2$ penalizes large leaf scores and therefore stabilizes the training process and reduces the risk of overfitting. In this work, loss function $L(\theta)$ was chosen as the MSE (see Eq. 29). The hyperparameters $\gamma$ and $\lambda$ can be chosen with cross-validation as described for Lasso regression.

Compared to traditional GBT implementations, XGBoost is optimized for performance by leveraging multi-core processors and GPU processors, making it highly scalable for large datasets. Furthermore, XGBoost prevents overfitting with a technique called *tree pruning* during the tree-building process, which reduces model complexity by removing unnecessary tree branches. XGBoost can capture linear and non-linear relationships between the predictor variables and the labels through its ensembles and the flexible structure of decision trees. For the predictive modeling tasks in this work, the PYTHON library *XGBoost*, created and maintained by CHEN et al. [CHEN16], was utilized.

**Principal component analysis** *Hauptkomponentenanalyse*

Principal component analysis is a method that, as BISHOP (2006) aptly described, represents an "*orthogonal projection of the data onto a lower dimensional linear space, known as the principal subspace, such that the variance of the projected data is maximized for the first principal component, and subsequent components maximize the remaining variance*" [BISH06]. The primary goal is to find a new representation $\mathbf{Z}_k$ of the data in the principal subspace, where $\mathbf{Z}_k \in \mathbb{R}^{n \times k}$ and $\mathbf{X} \in \mathbb{R}^{n \times p}$, where typically $k \ll p$, through a series of linear transformations. Figure 2.10 visualizes a projection from a 3-dimensional into a 2-dimensional space.

One way of performing PCA on data is by performing *Singular value decomposition* (SVD) on the centered model matrix, i.e., after subtracting the feature mean values. In the first step, the mean of each feature is subtracted from the dataset, centering it around the origin. The centered model matrix $\tilde{\mathbf{X}}$ is then factorized with SVD:

$$\tilde{\mathbf{X}} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathsf{T}}, \tag{32}$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthogonal matrices containing left and right singular vectors, respectively, and $\mathbf{\Sigma}$ is a diagonal matrix with singular values in descending order [STRA06]. The columns of $\mathbf{V}$ are called *principal directions*. When projected onto, these directions maximize

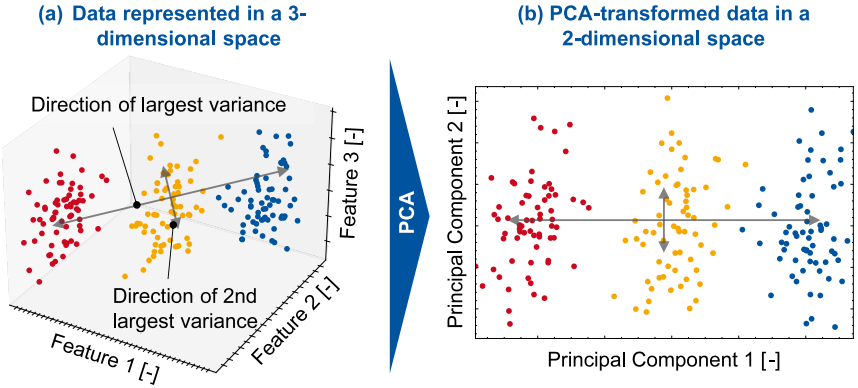



**Figure 2.10:** Visualization of (a) data in a 3-dimensional space and (b) the data after undergoing PCA transformation in a 2-dimensional space.
*Visualisierung von (a) Daten in einem 3-dimensionalen Raum und (b) ebendieser Daten, die einer PCA-Transformation unterzogen wurden in einem 2-dimensionalen Raum.*

the data variance [BISH06]. The first principal direction captures the largest amount of variance, the second direction captures the second largest amount of variance, and so on. Principal directions are orthogonal to each other since $\mathbf{V}$ is an orthogonal matrix.

Followingly, the data is projected to the coordinate system spanned by the principal directions, yielding the so-called *principal components* (PCs). By multiplying $\tilde{\mathbf{X}}$ from right with $\mathbf{V}$, one obtains the data transformed to the coordinates spanned by the principal directions:

$$\mathbf{Z} = \tilde{\mathbf{X}}\mathbf{V}. \tag{33}$$

For the purpose of dimensionality reduction, *Truncated SVD* refers to an incomplete SVD that retains only the top $k$ singular values and their corresponding singular vectors. This truncated representation can be particularly beneficial in ML scenarios involving large model matrices, where only a reduced representation is needed. The projection of the data to the lower-dimensional coordinate system spanned by $k$ principal directions is obtained with the multiplication of the centered matrix $\tilde{\mathbf{X}}$ with $\mathbf{V}_k$, where $\mathbf{V}_k$ denotes the first $k$ columns of $\mathbf{V}$:

$$\mathbf{Z}_k = \tilde{\mathbf{X}}\mathbf{V}_k, \tag{34}$$

and $\mathbf{Z}_k$ is of size $n \times k$. Another approach to perform PCA is to perform eigenvalue decomposition for the empirical covariance matrix $\mathbf{C}$ of $\tilde{\mathbf{X}}$, where

$$\mathbf{C} = \frac{1}{n-1}\tilde{\mathbf{X}}^{\mathsf{T}}\tilde{\mathbf{X}}. \tag{35}$$

The eigenvalues of $\mathbf{C}$ are the squared singular values of $\tilde{\mathbf{X}}$ divided by $n-1$:

$$\lambda_i = \frac{\sigma_i^2}{n-1}, \tag{36}$$

where $\sigma_i$ denotes the $i$-th singular value of $\tilde{\mathbf{X}}$. The proportion of variance of the data explained by each principal component is given by:

$$\text{Variance explained by } i\text{-th PC} = \frac{\sigma_i^2}{\sum_{j=1}^{p}\sigma_j^2} \tag{37}$$

where $\sigma_i$ is the $i$-th singular value. In practice, $k$ is chosen to retain a desired amount of cumulated variance, e.g., 80 %, or the components with a larger variance than the average variance. However, this varies by use case.

In summary, PCA can be performed by performing the SVD of the centered data matrix to identify the principal directions that maximize the variance of the data when projected onto them. The data is then represented in this new linear basis, providing a compact and often more interpretable view of the original dataset.

This work employs the PCA implementation from sklearn. Depending on the input data's shape and the chosen $k$, it either performs a full SVD or a truncated SVD as detailed in [HALK11].

Within this work, PCA is utilized to:

(i) visualize high-dimensional datasets from an industrial fineblanking production (Chapter 5) and

(ii) transform features from laboratory experiments in this study into a scalar TCM indicator, facilitating the qualitative analysis of this indicator's evolution across datasets (Chapter 8).

**Uniform manifold approximation** *Uniform manifold approximation*

The state of the art non-linear dimensionality reduction algorithm UMAP is renowned for visualization [BECH19] and preprocessing [ALLA20] of high-dimensional datasets. Its primary objective is to preserve both the global and local structures of high-dimensional data by capturing the inherent similarities between data points. This is achieved by constructing a low-dimensional representation using a graph-based approach, where the graph edges encode these similarities [MCIN18].

Creating a fuzzy topological representation of the input data marks the initial step of UMAP, specifically a *fuzzy simplicial complex* comprised of *fuzzy simplicial sets*. This involves constructing a weighted $k$-neighbor graph, where weights represent the probability or confidence of a point's membership within a particular fuzzy simplicial set, effectively encoding the high-dimensional similarities. Following this, UMAP constructs a low-dimensional simplicial complex. In this space, the distances between embeddings are transformed into low-dimensional similarities using a "*smooth approximation to the high-dimensional similarity function*" [DAMR21], ensuring that the low-dimensional representation captures the essence of the high-dimensional data structure for all pairs of points.

UMAP then optimizes the positions of the low-dimensional data points to minimize the *cross-entropy* between the high and low-dimensional simplicial complexes, as given by [MCIN18]:

$$C((A,\zeta),(A,\nu)) = -\sum_{a\in A}\left(\zeta(a)\log(\nu(a)) + (1-\zeta(a))\log(1-\nu(a))\right), \tag{38}$$

where $A$ is the set of all pairs of data points. For each pair $a \in A$, $\zeta(a)$ denotes the similarity between data points in the high-dimensional representation, and $\nu(a)$ denotes the similarity in the low-dimensional representation [MCIN18].

The term $\zeta(a)\log(\nu(a))$ is the *attraction component*, ensuring that points that are close or connected in the high-dimensional space remain close in the low-dimensional representation. Conversely, the term $(1-\zeta(a))\log(1-\nu(a))$ is the *repulsion component*, ensuring that points that are not closely connected in the high-dimensional space are pushed apart in the low-dimensional representation.

Figure 2.11 offers a visual representation of a fuzzy topological structure. Here, color decay indicates the confidence of a data point's membership within a specific fuzzy simplicial set, diminishing as the distance from the data point grows.

Circles' radii around data points adjust based on local data point density, capturing their diverse distances and relationships. This ensures each sample $x_i$ is locally connected to at least one other sample with a weight of 1 (highest confidence), termed the *local-connectivity constraint*. Notably, this method is computationally efficient, as membership calculations for a data point within a fuzzy simplicial set are restricted to its $k$-nearest neighbors [MCIN18].

Within this research and analogously to PCA, UMAP serves to

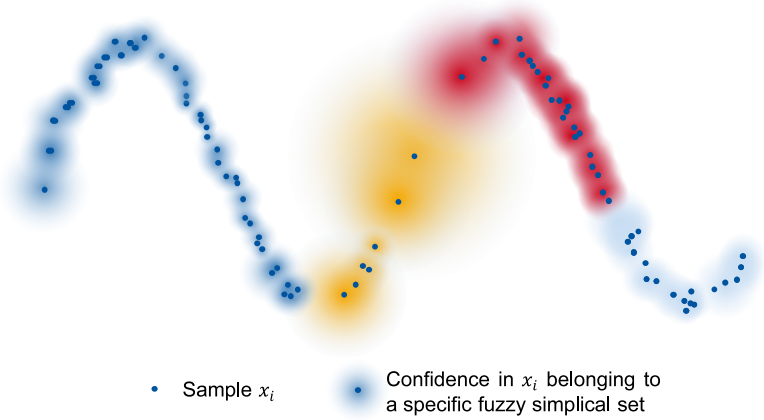



**Figure 2.11:** Intuitive visualization of a fuzzy topological representation of input data (adapted from [UMAP23]). Different colors represent different simplicial sets.
*Intuitive Visualisierung einer fuzzy-topologischen Repräsentation von Eingabedaten. Unterschiedliche Farben stehen für unterschiedliche Simpliziale Mengen.*

(i) visualize high-dimensional datasets stemming from industrial fineblanking production (Chapter 5) and to

(ii) transform selected features from laboratory experiments into a scalar TCM indicator (Chapter 8).

To this end, the PYTHON library *umap-learn* is employed. The library was developed and is maintained by the authors of [MCIN18].

**Autoencoders** *Autoencoder*

Autoencoders are ANNs that traditionally are used to learn efficient embeddings $\mathbf{H}$ of input data $\mathbf{X} \in \mathbb{R}^{n \times p}$ of $\mathbf{X}$. Efficient in this context means that the embedding $\mathbf{H} \in \mathbb{R}^{n \times \ell}$ (the input is said to be *encoded*, $\mathbf{H}$ is also referred to as *code*) of input data $\mathbf{X}$ has substantially lower dimensionality ($\ell \ll p$), but preserves important information about $\mathbf{X}$. Preserving important information means here that the *reconstruction* (or *decoding*) $\mathbf{R}$ of $\mathbf{X}$ from $\mathbf{H}$ yields a small *reconstruction error* $E(\mathbf{X}, \mathbf{R})$, where $E(\cdot\ ,\ \cdot)$ denotes a suitable error metric (e.g., MSE). Effectively, this means that the ANN has the task of *copying* the input $\mathbf{X}$ to $\mathbf{R}$ with minimal loss but is restricted in the capacity of its embedding of $\mathbf{X}$ in $\mathbf{H}$ [GOOD16].

The described AEC is a so-called *undercomplete* AEC since the embedding's dimension is lower than the input's and reconstruction's dimension. While undercomplete AECs are not the only type of AECs, the usage of AECs within this work is limited to undercomplete AECs. The schematics of an undercomplete AECs are visualized in Figure 2.12 where $e(\mathbf{X})$ denotes the encoder and $d(\mathbf{H})$ the decoder.

Within this work, AECs are used to detect *anomalies* within datasets, specifically to identify misegmentation in Section 4.3.2, i.e., samples $x_i$ that deviate substantially from the majority of the samples within $\mathbf{X}$. The general theory of ANNs (e.g., activation functions, optimizers, etc.) is omitted (see [GOOD16] and [BISH06] for the general theory of ANNs and AECs).

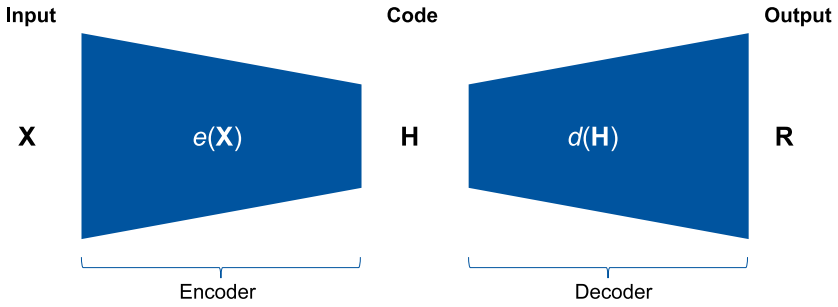



**Figure 2.12:** Schematic depiction of an undercomplete AEC (adapted from [NIEM21]).
*Schematische Darstellung eines unvollständigen AEC.*

## 2.3 Foundations of time series analysis

*Grundlagen der Zeitreihenanalyse*

Section 2.3 introduces the *concept of time series analysis*. Section 2.3.1 lays out definitions within the context of time series analysis and motivates *spectral-domain analysis*. Section 2.3.2 introduces important concepts and algorithms in time series analysis.

### 2.3.1 Introduction to and key definitions in time series analysis

*Einführung in und wichtige Definitionen der Zeitreihenanalyse*

A time series of length $N$ is a sequence of observations ordered in time:

$$x(t) = (x(t_0), x(t_1), \ldots, x(t_{N-1})), \tag{39}$$

with $x(t_k) \in \mathbb{R}^p$, $p = 1$ in case of a univariate, and $p > 1$ in case of a multivariate time series. In this context, $t_i$ denotes *the time of observation*.

A time series where the time intervals between $x(t_k)$ and $x(t_{k+1})$ are equal for all $k \in \{0, 1, \ldots, N-2\}$ is called *evenly spaced*. All generated time series are assumed to be evenly spaced for this work. This assumption is reasonable since these series were generated using appropriate measurement equipment.

Time series are a crucial research subject, occurring in a wide range of disciplines such as physics, finance, sociology, and manufacturing. Time series typically result from multiple processes acting simultaneously or with a delay (*lag*), potentially influencing each other [BOX15]. Comprehending the complex relationships between these processes is crucial for accurate analysis, and a key tool for understanding these relationships is *spectral analysis* [BLOO00].

Spectral analysis is a fundamental tool used to uncover complex relationships in time series. It works by decomposing the time series from its *time-domain representation* into its underlying *frequency-domain representation*, facilitating the interpretation of the time series as a composition of different contributing components or generating processes and allowing subsequent analysis of these (see Figure 2.13) [BRIL01].

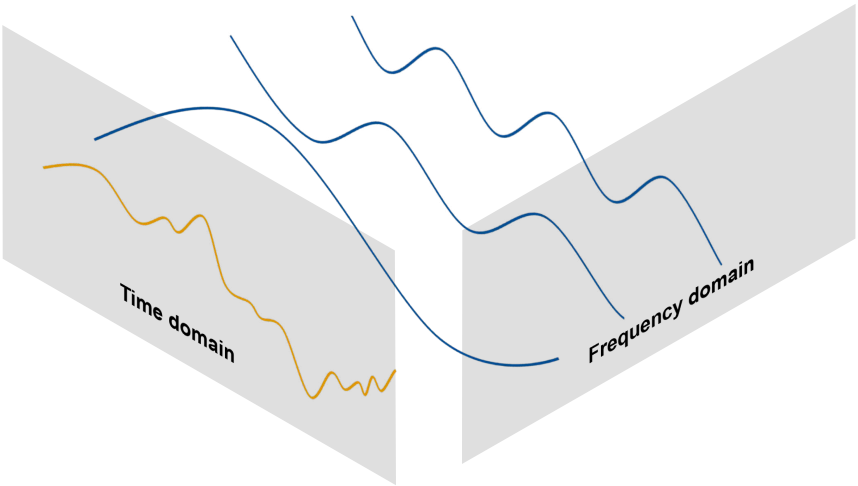



**Figure 2.13:** Schematic visualization of time-domain vs. frequency-domain representation of a signal, (adapted from [KONG20]).
*Schematische Visualisierung der Darstellung eines Signals im Zeitbereich und im Frequenzbereich.*

Such a transformation is subject to the *Shannon-Nyquist sampling theorem*. This theorem states that the sampling frequency $f_\mathrm{s}$ should be at least twice the maximum frequency present in the signal, known as the NYQUIST frequency $f_\mathrm{Nyquist}$:

$$f_\mathrm{s} \geq 2 f_\mathrm{Nyquist}. \tag{40}$$

This condition is necessary to prevent aliasing and ensure the perfect reconstruction of the signal [SHAN49]. In practice, this means the time series must be sampled at a rate sufficient to capture the full range of frequencies (*frequency components*) of the underlying signal. By identifying and analyzing these components, spectral analysis can reveal the dynamics of the system generating the time series.

For TCM, utilizing spectral analysis to analyze sensor data from a production line can help decompose signals from a complex manufacturing process into several fundamental subprocesses, such as friction or material deformation. This decomposition can subsequently be used to identify patterns and anomalies that are indicative of potential issues with the tool, such as unexpectedly high tool wear [UNTE21]. Revealing the underlying dynamics of a system through spectral analysis can provide a more accurate and comprehensive understanding of the system, leading to better-informed decision-making and improved performance [BACC15].

### 2.3.2 Selected concepts and methods in time series analysis

*Ausgewählte Konzepte und Methoden in der Zeitreihenanalyse*

Following, selected methods in time series analysis are presented that were used in the time series analysis carried out within this work due to their scalability and versatility.

**Fast fourier transform** *Schnelle Fourier-Transformation*

One of the most widely used algorithms across all scientific domains is the *fast Fourier transform* (FFT). The FFT is an algorithm for efficiently computing the *discrete Fourier transform* (DFT), a mathematical operation that transforms a discrete time-domain signal into its frequency-domain representation. For the $k$-th *frequency bin* or distinct frequency interval, the *Fourier coefficient* $\hat{x}(k) \in \mathbb{C}$ is calculated from the time-domain signal amplitude $x(t_j)$ with

$$\hat{x}(k) = \sum_{j=0}^{N-1} x(t_j) e^{-2\pi i \frac{jk}{N}} \text{ for } k \in \{0, 1, 2, \ldots, N-1\}, \tag{41}$$

where $e$ denotes EULER's number, $\pi$ denotes the number Pi, and $i$ denotes the imaginary unit.

Figure 2.14 presents an AE signal during the stripping phase of a fineblanking process, showcasing both its time-domain and frequency-domain representations. Specifically, the visualization of the frequency-domain shows the *magnitude*, or amplitude, for each respective frequency bin.

Due to its computational efficiency and scalability for large datasets, the FFT is a valuable tool for spectral analysis, which is why it was chosen for this work instead of other spectral analysis methods such as the *Wavelet Transform*.

Fast Fourier transforms performed in this work employ the *fft* module from the PYTHON package *NumPy* [HARR20]. This module implements the COOLEY-TUKEY FFT algorithm, which is highly efficient due to its ability to recursively decompose a DFT of any composite size (typically a power of two) into many smaller DFTs, significantly reducing the computational complexity [COOL65].

Within the analyses of this work, FFT is utilized in two ways:

(i) Explicitly in Chapter 5, Chapter 6, and Chapter 8 as an integral building block of WELCH's method [WELC67], to *decompose AE signals* into their frequency components for further analysis or visualization. This approach allows for trend recognition and feature extraction within specific frequency ranges, and importantly, it enables the reduction of noise by averaging the FFTs of overlapping signal segments [SOLO91].

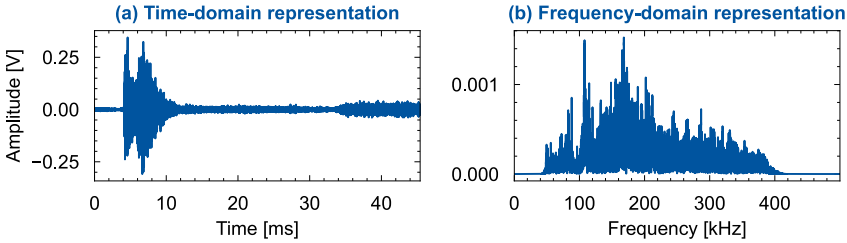



**Figure 2.14:** Time-domain vs. frequency-domain representation (real part of the DFT) of an AE signal acquired during the stripping phase of fineblanking.
*Darstellung des Zeitbereichs im Vergleich zum Frequenzbereich (Realteil der DFT) eines akustischen Emissionssignals, das während der Abstreifphase beim Feinschneiden aufgezeichnet wurde.*

(ii) Implicitly in Chapter 5, Chapter 7, and Chapter 8, as the spectral features extracted from times series with the PYTHON library *TSFEL* perform FFTs.

(iii) Implicitly in Chapter 4, as the utilized algorithm (MASS) performs FFTs for efficient similarity search in time series.

While the FFT is efficient and versatile, it does not adequately handle signals that exhibit variations in their statistical properties over time (such as mean, variance, etc.), i.e., signals generated by *non-stationary* stochastic processes [BOX15] (followingly denoted as *non-stationary signals*). If a sheet-metal stamping stroke is modeled as a stochastic process, it is reasonable to assume that this process is non-stationary and, thus, generates non-stationary signals (e.g., as stated by [SHAN17]) since a stroke represents a sequence of different subprocesses that vary in their properties (e.g., process forces, contact surfaces, etc.). Hence, the properties of a realization of the stochastic process (*i.e., a force or AE signal of a specific stroke*) are highly dependent on the point in time when the process is observed or the *time origin*. Therefore, it becomes essential to utilize a method capable of effectively portraying the frequency content of non-stationary signals. WELCH's method (or *WOSA*, WELCH*'s Overlapped Segmented Average*) performs FFTs on shorter, overlapping segments of the signal after applying a window function to them, and subsequently averages them. In practice, this leads to a more reliable estimate of the spectral content of an underlying signal. The specifics of this will be discussed as follows.

**Power spectral density, power spectrum, and WELCH'S method** *Spektrale Leistungsdichte, Spektrale Leistung, und WELCH- Methode*

The *power spectral density* (PSD) and the *power spectrum* (PS) are foundational tools for analyzing frequency-based characteristics of signals. The PS represents the *unnormalized power of each frequency bin* within the signal. The intuitive estimation of the PS of a signal uses its DFT:

$$P_{XX}(k) = |\hat{x}(k)|^2, \tag{42}$$

for each $k \in \{0, \ldots, N-1\}$, where $P_{XX}(k)$ denotes the PS at the $k$-th frequency bin, and $\hat{x}(k)$ the Fourier coefficient of the $k$-th frequency bin. In contrast, the PSD delineates how a signal's *power* is distributed across different frequencies, effectively normalizing the PS. In practice and for digital signals, the PSD of a time-domain signal $x(t)$ at the $k$-th frequency bin can be estimated as follows:

$$S_{XX}(k) = \frac{1}{N f_\mathrm{s}} |\hat{x}(k)|^2 \tag{43}$$

for each $k \in \{0, \ldots, N-1\}$, where $S_{XX}(k)$ denotes the PSD at the $k$-th frequency bin, $N$ the number of samples in the signal, $f_\mathrm{s}$ the sampling frequency, and $\hat{x}(k)$ Fourier coefficient of the $k$-th frequency bin. This equation normalizes the power over the duration of the signal, making the PSD a quantity that can be compared between signals of different lengths.

However, directly estimating the PSD and PS using the DFT presents inherent challenges. The DFT, like Fourier Transforms in general, fits sinusoidal components to the entire signal, implicitly assuming that the signal's frequency content is stationary over time. This assumption is often violated in real-world scenarios, where signals may exhibit time-varying behavior. For such non-stationary signals, the DFT may not provide an accurate or complete representation of the signal's true frequency content. Additionally, sensitivity to noise can distort the true spectral content when using the DFT, which is especially critical in an industrial setting.

A more robust and one of the most used methods for PSD estimation is WELCH's method due to its variance reduction capabilities and efficient computation via FFTs [STOI05].

In WELCH's method, the signal is divided into overlapping segments with a fixed window size and overlap. Then, a window function is applied to each segment, computing a *modified periodogram* for each segment. The resulting modified periodograms are then averaged [WELC67]. By doing so, WELCH's method acts as a noise reduction technique, leading to a smoother and more reliable estimate of both the PSD and the PS at the cost of a lower frequency resolution. Figure 2.15 visualizes the PSD estimated with (a) the DFT (see Eq. 43) and (b) WELCH's method with a window size of 512 for a time-domain signal from an industrial setting of length 45,400 and a sampling frequency $f_s$ of 1 MHz. Clearly, (b) represents a less noisy and smoother estimate of the PSD in comparison to (a).

The choice of the window size in WELCH's method is a critical parameter that represents a trade-off between frequency resolution and variance reduction. In this work, smaller window sizes were chosen for visualization purposes, hence achieving a substantial variance reduction. Since the signal is divided into more segments with smaller window sizes, more averaging of the modified periodograms is performed, and consequently, smoother PSD plots result. In turn, the frequency resolution is substantially reduced. These smooth PSD plots allow the identification of broader trends in large datasets. Figure 2.16 demonstrates the trade-off between frequency resolution and variance reduction for different window sizes. While a window size of (a) 16 leads to a substantial reduction in variance (while still preserving important information about the spectrum), the frequency resolution is coarse. In contrast, a window size of (d) 16,384 does not substantially reduce the variance, yet it preserves a high-frequency resolution. Window sizes of (b) 256 and (c) 1,024 represent more balanced trade-offs between the variance and the frequency resolution.

For visualization purposes, the PSD was employed within this work to broadly compare spectral content across a wide range of signals of different lengths and sampling rates. The PS was utilized to extract specific features from the spectrum, such as the absolute power of specific frequency components. The PS is particularly useful for this purpose, as it directly reflects the signal's energy content in various frequency bands (see Section 4.4). Comparisons of the resulting features between signals of different lengths or with different sampling rates were avoided. Larger window sizes were used for feature extraction purposes with subsequent aggregation of the features, such as calculating the mean frequency or the relative bandpower within a certain frequency range (see Section 4.4) since the aggregation

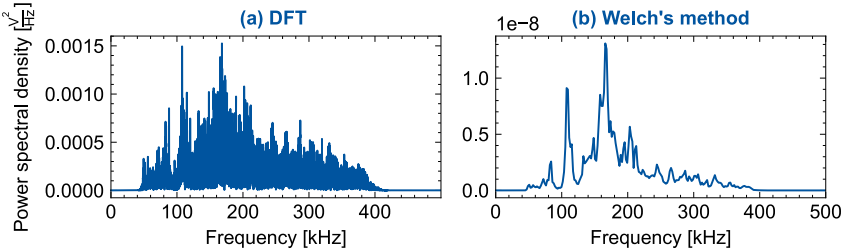



**Figure 2.15:** Visualization of PSD estimates using (a) the DFT of a time-domain signal and (b) WELCH's method with a window length of 512 for a signal of length 45,400.
*Visualisierung von PSD-Schätzungen mit (a) der DFT eines Zeitsignals und (b) der Methode von WELCH mit einer Fensterlänge von 512 für ein Signal der Länge 45.400.*

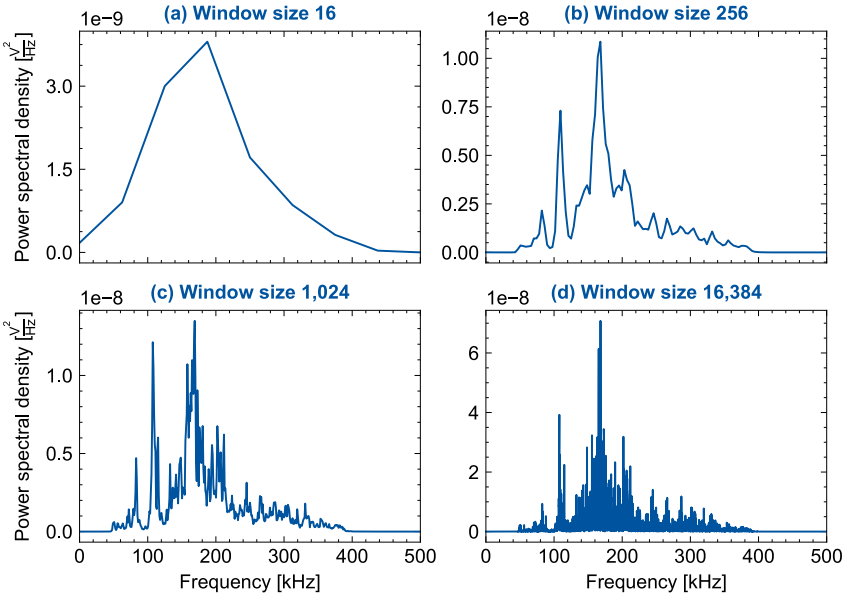



**Figure 2.16:** Visualization of the impact of the window size in WELCH's method for window sizes (a) 16, (b) 256, (c) 1,024, and (d) 16,384 for a signal of length 45,400.
*Visualisierung der Auswirkungen der Fenstergröße bei der Methode von WELCH für die Fenstergrößen (a) 16, (b) 256, (c) 1,024 und (d) 16,384 für ein Signal der Länge 45,400.*

makes the features less sensitive to fluctuations in the raw spectral data. Therefore, a higher frequency resolution was obtained.

To this end, the function *scipy.signal.welch* from the PYTHON library *SciPy* [VIRT20] was used, which allows estimations of both the PSD and PS of a time-domain signal with WELCH's method. A HANNING window was chosen as a window function, which, together with a segment overlap of 50 %, represents a commonly used approach [HEIN02]. The results of these spectral analyses are presented in Chapter 5, Chapter 6, and Chapter 8.

**Mueen's algorithm for similarity search** *Mueen's Algorithmus zur Ähnlichkeitssuche in Zeitreihen*

In many applications involving time series data, such as time series classification (i.e., the task of assigning a label $y$ to a time series $x(t)$), it is essential to identify and align similar patterns between several observed time series to facilitate accurate analysis and interpretation of the underlying processes [HU13]. However, as time series data may contain shifts, it becomes challenging to identify and align these patterns accurately.

A data mining technique for handling such cases is *Mueen's algorithm for similarity search* (MASS), which enables the identification of similar subsequences of a fixed length within the time series data [MUEE22]. Specifically, MASS operates by comparing a query sequence $x_Q(t)$ of length $M$ to a longer time series, $x_T(t)$ of length $N$ with $N > M$, in order to identify subsequences within $x_T(t)$ of lengths $M$ that are most similar to $x_Q(t)$, see Figure 2.17.

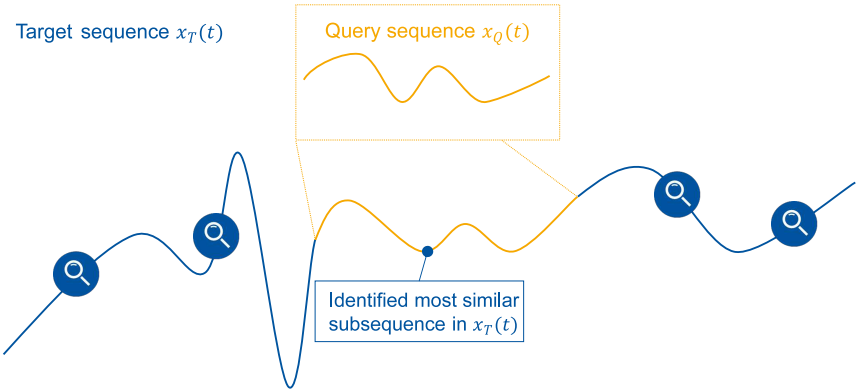



**Figure 2.17:** Schematic visualization of similarity search in sequences, e.g., time series.
*Schematische Visualisierung der ähnlichkeitssuche in Sequenzen, z. B. Zeitreihen.*

Without considering early abandoning techniques that rely on the data having a certain structure, $N - M + 1$ subsequences of $x_T(t)$ need to be compared to the query, resulting in a computational complexity of $O(NM)$ if this would be done in a brute-force implementation, e.g., with a *for*-loop. Although MASS calculates the Euclidean distance of every standardized subsequence to the query, it reduces the computational complexity to $O(Nlog(N))$ [MUEE09], which is especially significant for long sequences, such as the time series analyzed in this work. MASS achieves this by exploiting the speed-up of a convolution of two sequences in the time domain by calculating the element-wise product of their DFTs and subsequently transforming the result back into the time-domain. Specifically, MASS i) reverses $x_Q(t)$ to $x_Q(t)^{(R)}$ and zero-pads both $x_T(t)$ and $x_Q(t)^{(R)}$ to length $2N$, ii) performs FFTs of the resulting sequences, iii) multiplies both frequency-domain sequences element-wise, and vi) performs an inverse fast Fourier transform (IFFT) of the result. Finally, the *distance profile* $D$ is calculated as a sequence of the Euclidean distances of all standardized subsequences of length $M$ in $x_T(t)$ and the standardized query $x_Q(t)$.

Algorithm 1 presents pseudocode for MASS v2 inferred from the MATLAB and PYTHON code provided in [MUEE22]. In Algorithm 1, $\odot$ represents the element-wise product, $\Sigma_{x_T}$ the moving standard deviation and $M_{x_T}$ the moving mean (both with a window length of $M$) of $x_T(t)$.

The distance profile $D$ of length $N - M + 1$ is calculated with

$$D = \sqrt{2M(1 - \frac{z(t)[M:N] - M\mu_{x_Q}M_{x_T}[M:N]}{M\sigma_{x_Q}\Sigma_{xT}[M:N]})}, \tag{44}$$

where $M$ denotes the length of $x_Q(t)$, $[M:N]$ denotes a subsequence of length $N - M + 1$ starting at index $M$ and ending at index $N$ of the corresponding sequence, $z(t)$ the IFFT of the element-wise product of $x_T(t)^{(P)}$ (zero-padded $x_T(t)$ to length $2N$) and $x_Q(t)^{(R,P)}$ (reversed and zero-padded $x_Q(t)$ to length $2N$), $\mu_{x_Q}$ and $\sigma_{x_Q}$ the mean and standard deviation of $x_Q(t)$, and $M_{x_T}$ and $\Sigma_{x_T}$ the moving mean and standard deviation of $x_T(t)$ [YEH18].

For the query and target sequence lengths used within this work, a time complexity reduction from $O(NM)$ to $O(N \log N)$ represents a substantial improvement. This work employs

**Algorithm 1** Mueen's algorithm for similarity search v2

**Ensure:** Time series $x_T(t)$, Query $x_Q(t)$
**Require:** Distance Profile $D$
1: $N \leftarrow$ Length$(x_T(t))$
2: $M \leftarrow$ Length$(x_Q(t))$
3: $\Sigma_{x_T}, M_{x_T} \leftarrow$ MovStDevMean$(x_T(t), M)$
4: $\mu_{x_Q}, \sigma_{x_Q} \leftarrow$ StDevMean$(x_Q(t))$
5: $x_Q(t)^{(R)} \leftarrow$ Reverse$(x_Q(t))$
6: $x_Q(t)^{(R,P)} \leftarrow$ ZeroPad$(x_Q(t)^{(R)}, 2N)$
7: $x_T(t)^{(P)} \leftarrow$ ZeroPad$(x_T(t), 2N)$
8: $z(t) \leftarrow$ IFFT(FFT$(x_Q(t)^{(R,P)}) \odot$ FFT$(x_T(t)^{(P)}))$
9: $D \leftarrow$ DistProfile$(z(t), M, \Sigma_{x_T}, M_{x_T}, \sigma_{x_Q}, \mu_{x_Q})$
10: **return** $D$

MASS to i) identify segments of interest within long and noisy AE signals and ii) align the starting point of shifted signals in Chapter 4. The implementation of the algorithm in the PYTHON package *stumpy* [LAW19] is utilized for this purpose, allowing for fast and accurate pattern matching across the datasets.

**Digital BUTTERWORTH filters** *Digitale BUTTERWORTH-Filter*

Analog signals are vulnerable to disruptions in industrial settings, specifically due to mechanical, electrical, and environmental noise [SEEM16]. Such disturbances can introduce undesired frequencies into the signal, ranging from high frequencies originating from interfering noises to low frequencies emerging from oscillations of mechanical actions. One way to mitigate these effects and enhance the *signal-to-noise ratio* (SNR) is to use digital filters. These can function as high-pass, low-pass, band-pass, or band-stop filters, each effectively attenuating specific unwanted frequency components [PROA07].

The design and application of digital BUTTERWORTH *filters*, also known as *maximally flat magnitude filters*, is a digital signal processing (DSP) technique. BUTTERWORTH filters are widely used to reduce noise in various kinds of time series data. They are engineered to have a smooth frequency response with no ripples in both the *passband*, i.e., frequency components allowed to pass through the filter, and the *stopband*, i.e., frequency components to attenuate. This ensures minimal distortion in the filtered signal data, which preserves the shape of the signal while excluding noise from unwanted frequency bands [BUTT30]. Hence, they are suitable for a broad range of applications in time series analysis.

A BUTTERWORTH filter, characterized by its order $n$, cutoff frequency $\omega_c$, and type (low-pass, high-pass, band-pass, or band-stop), seeks to transform an input signal, $x(t)$, into a filtered output signal, $y(t)$, by attenuating specific frequency components of $x(t)$. Simultaneously, it presents a uniform sensitivity for the desired frequencies [BUTT30]. Figure 2.18 depicts the frequency responses of BUTTERWORTH *infinite impulse response* (IIR) filters with orders of $n = 2$, $n = 4$, and $n = 6$ and a cutoff frequency $\omega_c$ of 1 MHz. The *gain* here signifies the behavior of a filter, indicating whether it permits frequency components to pass through or whether it attenuates them.

This work uses band-pass BUTTERWORTH filters as a DSP technique to reduce noise in the acquired AE data in Chapter 4. The PYTHON module *scipy.signal* provides a convenient implementation of BUTTERWORTH filters, which is utilized to design and specify the filters to the data.

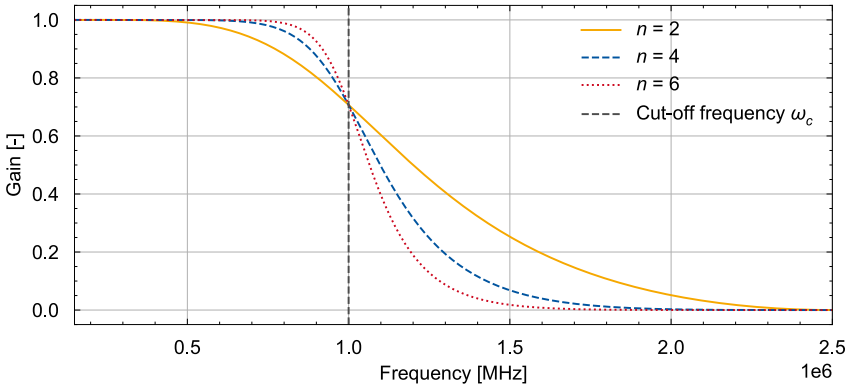



**Figure 2.18:** Visualization of the impact of the filter order $n$ on the frequency response of a signal.
*Visualisierung des Einflusses der Filterordnung n auf den Frequenzgang eines Signals.*

## 2.4 State of the art for indirect tool condition monitoring of sheet-metal stamping tools

*Stand der Technik zur indirekten Zustandsüberwachung von Stanzwerkzeugen*

The previous sections focused on laying the theoretical foundation needed to grasp the state of the art of indirect TCM in sheet-metal stamping processes and identifying research gaps. Following, this state of the art will be presented with a focus on approaches based on ML.

In past research endeavors, both supervised and unsupervised ML methods were used in indirect TCM approaches in sheet-metal stamping processes. Unsupervised machine learning approaches to TCM do not require a costly label generation since these approaches can be utilized to detect structural changes within the generated process signals [NIEM21].

Prior works showed that UMAP and PCA embeddings for AE signals [UNTE21] and for force-displacement signals [NIEM22] in fineblanking resemble the expected tool wear. The same was shown for reconstruction error sequences of AECs for both AE and force-displacement signals [NIEM21]. However, these works lack validation in the form of an approximated ground truth of the (either indirectly or directly) observed tool wear as a benchmarking tool for the developed approaches [UNTE23].

NIEMIETZ et al. derived wear estimators from short- and long-term variations of suitable representations extracted from domain-specifically preprocessed force-displacement signals to model the punch wear in four fineblanking experiments. The validity of the estimators was examined with sparse wear stage information, specifically the linear interpolation of only four supporting points per experiment since the work relies on direct examination of the punch. The supporting points were extracted from scanning electron microscope (SEM) images through the usage of a contour detection algorithm. The authors state that existing approaches to inferring punch wear information largely depend on artificially introduced wear states, i.e., punch or die chamfer manipulation, treating the wear state as binary (a similar argument was presented by [KUBI22]). This highlights the need for further research with industrial-grade process setups and emphasizes the importance of modeling and understanding transitions between wear stages [NIEM22].

A similar approach to [NIEM21] was followed by ASAHI et al. The authors used the reconstruction error sequence of an AEC, which was trained on collected time series data from

various sensor sources during a sheet-metal stamping process with an unworn punch state. The authors showed that the reconstruction error of the AEC, calculated for each process cycle, can be interpreted as an anomaly score for the respective process cycle [ASAH21]. However, the specific wear mechanisms observed were not laid out in the work. The authors used the reconstruction error as an input for classification models with three class labels representing three consecutive wear stages.

KUBIK et al. used five stages of artificially introduced abrasive *punch wear states* (i.e., different severities of wear) as class labels in a sheet-metal stamping process. The authors used support vector machines (a support vector machine is an ML algorithm) as classifiers to model the correlation between extracted features from force-displacement signals and the wear states. The signals were gathered for varying punch wear states, stroke frequencies, and sensor positions. KUBIK et al. evaluated the model performances based on their i) classification accuracies and ii) class separabilities. The authors emphasize the importance of leveraging domain knowledge during data acquisition, preprocessing, and transformation to optimize for model performance, such as adequate filtering and feature engineering. Moreover, the authors underscore the need to evaluate different process signals, such as AE, as correlation variables for the punch wear. Furthermore, the authors highlight that the application of ML methods is limited to describing discrete fault states, and the actual quantification of wear conditions is hardly found within literature [KUBI22].

Acoustic emission signals are based on transient elastic waves generated by the sudden release of energy, e.g., crack propagation, plastic deformation, or impacts [DEUT17]. In contrast to force sensors, AE sensors are inexpensive and easy to mount and maintain, making them suitable for industrial usage [UNTE21]. Compared to force-displacement signals, extracting relevant information from AE signals typically requires more effort due to their complex structure.

KOLLMENT et al. suggested segmenting the AE signal based on the cyclical recurring process phases [KOLL18]. The SNR can be further enhanced through the usage of suitable transformation techniques, such as frequency-domain analysis, see Section 2.3.2. Past research from BACCAR et al. showed that under sliding conditions, the behavior of the frequency components of AE signals changes with the wear state [BACC15].

SHANBHAG et al. used various frequency characteristics of AE signals from a semi-industrial sheet-metal stamping process setup that simulates industrial sheet-metal stamping conditions to detect the onset and severity of adhesive wear on the die [SHAN20]. The authors carried out stamping and scratch tests to simulate galling conditions between the die and the workpiece in industrial sheet-metal stamping. SHANBHAG et al. found that the mean frequency feature, a scalar value representing the averaged frequency content of a signal weighted by the power of the frequency bins, can be used to qualitatively identify the onset and severity of galling on the die [SHAN20]. SHANBHAG observed a decrease in the mean frequency of the acquired signals with the onset of galling between the die and the workpiece. With reference to BASSIUNY [BASS07], SHANBHAG et al. stated in earlier works that this decrease could potentially be attributed to "*the number of resonant frequency components that are exhibited at lower frequency range for the worn parts*" [SHAN18]. However, BASSIUNY artificially created fault conditions in a sheet-metal stamping process (e.g., misfeed, sheet metal thickness deviations, etc.) and analyzed a strain signal. Furthermore, SHANBHAG et al. identified the frequency range 100–200 kHz as the frequency range that contained the most relevant information about the onset and severity of galling for their process setup, both for the stamping and the scratch tests [SHAN20].

## 2.5 Identified research deficits

*Identifizierte Forschungsdefizite*

From the review of the state of the art in indirect TCM in sheet-metal stamping processes, several specific *research gaps* (RG1–RG5) were identified and addressed with the research questions formulated in Chapter 3.

**Lack of a systematic domain-specific preprocessing strategy for TCM with AE signals in the context of punch wear (RG1)** *Fehlen einer systematischen domänenspezifischen Vorverarbeitungsstrategie für TCM mit AE-Signalen im Zusammenhang mit Stempelverschleiß*

Since the punch is a critical tool component and its condition is highly relevant for the planning of maintenance intervals, attention towards leveraging an easy-to-acquire process signal as an indirect TCM source could potentially lead to novel insights into punch wear dynamics and further the development of indirect TCM systems for industrial sheet-metal stamping. However, a *systematic domain-specific preprocessing strategy for TCM with AE signals is notably absent* in current literature.

**Insufficient data from industrial-grade sample sizes and process setups in indirect TCM research (RG2)** *Unzureichende Daten(mengen) aus industriellen Settings in der Forschung zum indirekten TCM*

Current research in indirect TCM in sheet-metal stamping processes *lacks sufficient data obtained from industrial-grade sample sizes and process setups*. This limitation restricts the practical applicability of indirect TCM methodologies and their potential adoption in real-world sheet-metal stamping scenarios. Expanding research to include larger, more representative sample sizes and process setups could help address this gap and enhance the reliability and generalizability of findings.

**Lack of a suitable proxy for isolating punch wear, limited number of class labels (RG3)** *Fehlen eines geeigneten Proxys für die isolierte Betrachung des Stempelverschleißes, geringe Anzahl von Klassen*

The *lack of an identified suitable proxy for isolating punch wear* poses a challenge. While the quality of the sheared surface of the produced part can be utilized as a proxy for the die wear, no such proxy exists for the punch wear in the reviewed literature. The requirement to disassemble and reassemble the tool for punch wear assessment disrupts and influences both the process setup and, thus, the process signals. An ideal proxy could provide a non-intrusive method for monitoring punch wear, increasing the feasibility of continuous monitoring in industrial settings. Furthermore, this proxy could be utilized to validate unsupervised approaches, shedding light on the applicability and used as a benchmarking tool.

The *limited number of available class labels employed in supervised ML approaches* to indirect TCM fails to account for the transient nature of wear processes within sheet-metal stamping. This neglect of transition stages between wear states impedes meaningful inference from model decisions and undermines these models' predictive accuracy and potential applicability. Expanding the label spectrum, i.e., with a suitable dataset for regression instead of classification, could potentially capture the progression of wear states and offer more nuanced, accurate insights into punch wear dynamics.

**Lack of explainable predictive modeling of the punch wear from AE signals in industrial-grade process setups (RG4)** *Mangel an interpretierbaren Vorhersagemodellen für den Stempelverschleiß aus AE-Signalen in industriellen Prozess-Umgebungen*

A notable deficit in the current body of knowledge pertains to the *lack of explainable predictive modeling of the punch wear from AE signals in industrial-grade process setups*. While AE signals offer a promising avenue for indirect TCM methodologies, the correlation between these signals and the punch wear remains inadequately explored in industrial-grade process settings. This gap presents a significant impediment to the development of effective, reliable, and trustworthy monitoring strategies. Addressing this research gap could pave the way for more robust and reliable TCM systems, further augmenting the potential for proactive maintenance and reducing downtime in industrial sheet-metal stamping processes.

**Lack of a verified and industrial applicable scalar indicator for indirect TCM (RG5)** *Fehlen eines verifizierten und industriell anwendbarem skalaren TCM-Indikators*

While various works for laboratory settings exist, *no transfer of the results of laboratory experiments and the subsequent analyses to different process settings*, i.e., different geometry, lubrication, material, press, etc., exists for indirect TCM in sheet-metal stamping. This impairs the applicability of AE as an indirect TCM technology. A systematic transfer of laboratory results to an industrial setting can potentially allow insights into the dynamics of industrial production and provide a blueprint for the development of easy-to-interpret TCM indicators.

Chapter 3 will formulate the research questions and lay out the research methodology employed within this work.

# 3 Objective and research methodology

*Zielsetzung und Forschungsmethodik*

From the reported initial situation within Chapter 2, the following **economical problem statement** emerges for the sheet-metal stamping industry: *The processing of higher-strength steels in sheet-metal stamping leads to increased failure rates and reject production due to unpredictably high tool wear rates, resulting in reduced tool life, poor resource utilization, and impaired production planning capabilities.*

The increasing availability of computational resources and advances in the development of highly flexible ML methods lead to an increase in research activity in the domain of data-driven indirect TCM. However, this increase in research activity has not yet led to the development and adoption of a reliable indirect TCM system in the sheet-metal stamping industry to this day. While the punch represents a critical tool component, research on indirect punch wear monitoring systems shows substantial deficits. Thus, two **scientific problem statements** arise: *Firstly, despite existing research on indirect TCM, there is a notable lack of studies focusing on developing process images suitable for indirect TCM under industrial conditions. Secondly, the scientific literature is missing a systematic methodology and comprehensive analysis of the development of punch wear in sheet-metal stamping, particularly in the context of practical industrial settings.*

To overcome the aforementioned scientific and economic problem, the following **research objective** is formulated:

*The objective of this work is to achieve a data-driven and explainable model to estimate the punch wear in sheet-metal stamping processes with acoustic emission signals.*

To accomplish the research objective, the following **research hypothesis** is developed:

**With a combination of both qualitative and quantitative analysis of acoustic emission signals, it is possible to model and predict the punch wear state in sheet-metal stamping using suitable machine learning methods.**

To test the research hypothesis, the following research questions (RQ1–RQ5) need to be answered:

**RQ1** What is the process for designing a domain-specific preprocessing pipeline to isolate relevant process phases from acoustic emission signals in the context of punch wear monitoring in sheet-metal stamping?

**RQ2** What recurring patterns and differences can be qualitatively identified in the evolution of preprocessed industrial acoustic emission signals over several punch life cycles?

**RQ3** How can punch wear information be made available as a proxy variable?

**RQ4** How can machine learning models be used to predict the punch wear in sheet-metal stamping from acoustic emission signals and evaluated from an explainable AI perspective to derive feature importances?

**RQ5** How can the findings from RQ1–RQ4 be used to develop a scalar indicator for indirect tool condition monitoring in industrial sheet-metal stamping processes?

CHATGPT (GPT-4, `https://chatgpt.com/`) was used in this chapter to proofread texts generated by the author and for support with LATEX formatting. DEEPL (`https://www.deepl.com/translator/`) was used for translation purposes.

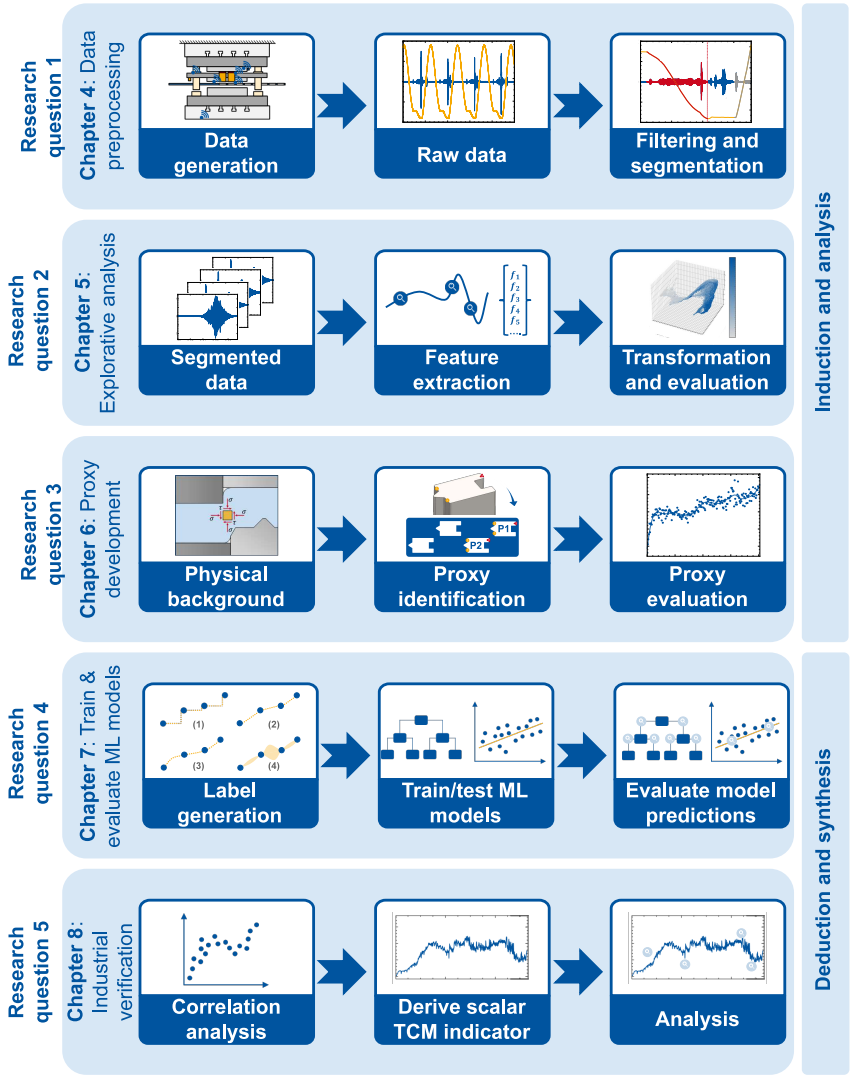



**Figure 3.1:** Illustration of the developed approach and research methodology.
*Illustration des entwickelten Ansatzes und der Forschungsmethodik.*

The work carried out to answer **RQ1** is documented in Chapter 4. The primary challenges expected in the development of a domain-specific preprocessing strategy for punch wear monitoring with AE data are i) the high volume of data generated by AE sampling rates in industrial and industrial-grade manufacturing settings and ii) the complex nature of the AE signal during sheet-metal stamping, which complicates the extraction of relevant process phases. Therefore, a preprocessing pipeline is to be developed and evaluated in three laboratory fineblanking experiments (a sequence of consecutive strokes will be referred to

as a *stroke series*). This pipeline will be extended and used to preprocess five industrial fineblanking stroke series.

The resolution of **RQ2** is the subject of the work detailed in Chapter 5. Conducting exploratory research on data from industrial settings presents substantial challenges due to the difficulty in controlling the environment and typically minimal context information. Thus, the central hurdle to addressing **RQ2** lies in overcoming these challenges and identifying patterns within and between stroke series. This will be accomplished through a qualitative analysis i) of low-dimensional embeddings and ii) of the evolution of domain-specific features for five industrial stroke series.

To address the existing research gap on isolated analysis of punch wear in sheet-metal stamping, the work detailed in Chapter 6 is designed to answer **RQ3**. The quality of the sheared surface of the scrap web segments (*proxy variable*) will be studied in terms of its evolution over completed process cycles for three different laboratory stroke series conducted with higher-strength steels. These series feature different sampling intervals of the scrap web segments. Specifically, two preparative experiments (low scrap web segment sampling rate) and one main experiment (high scrap web segment sampling rate) will be conducted. Additionally, the evolution of the frequency components of the AE signals will be assessed for the three laboratory stroke series.

Chapter 7 details the efforts to address the main research question **RQ4**. A data experiment will be conducted to explore the correlation between the evolution of the observed punch wear proxy variable and the acquired AE signals for the main experiment by predicting said punch wear proxy variable from the AE data, further differentiated into the two relevant process phases for the punch wear: shearing and stripping. The primary challenge in resolving **RQ4** stems from reducing bias that could be introduced through i) the employed ML models and ii) the imputation approach used to account for missing data. This will be addressed by using a large number of training and test sets and diverse model configurations coupled with a carefully designed data preprocessing strategy to avert data leakage. Additionally, a feature importance analysis across all models will be performed to derive a set of important features for further analysis in the context of research question **RQ5**.

Finally, the findings of Chapter 8 will be utilized to answer **RQ5**. Scalar indicators are sought after in industry due to their interpretability. A scalar TCM indicator will be developed by assessing the correlation of a set of scalar indicators and the observed punch wear proxy variable in the laboratory setting. The resulting correlations will rank the TCM indicators, and the top ranking indicator will be calculated for the industrial stroke series. The main challenge in interpreting the results lies in the insufficient context information accompanying the industrial data. This issue, however, reflects authentic industrial conditions and will be addressed by using available information to formulate assumptions about the stroke series.

*Inductive reasoning* is applied to **RQ1–RQ3**, starting with specific observations and moving towards broader generalizations. For **RQ1**, patterns in specific AE signals are identified to develop a preprocessing strategy. Resolving **RQ2** requires inductive analysis to detect patterns in industrial data and apply these findings to broader TCM. Unraveling **RQ3** also calls for inductive reasoning as the study of the punch wear proxy and its development across the lifespan of a punch will start from specific observations. *Deductive reasoning* is used for addressing **RQ4** and **RQ5**, applying general principles to specific instances. Answering **RQ4** begins with the theory of a correlation between punch wear and AE signals and tests it with a specific dataset. Tackling **RQ5** employs a general principle that a scalar indicator can encapsulate the tool condition, applying it to the specific findings of **RQ1–RQ4**. This combines bottom-up insights from induction with the top-down structure of deduction, offering a comprehensive view of indirect TCM in sheet-metal stamping processes.

# 4 Generation, description, and preprocessing of the acoustic emission datasets

*Generierung, Beschreibung, und Vorverarbeitung der Körperschall-Datensätze*

Chapter 4 provides an *overview of the datasets* generated and analyzed within this work, as well as the *preprocessing strategy* that has been developed and employed. Specifically, the strategy includes time series segmentation, digital filtering, and feature extraction. These datasets serve as the foundation for this work, and their careful preprocessing is critical for testing the research hypothesis. Furthermore, the heterogeneity of the datasets ensures that the information extracted is robust under various process conditions.

Specifically, Section 4.1 lays out *details on the generated datasets*. The datasets generated and analyzed stem from *laboratory* and *industrial* fineblanking lines. The laboratory datasets are described in Section 4.2, and the industrial datasets in Section 4.3.

## 4.1 Dataset overview

*Überblick über die Datensätze*

In total, 398,047 fineblanking process cycles were analyzed for this work, of which 71,704 process cycles were gathered in a laboratory setting (*laboratory datasets* L1, L2, and L3) and 326,343 in an industrial setting (*industrial datasets* I1, I2, I3, I4, and I5). The laboratory datasets denote datasets generated at the fineblanking plant of the *Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University*, depicted in Figure 4.1.




**Figure 4.1:** Photograph of the fineblanking plant at WZL of RWTH Aachen University.
*Foto der Feinschneidanlage am WZL der RWTH Aachen University.*

The methodology presented within Section 4.2.2 was published in [UNTE23]. Furthermore, the measurement setup documented in Section 4.2.1 and details about dataset L3 were published in [UNTE23]. In Section 4.3.2, the methodology and details on the industrial measurement setup and datasets represent original contributions. The approach of using the reconstruction error of an AEC to identify failed segmentations was derived from the work presented in [NIEM21]. CHATGPT (GPT-4, `https://chatgpt.com/`) was used in this chapter to proofread texts generated by the author and for support with LATEX formatting. DEEPL (`https://www.deepl.com/translator/`) was used for translation purposes.

While the environment is controlled, i.e., the plant is supervised and operated by scientists and technicians with the aim of scientific research, it is an industrial-grade fineblanking plant. Therefore, the conditions are comparable to industrial conditions. All laboratory datasets were generated with the same punch geometry. However, series L1 and L2 were carried out with a different tool than series L3.

The industrial datasets are subject to a non-disclosure agreement (NDA). Specifically, the i) time of the data acquisition, ii) details on the geometry and the tool cavity, and iii) the company name remain unmentioned within this work. However, the results of the data analysis and the provided digital metadata about machine standstills (time of occurrence, reason, and duration) received permission to be presented within this work and are sufficient to answer the research questions involving the industrial datasets, specifically RQ1, RQ2, and RQ5.

Table 4.1 provides an overview of all gathered and analyzed datasets. Column *Material/Thickness* presents the fineblanked material, column *Stroke frequency* presents the (for all stroke series unchanged) stroke frequencies, column *Num. of strokes* the number of *analyzed* strokes, and column *Wear information* presents a binary indicator of whether information on the punch wear was gathered in any shape or form.

For all laboratory datasets, the number of analyzed strokes (as presented in Table 4.1) corresponds to the number of strokes that were carried out within this stroke series. However, this does not hold true for the industrial datasets. The industrial datasets were extracted from a pool of roughly 2,000,000 strokes (AE signal and distance signal) that were gathered over the span of two years and written to cloud storage. To reduce the resulting cost, the secondary process signal (distance signal) was sampled at 220 Hz, and the AE signal was sampled at 400 kHz. The large difference in sampling rates impairs the alignment of the AE signals since the exact start of the relevant process phases can not be inferred solely by utilizing the distance signals. A small percentage of the strokes was omitted (minimum 0.28 % (I5), maximum 0.89 % (I3)) since a time series segmentation was not possible without an unreasonable amount of manual labor. The specifics of this are laid out in Section 4.3.2.

The respective experiments for datasets L1 and L2 were not planned and operated by the author of this work. However, the results of the analyses of these datasets presented within this work were achieved solely by the author. Dataset L3 was planned and operated solely by the author with the assistance of a technician.

**Table 4.1:** Overview of all generated and examined datasets.
*Überblick über alle generierten und untersuchten Datensätze.*

| Name | Material/Thickness | Stroke frequency | Num. of strokes | Wear information |
|---|---|---|---|---|
| L1 | 100Cr6/6 mm | 50 $min^{-1}$ | 30,284 | Yes |
| L2 | 42CrMo4/6 mm | 50 $min^{-1}$ | 970 | No |
| L3 | 58CrV4/5 mm | 50 $min^{-1}$ | 40,450 | Yes |
| I1 | DC01/1.77 mm | 26 $min^{-1}$ | 52,753 | No |
| I2 | DC01/1.77 mm | 26 $min^{-1}$ | 89,590 | No |
| I3 | DC01/1.77 mm | 26 $min^{-1}$ | 108,392 | No |
| I4 | DC01/1.77 mm | 26 $min^{-1}$ | 39,414 | No |
| I5 | DC01/1.77 mm | 26 $min^{-1}$ | 36,194 | No |

While the industrial datasets I1, I2, I3, I4, and I5 were operated by a company, the measurement setup and data acquisition were planned with the participation of the author. The results of the analyses of these datasets presented within this work were achieved solely by the author.

## 4.2 Laboratory datasets

*Labordatensätze*

Section 4.2.1 presents the experimental setup for the laboratory datasets. Section 4.2.2 presents the developed and utilized *time series segmentation approach*. Finally, within Section 4.2.3, the *exploratory feature extraction process* is described.

### 4.2.1 Experimental setup

*Experimenteller Versuchsaufbau*

All strokes within the laboratory datasets were carried out on a servomechanical fineblanking press of type FEINTOOL XFT 2500 speed. The tools used for the laboratory datasets had two cavities laterally offset to each other. Figure 4.2 depicts the part geometry and layout of the scrap web.

**Measurement setup** *Messaufbau*

The sensor positions and the resulting *raw signals* within the tool are presented in Figure 4.3. In this context, the term raw signal refers to the fact that the signals are not segmented into single strokes or single process phases.

The AE signals were acquired with an AE sensor of type KISTLER 8152C0 with a frequency response of 50–400 kHz. The AE signals were subsequently processed with an AE piezotron coupler of type KISTLER 5125c. The distance signals were acquired with a laser triangulation sensor of type MICRO-EPSILON optoNCDT 1900 (ILD1900-25) with a sampling frequency of 10 kHz. The AE signal was digitized with an analog-to-digital converter of type NATIONAL INSTRUMENTS NI-9223 and a sampling frequency of 1 MHz, effectively *oversampling* the signal to reduce noise and improve the resolution of the digital signal. Whenever the fineblanking press was stopped, e.g., due to coil changes or machine errors, the measurements were

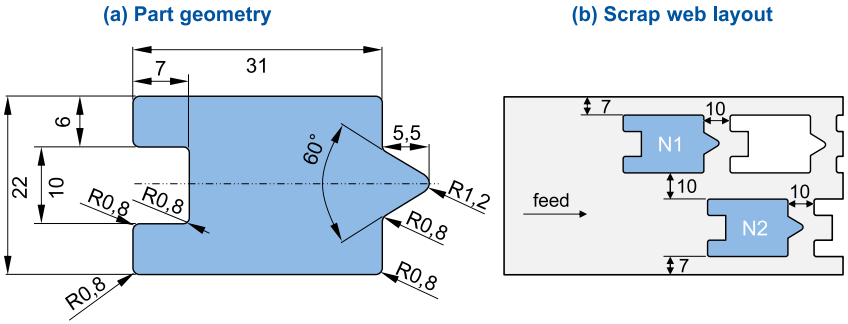



**Figure 4.2:** Depiction of laboratory part geometry and scrap web layout.
*Darstellung der Teile-Geometrie der Labor-Experimente und des Stanzgitter-Layouts.*

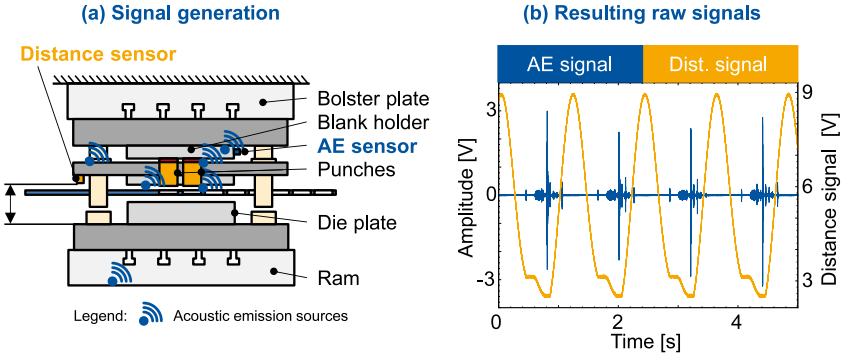



**Figure 4.3:** Depiction of the measurement setup and the resulting signals.
*Darstellung des Messaufbaus und der resultierenden Signale.*

stopped. The AE and distance signals were written with shared timestamps to *tdms* files with NATIONAL INSTRUMENTS NI LabVIEW.

**Digital filtering** *Anwendung digitaler Filter*

In the postprocessing, a BUTTERWORTH band-pass filter with a pass band of 50–400 kHz and order 6 was applied to the AE signals as an IIR filter to prevent aliasing from frequencies outside of the sensors frequency response. Following this, the specific experimental setup for every laboratory dataset is laid out.

**Laboratory dataset L1** *Labordatensatz L1*

Dataset L1 contains 30,284 strokes and was acquired within three workdays. In total, six coils of higher-strength sheet-metal material of grade 100Cr6 (DIN 1.3505, AISI 52100), width 68 mm, and thickness 6 mm were processed. The material came from a single supplier. The punch steel was chosen as a powder metallurgical cold work steel of grade *S390MC*. The coating was chosen as an aluminum chrome-nitride coating. The chosen lubricant was a mineral oil-based chlorine-free lubricant of type HOLIFA VP 1150/200.

After each workday, i.e., after the 12,800th, after the 25,675th, and after the 30,284th stroke, the tool was disassembled, and the punches were analyzed via SEM (see Section 6.3.1). Chipping (localized fracture of the punch material) occurred at the 15,881st stroke on one punch, resulting in a substitution with an identical punch. Furthermore, the screw of the unsubstituted punch broke at the last stroke. After this breakage, the experiment was ended. The punches showed severe wear at the end of the experiment but no further failures.

For every 2,500th stroke, scrap web segments were sampled from the waste container for later analysis. The results of the analyses of the sheared surfaces of the scrap web segments and the SEM analyses are presented in Section 6.3.1. After analysis, the punch positions in the tool were switched to account for possible differences in the wear evolution of the punches due to the cavity. The used blank- and counter holder forces and results of the mechanical tests for the material are presented in Table A.1 in Appendix A.1.

**Laboratory dataset L2** *Labordatensatz L2*

Dataset L2 contains 970 strokes and was acquired within one workday. Only a fraction of one coil of higher-strength sheet-metal material of grade 42CrMo4 (DIN 1.7225, AISI 4140),

width 68 mm, and thickness 6 mm was processed. The remaining process setup was chosen to be identical to dataset L1.

After 970 strokes, a punch breakage occurred (see Figure 4.4). Since the punch breakage occurred unusually early into the series and was suspected to have occurred due to alignment issues within the tool and resulting friction between punch and punch guidance, it was decided to end the experiment. A detailed discussion of the punch breakage can be found in Section 6.3.2. Table A.1 in Appendix A.1 presents the blank- and counter-holder forces used, along with the results of the mechanical tests for the material.

**Laboratory dataset L3** *Labordatensatz L3*

Dataset L3 contains 40,450 strokes and was acquired within three workdays. Eight coils of higher-strength sheet-metal material 58CrV4 (DIN 1.8161), width 68 mm, and thickness 5 mm were processed. The punch steel was chosen as a powder metallurgical cold work steel of grade *K490* with a titanium aluminum carbon-nitride coating. A mineral oil-based chlorine-free lubricant of type WISURA FMO 5020 was chosen. During the experiment, no punch breakages or shearing edge failures occurred. The experiment was ended to not endanger tool components since a high amount of metal shavings was observed within the tool.

For every 200th stroke, scrap web segments were sampled from the waste container for later analysis. The results of this analysis are presented within Chapter 6. The tool was not disassembled during the stroke series L3, hence leaving the punch positions unchanged in contrast to L1. In Appendix A.1, Table A.1 showcases the blank holder and counter holder forces for the experiment, as well as the material's mechanical test results.

### 4.2.2 Time series segmentation
*Zeitreihensegmentierung*

The segmentation of the resulting AE signals was done to reduce the amount of redundant information contained in the signal, especially to minimize the risk of spurious correlations in later analysis steps. Since this work seeks to identify the conditions of the tribological system

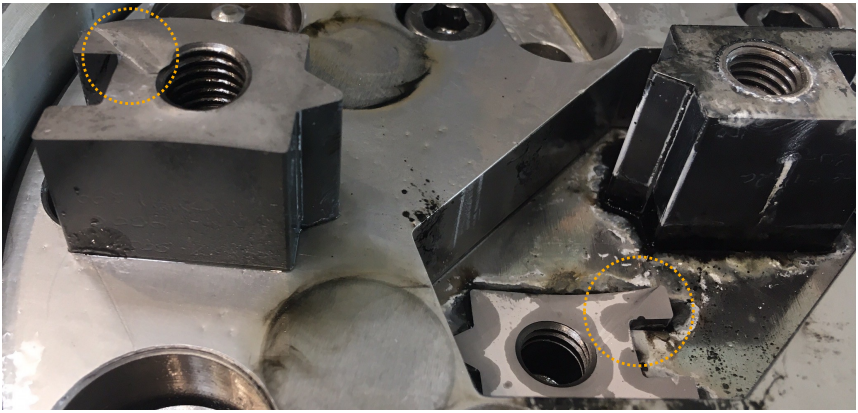


**Figure 4.4:** Fotographic documentation of the punch breakage in stroke series L2.
*Fotografische Dokumentation des Stempelbruchs in der Hubserie L2.*

between the punch and the sheet metal, only AE signals emitted during process phases where a relative motion between the punch and the sheet metal occurs are of interest. Thus, a *segmentation pipeline* (see Figure 4.6) was developed.

The pipeline has two main objectives:

(i) *Extract* the AE signal emitted during the *shearing and stripping phase* of the process (see Section 2.1) from the raw AE signals.

(ii) Ensure that the signals are *aligned*, i.e., the extracted shearing and stripping segments of strokes $i$ and $k$ with $i \neq k$ should approximately describe a similar state of the process phase at each point in time.

**Objective 1: AE signal segmentation** *Ziel 1: Segmentierung des AE Signals*

Objective 1 was achieved by developing a *threshold-based* segmentation approach. In a *first step* (*rough segmentation*), the data load was reduced by extracting strokes, i.e., one closing and the subsequent opening cycle of the tool, from the raw signal. To this end, *the top dead center* (TDC) was identified as a local minimum in the distance signal that marks the end of the shearing phase (see [UNTE21] and [UNTE23]). Since the distance signal was subject to severe noise, it was smoothed with a rolling median filter. The shearing and stripping phases were then identified based on the fact that their position relative to the TDC is approximately constant (*fine segmentation*). To enable an efficient search for the TDC (the largest written tdms files had a size of 144 gigabytes, representing approximately 100 min of production time or 5,000 strokes), the search area was limited to an upper and a lower boundary of the distance signal. Figure 4.6 schematically visualizes the segmentation processes. The TDC (i) was searched within an upper and a lower boundary of the distance signal (ii). The fine segmentation of a stroke was then inferred from the position of the TDC. This pipeline was implemented in PYTHON with the packages NumPy, *pandas*, and SciPy.

**Objective 2: Segment alignment** *Ziel 2: Ausrichtung der Segmente*

While all strokes were i) identified and ii) segmented with this approach, *misalignment* due to the different sampling rates of the AE and distance signals occurred. Misalignment in this context means that for each stroke $i$, the shearing and stripping phases of all other strokes $k$ with $i \neq k$ were shifted in time relative to stroke $i$. This shift was equal for both process phases and varied stroke-wise [UNTE23]. Figure 4.7 depicts the misalignment of the stripping phases and the subsequent alignment by using MASS (see Section 2.3).

For the purpose of visualization, Figure 4.7 shows transformed stripping signals, specifically, the *root mean square* (RMS) of an excerpt of the stripping segments, calculated with a sliding window of size 500. The resulting visualization depicts the initial bursts in the AE signal

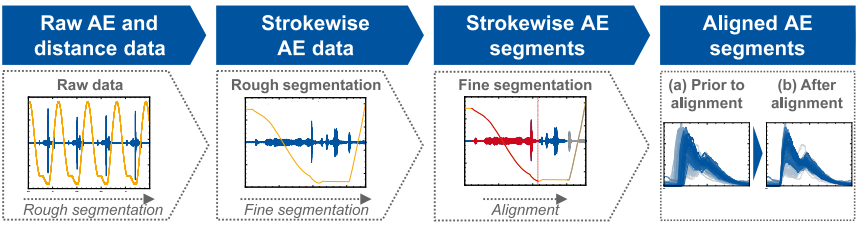



**Figure 4.5:** Developed segmentation pipeline for the laboratory stroke series (a similar figure was published in [UNTE23]).
*Entwicklung einer Segmentierungspipeline für die Hubserie aus der Laborumgebung.*

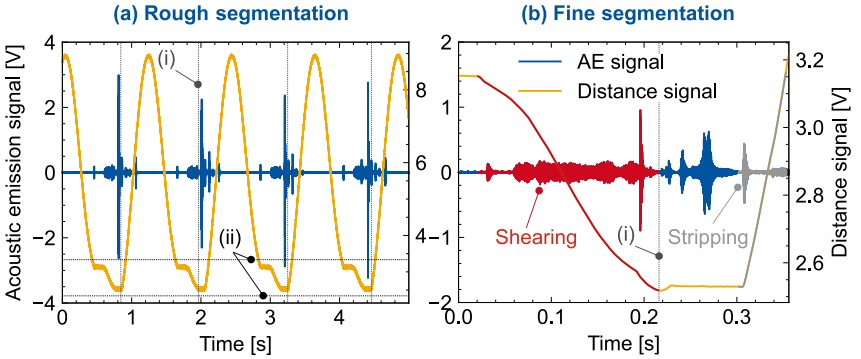



**Figure 4.6:** (a) Threshold-based rough segmentation and subsequent (b) fine segmentation.
*(a) Schwellenwertbasierte Grobsegmentierung und anschließende (b) Feinsegmentierung.*

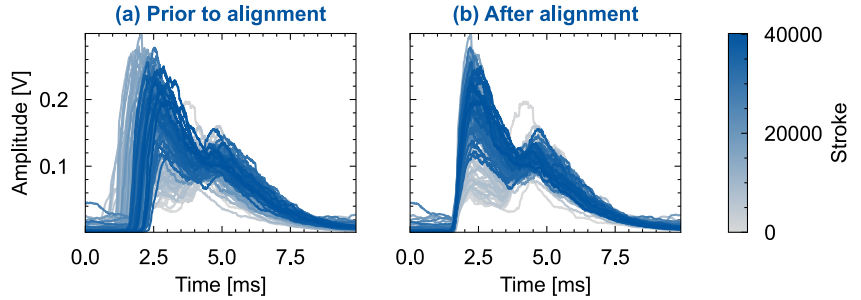



**Figure 4.7:** Moving RMS (window size 500) of the initial AE burst during stripping for every 300th stroke of dataset L3 (a) prior to and (b) after alignment.
*Gleitendes Quadratisches Mittel (Fenstergröße 500) des anfänglichen AE-Bursts während des Abstreifens für jeden 300. Hub des Datensatzes L3 (a) vor und (b) nach der Ausrichtung.*

within the stripping segments that result from surpassing the clamping force between the scrab web sheared surface and the lateral surface of the punch when the stripping process is initialized [VOIG21].

An implementation of MASS from the PYTHON library stumpy was used to quantify the shifts within the sequence of stripping phases to a reference stroke $x_Q(t)$ since the initial burst in the stripping segments was reliably detectable. To this end, the starting index of the sequences $x_T(t)[i : i + M - 1]$ with the lowest distance in the resulting distance profiles to $x_Q(t)$ was extracted for every stroke. This information was then used to extract the aligned shearing and stripping segments from the rough segmentation.

### 4.2.3 Exploratory feature extraction
*Explorative Extraktion von Features*

The first step sought to mitigate the curse of dimensionality (see Section 2.2) and employ feature extraction as a form of dimensionality reduction. To this end, the PYTHON library *Time series feature extraction library* (TSFEL) was used. The library offers computationally efficient calculation of comprehensive and salient features for time series from the *spectral*, *temporal*, and *statistical* domain. A full list of features can be accessed in [BARA20]. The library offers the computation of features within user-defined window ranges. However, this work used the full signal length as a window size to reduce the resulting data load. All available features were extracted from the stripping and shearing segments. The resulting datasets were used within Chapter 7.

## 4.3 Industrial datasets
*Industriedatensätze*

Section 4.3.1 lays out the experimental setup for the generation of the industrial datasets. The developed segmentation pipeline for the industrial datasets is presented in Section 4.3.2.

### 4.3.1 Experimental setup
*Experimenteller Versuchsaufbau*

While the experimental setup for the laboratory datasets was laid out in detail in Section 4.2.1, the documentation for the industrial datasets is subject to an NDA. However, the author has permission to publish the measurement setup and the results of his analyses, including the shape of the AE signal, trends in the data, and corresponding documented events. This information is sufficient to answer the research questions involving industrial datasets, i.e., RQ1, RQ2, and RQ5.

**Measurement setup** *Messaufbau*

Similar to the experimental setup for the laboratory datasets documented in Section 4.2.1, the AE signals were acquired with an AE sensor of type KISTLER 8152C0, mounted in a similar position as depicted in Figure 4.3, subsequently processed with an AE piezotron coupler of type KISTLER 5125C, and digitized with an analog-to-digital converter of type NATIONAL INSTRUMENTS NI-9223. The signals were effectively oversampled with a sampling frequency of 1 MHz. The signals were downsampled with SciPy's *resample* function from the *signal* module to a sampling frequency of 400 kHz, yielding a NYQUIST frequency of 200 kHz to i) *handle the data efficiently and in time* and ii) *reduce resulting cloud storage cost*.

The datasets were gathered over the span of 2 years. Continuous AE and distance measurements with high sampling frequencies were unviable because of the potential resulting data load. Manually triggering the measurements was also unviable in an industrial setting since workers are expected to focus on their duties and obligations during their shifts. Thus, the company SENSEERING that led the data acquisition implemented a trigger algorithm based on a thresholding approach similar to the one laid out in Section 4.3.2. To this end, a distance sensor of type LEUZE ODSL 8/V66-500-S12 was used with a sampling frequency of 220 Hz. The implications of the relatively low sampling rate compared to the AE signal's sampling rate are discussed in Section 4.3.2.

Acoustic emission and distance data were only written whenever the distance signal fell below and subsequently surpassed pre-defined thresholds that were explored and defined by SENSEERING, the fineblanking company, and the author of this work. To further reduce the data load, the involved parties decided only to write AE data containing the shearing phase of the process and thus not acquire AE data for the stripping phase of the process.

**Digital filtering** *Anwendung digitaler Filter*

The AE data was filtered with a high-pass BUTTERWORTH IIR filter of order 6 and a cutoff frequency of 50 kHz, representing the lower boundary of the sensor's frequency response. Since the data was digitally downsampled in postprocessing, all contained frequencies within the signal fall into the sensor's frequency response.

**General context of the industrial data** *Grundsätzlicher Kontext der industriellen Datensätze*

All industrial stroke series produced an identical safety-critical part for the automotive industry. The fineblanked material grade was DC01 (DIN 1.0330, AISI 1008) for all industrial stroke series with a sheet metal thickness of 1.77 mm. The author does not have information on the specific number of suppliers for the material. However, the number is larger than one. Hence, according to industry experts, more substantial fluctuations in material properties are expected than in the case of a single supplier. However, the effect of the number of suppliers can not be examined due to the lack of metadata.

As presented in Table 4.1, the industrial datasets do not contain information about the wear state of the punch. Tools are dismounted and maintained based on one of two criteria. The first is when a *pre-defined stroke threshold* is reached. The second is when a worker observes either *product impairment or an abnormal frequency of machine standstills*, which industry experts associate with tool wear phenomena. The company employs two tools for the same product to ensure uninterrupted production. While Tool 1 is under maintenance, Tool 2 is mounted and used. During maintenance, punches are reground, and the coating is renewed. If this proves insufficient, the punches are replaced.

The company keeps a machine logbook that documents i) machine standstill times, ii) machine standstill reasons, and iii) machine standstill durations. Based on this information, *two key assumptions* were made prior to the analysis of the industrial datasets:

(i) The datasets differ substantially in the number of strokes. It is likely that *problems with the tool occurred within the shorter datasets*. These problems *may* relate to the punch wear. Since the punches were not examined in detail, this can not be ascertained with certainty ex-post.

(ii) Machine standstills, i.e., downtime reasons and stand still times, are inherently biased and dependent on the worker. However, standstills with logged standstill reason *Tool may* relate to the punch wear.

While the assumptions are made under uncertainty, the datasets represent the most extensive evaluation of AE signal evolution known to the author under realistic industrial conditions. The following chapter will outline the developed segmentation pipeline for the industrial datasets.

### 4.3.2 Time series segmentation
*Zeitreihensegmentierung*

Figure 4.8 visualizes the AE signal of the 1st and the last (108,392nd) stroke of the industrial dataset I3. Clearly, the acquired AE signal of the 108,392nd stroke is substantially shorter than the AE signal of the 1st stroke, representing the variation within the AE signal lengths within the industrial datasets. This variation is, on the one hand, caused by the low sampling rate of the distance sensor and, on the other hand, by signal disturbances caused by the industrial conditions (e.g., lubricant smearing) in said signal.

Furthermore, it is apparent that the *shape of the signal changes*, i.e., i) bursts appear within the AE signal of the 108,392nd stroke that can not be observed in the AE signal of the 1st stroke and ii) the amplitude substantially increased, potentially caused by an increase of wear-related friction increase. Since these changes occurred unpredictably, at least in the exploratory stage of the data analysis, a segmentation purely based on the AE signal required more effort than in the laboratory case. A pipeline was developed with the following steps that were carried out for every industrial dataset:

S1 Identify a *reference shearing segment* $x_Q(t)$ within the dataset based on the known tool design and known process steps. To this end, the company provided an analog voltage signal after the data acquisition phase for a single stroke. This analog signal was used to identify the shearing phase for reference strokes in a pristine tool condition.

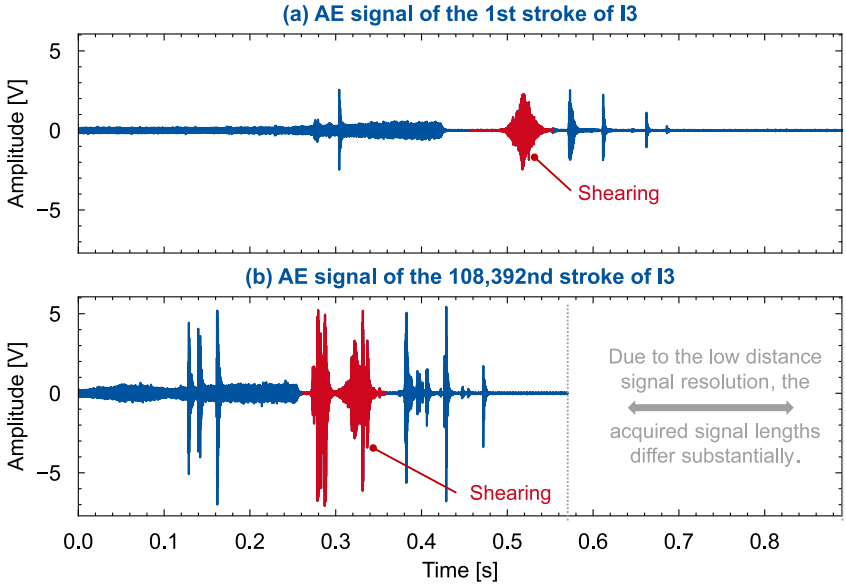



**Figure 4.8:** Visualization of the acquired AE signal of (a) the 1st and (b) the 108,392nd stroke of the dataset I3.
*Visualisierung des erfassten AE-Signals von (a) dem 1. und (b) dem 108.392. Hub des Datensatzes I3.*

S2 Use MASS to find the most similar subsequence to $x_Q(t)$ in $x_T(t)$ for every single stroke, i.e., the minimum of the distance profile $D$.

S3 Manually screen the first 1,000 identified subsequences from step S2 and remove failed segmentations (Figure 4.9 shows two examples of failed segmentations). *Failed segmentations were removed from the dataset*.

S4 Train an AEC on the successful segmentations that step S3 yielded, reconstruct all subsequences contained in the dataset, and evaluate the MSE of the reconstructions. Manually screen the 5,000 subsequences with the highest MSE. *Failed segmentations were removed from the dataset*.

S5 Manually screen all subsequences. *Failed segmentations were removed from the dataset*.

Failed segmentations were *qualitatively* defined. The criteria were chosen as follows:

C1 The characteristic shape that results from plastomechanic and tribological mechanisms [VOIG21] *must be present in the extracted segment* (see Figure 4.9, (i) and (ii)).

C2 Said characteristic shape has to be present *approximately in the same time frame*, i.e., between 0.045 s and 0.080 s. However, slight shifts of up to 0.01 s were accepted.

The chosen criteria represent a trade-off between the manual effort required to clean the dataset and the resulting dataset quality. A near-perfect segment alignment, as depicted for the laboratory datasets in Figure 4.7, was not achievable. However, occurring shifts within the resulting segments were minimal and accepted for the exploratory analysis carried out within this work.

Steps S3 and S4 were carried out to reduce the amount of failed segmentations that need to be manually detected since the pipeline was developed under the assumption that the final datasets may still contain a small amount of failed segmentations due to human error. However, a high reconstruction error does not necessarily imply that the subsequence represents a failed segmentation. The utilized AEC architecture for this task was published in [NIEM21].

*Convolutional* layers were chosen due to their *shift invariance* (or translation invariance) since they "force the replication of weight configurations across space" [LECU95]. If certain learned features (the AEC essentially extracts suitable features from the signals) are not present in the signal that were learned as salient by the AEC, signal reconstruction should yield a high error. However, successful segmentations also differed substantially in their signal characteristics, as can be seen in Figure 4.9. The changes that the acquired AE signals underwent with an accumulation of process cycles made a further manual screening step necessary. This manual screening was done based on the plots of the resulting subsequences.

Table 4.2 presents the resulting number of strokes and the corresponding percentage of the dataset size for the datasets I1, I2, I3, I4, and I5. As can be inferred, a comparable percentage of the total dataset sizes was removed due to failed segmentation. Since the percentages remain under 1 % of the total dataset sizes, and this work seeks to identify trends in the evolution of the AE signals of the industrial datasets, omitting the failed segmentations was accepted.

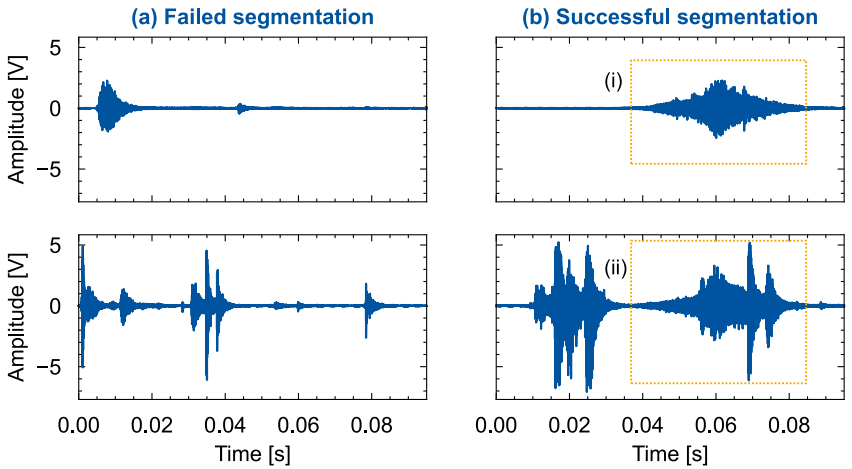



**Figure 4.9:** Visualization of (a) failed vs. (b) successful segmentations within dataset I3. The characteristic shape within the shearing phase is annotated with (i) (1st stroke from I3) and (ii) (108,392nd stroke from I3). The remainder of the signal resulted from the initial contact between the punch and the sheet metal; the increased amplitude may be caused by a coating failure and subsequent adhesion mechanisms that lead to increased friction between the punch and the scrap web. The superimpositions in the signal segment (ii), compared to (i), could also be wear-induced and result from elevated friction between the punch and the scrap web.
*Visualisierung (a) fehlgeschlagener vs. (b) erfolgreicher Segmentierungen im Datensatz I3. Die charakteristische Form innerhalb der Schneidphase ist mit (i) (1. Hub von I3) und (ii) (108.392. Hub von I3) gekennzeichnet. Der Rest des Signals resultierte aus dem initialen Kontakt des Stempels mit dem Blech; die erhöhte Amplitude könnte auf ein Versagen der Beschichtung zurückzuführen sein, welches durch Adhäsions-Mechanismen zu einer höheren Reibung zwischen Stempel und Stanzgitter führte. Die Überlagerungen im Signalsegment (ii) im Vergleich zu (i) könnten ebenfalls verschleißinduziert und durch eine erhöte Reibung zwischen Stanzgitter und Stempel hervorgerufen worden sein.*

**Table 4.2:** Overview of the removed number (num.) of strokes per dataset and corresponding percentage (Pctg.) of the dataset size.
*Überblick über die Anzahl und den Anteil der entfernten Hübe pro Datensatz.*

| Dataset name | Total num. of strokes | Removed strokes | Pctg. of strokes removed (%) |
|---|---|---|---|
| I1 | 53,086 | 333 | 0.63 |
| I2 | 90,213 | 623 | 0.69 |
| I3 | 109,363 | 971 | 0.89 |
| I4 | 39,688 | 274 | 0.69 |
| I5 | 36,297 | 103 | 0.28 |

## 4.4 Feature extraction

*Feature Extraktion*

Following, the feature extraction approach is laid out. Specifically, the approach for the PSD estimation as well as the calculation of the features *mean frequency* and *total, absolute, and relative bandpower*. Finally, the feature extraction process with the PYTHON library TSFEL is briefly elaborated.

**Power spectral density estimation** *Schätzung der spektralen Leistungsdichte*

Similar to the feature extraction process for the laboratory datasets documented in Section 4.2.3, the PSD was estimated for every stroke of the industrial datasets utilizing WELCH's method (see Section 2.3). The resulting evolution of the frequency components is presented within Section 5.2.

**Mean frequency** *Mittlere Frequenz*

SHANBHAG et al. proposed the mean frequency estimate feature from MATLAB as a suitable feature to detect the onset and severity of galling. With the onset of galling between the die and the parts, the authors observed a steep decline in the feature and state that the exhibited galling was potentially causal for this steep decline [SHAN18].

The MATLAB documentation does not explicitly describe how the mean frequency feature is estimated [MATH23]. However, it links to [PHIN12], where the authors state that the mean frequency feature, mainly used in surface electromyography research, is calculated as the product of the PS and the *frequencies*, averaged by the total sum of the PS. For a discrete-time signal, this can be formalized to

$$\mathrm{MNF} = \frac{\sum_{k=0}^{k_{\mathrm{Nyquist}}} f_k P_{XX}(k)}{\sum_{k=0}^{k_{\mathrm{Nyquist}}} P_{XX}(k)}, \tag{45}$$

where $k_{\mathrm{Nyquist}}$ denotes the bin corresponding to the NYQUIST frequency, $f_k$ denotes the frequency value at frequency bin $k$, and $P_{XX}(k)$ denotes the power at frequency bin $k$. The author has implemented a PYTHON function to calculate the mean frequency feature for each stroke of the industrial datasets from the PS estimated with WELCH's method. The results of the mean frequency analyses are reported within Section 5.2.2.

**Total, absolute, and relative bandpower** *Gesamt-, absolute und relative Bandleistung*

Information about signal characteristics within specific frequency bands can be usefully illuminated by the concept of *bandpower*, often known as *spectral bandpower* [DRES04]. In the context of a discrete signal, the *total bandpower* (TBP) is computed as the *sum of the PS values across all frequency bins*:

$$\mathrm{TBP} = \sum_{k=0}^{k_{\mathrm{Nyquist}}} P_{XX}(k), \tag{46}$$

where $P_{XX}(k)$ is the PS value at frequency bin $k$. Next, the *absolute bandpower* (ABP) is defined as the *sum of the PS values within a defined frequency band*:

$$\mathrm{ABP} = \sum_{k=k_{\mathrm{min}}}^{k_{\mathrm{max}}} P_{XX}(k), \tag{47}$$

where $k_{\text{min}}$ and $k_{\text{max}}$ indicate the respective lower and upper bounds (in terms of frequency bin indices) of the desired frequency band. The *relative bandpower* (RBP) is then computed as the *ratio of the ABP within a specific frequency range to the TBP across all frequency bins*:

$$\text{RBP} = \frac{\sum_{k=k_{\text{min}}}^{k_{\text{max}}} P_{XX}(k)}{\sum_{k=0}^{k_{\text{Nyquist}}} P_{XX}(k)} \tag{48}$$

where $k_{\text{max}} \leq k_{\text{Nyquist}}$. These features, TBP, ABP, and RBP, are particularly useful when signals have distinct behaviors or characteristics at different frequency ranges, see Section 2.4. Therefore, evaluating the contribution of certain frequency bands is a promising approach. As ABP, TBP, and RBP are easy-to-interpret scalar features, they are particularly suited for exploratory data analysis.

A PYTHON function has been implemented to calculate TBP, ABP, and RBP by the author using WELCH's method. The ABP feature was used for the works documented within Chapter 6.3.2, and ABP, TBP, and RBP were used for the works documented in Chapter 8.

**TSFEL feature extraction** *Feature Extraction mittels TSFEL*

Analogous to the feature extraction process for the laboratory datasets documented in Section 4.2.3, TSFEL was used for the industrial datasets. All available features were extracted for the extracted segments with the full signal length as the window size. The TSFEL feature sets were used within Chapter 5, Chapter 7, and Chapter 8.

## 4.5 Intermediate conclusion and unresolved challenges

*Zwischenfazit und offene Herausforderungen*

Research question **RQ1** was formulated as **"What is the process for designing a domain-specific preprocessing pipeline to isolate relevant process phases from acoustic emission signals in the context of punch wear monitoring in sheet-metal stamping?"**. To address this, AE data was generated in both a laboratory and an industrial fineblanking setting and subsequently processed. A distance signal was also acquired to map the AE data to specific process phases. Since relative motion between the punch and the sheet metal occurs only during the shearing and stripping phases, the AE signal emitted during these phases needed to be isolated using a suitable algorithmic approach.

The segmentation of the AE signals in the laboratory setting was accomplished through a pipeline consisting of i) a *rough segmentation* to identify the TDC within the distance signal as a local minimum, ii) a *fine segmentation* to extract shearing and stripping phases from the rough segmentation based on their consistent relative positions to the TDC, and iii) an *alignment step* with MASS, a time series similarity search technique within data mining. The alignment step searched for the initial burst within the stripping segments and aligned shearing and stripping segments globally for each dataset by finding the most similar subsequences to a reference segment. This alignment was crucial to compensate for the substantial mismatch in sampling rates between the AE signal (1 MHz) and the distance signal (10 kHz), which leads to potential timing discrepancies in the resulting segments.

In the industrial setting, data was preprocessed to reduce the load, capturing only brief timeframes around the shearing phase. Although the pipeline could be automated in the laboratory setting, the industrial setting required more manual effort because for 0.28–0.89% of the strokes the extracted segments did not represent the shearing phase. An algorithmic error detection approach was developed, using an autoencoder trained on a

subset of successful segmentations. This autoencoder effectively learned the characteristics of successful segmentations, allowing it to distinguish between successful and failed ones by evaluating the reconstruction error of all segments. The segments were then sorted by their reconstruction error and the sorted segments were manually screened for failed segmentations. Failed segmentations were manually omitted. Sorting the segments by their reconstruction error allowed to handle a large amount of data by minimizing manual efforts, since most failed segmentations tended to cluster around the maximum reconstruction error.

Failed segmentations in the industrial setting were partly due to the substantially lower resolution of the distance signal (400 kHz AE signal vs. 220 Hz distance signal). Additionally, severe disturbances in the distance signal, potentially caused by industrial conditions such as lubricant residues, and the lower quality of the distance sensor compared to that in the laboratory setting, further complicated data analysis. Moreover, the industrial tool represents a progression die that combines multiple process steps, resulting in a more complex AE signal. Thus, segmentation based solely on the AE signal and minimal distance signal information remains challenging and is a crucial area for future research.

The research question RQ1 is considered answered with the developed pipeline. For a broader application in industrial settings, employing a higher-quality distance sensor is recommended, or identifying the TDC by using machine signals, such as an analog voltage signal when the TDC is reached. While these methods provide greater accuracy, they are costly and could potentially be replaced by time series segmentation algorithms based on machine learning that leverage the existing data to autonomously identify segments of interest, compensating for signal quality and operational condition variabilities.

# 5 Exploratory analysis of industrial acoustic emission datasets

*Explorative Analyse industrieller Körperschall-Datensätze*

Chapter 5 aims to conduct *exploratory analyses on the industrial datasets* I1, I2, I3, I4, and I5 with minimal contextual information. The objective in the context of testing the research hypothesis is to assess whether such high-level, qualitative analyses are sufficient for identifying emerging trends in the acquired and preprocessed AE data.

Section 5.1 conveys the *methodology* and the *results of the qualitative analysis of the resulting low-dimensional embeddings of the AE segments* for the industrial datasets that were generated with the segmentation pipeline documented in Section 4.3.2. The *methodology* and the *results* of the *qualitative evolution of domain-specific AE features* are documented in Section 5.2. Section 5.3 concludes Chapter 5.

## 5.1 Structural evolution of low-dimensional acoustic emission signal embeddings

*Strukturelle Entwicklung niedrig-dimensionaler Körperschallsignal-Embeddings*

Section 5.1 documents the qualitative analysis of the low-dimensional embeddings that were generated from the shearing segments of the industrial fineblanking stroke series. The methodology is conveyed in Section 5.1.1. Finally, the results are presented in Section 5.1.2. Section 5.1.3 concludes the qualitative analysis of the structural evolution of the low-dimensional AE signal embeddings.

### 5.1.1 Methodology

*Methodik*

In the first step, the dimensionality of the industrial datasets was reduced through feature extraction with the library TSFEL. To this end, the full signal length was used as a window size to reduce computational load in subsequent analyses.

The resulting descriptive representations were filtered for features with variance 0 using sklearn's *sklearn.feature_selection.VarianceThreshold* function. The remaining features were then standardized using sklearn's *sklearn.preprocessing.StandardScaler* function. A *StandardScaler* object was fit to each dataset, and the data was transformed subsequently.

Projection of the dataset-wise z-scores into 3-dimensional spaces was then carried out using UMAP and PCA, applied individually to each dataset. For UMAP embeddings, the procedure adhered to the standard parameters as documented in [UMAP23].

Visualization of the resulting 3-dimensional embeddings was achieved with *matplotlib*. A colormap was employed to illustrate the relationship between the spatial positioning of data points and the temporal (i.e., amount of prior process cycles) position of the stroke in the ordered stroke series. For a given datapoint in the coordinate system, the color of the datapoint indicates the position of the respective stroke that it represents. The full pipeline is depicted schematically in Figure 5.1.

---

The methodology presented within Section 5.1.1 was published in [UNTE21]. However, the i) analyzed datasets and ii) results presented are original and yet unpublished. The remainder of Chapter 5 also represents unpublished material. CHATGPT (GPT-4, `https://chatgpt.com/`) was used in this chapter to proofread texts generated by the author and for support with LATEX formatting. DEEPL (`https://www.deepl.com/translator/`) was used for translation purposes.

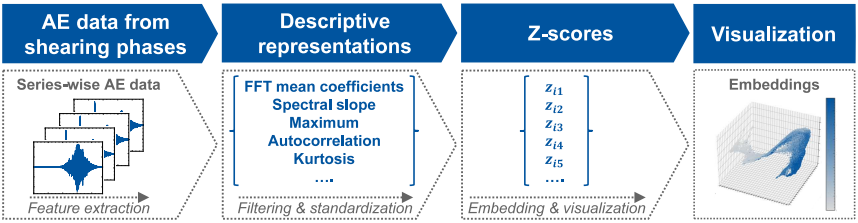



**Figure 5.1:** Developed pipeline for the qualitative analysis of the structural evolution of low-dimensional embeddings for the industrial shearing segments.
*Entwickelte Pipeline zur qualitativen Analyse der strukturellen Entwicklung von niedrigdimensionalen Embeddings für die Schneidsegmente aus der industriellen Umgebung.*

### 5.1.2 Results
*Ergebnisse*

In [UNTE21], the author of this work published the results of a qualitative analysis similar to the analysis carried out in this work. The results were generated from an industrial dataset as well. However, i) the dataset size was limited to 14,030 strokes, ii) at the end of the series, excessive punch and die wear were provoked through lubricant reduction, and iii) the experiment was supervised and closely monitored by scientists.

Following, the results of the qualitative analysis of AE signals generated by 326,433 fineblanking strokes are presented. The findings are summarized in Table 5.1. The data generation was not monitored by scientists. Attaching and removing the AE sensor when the tools went into maintenance was carried out solely by the producing company.

**Qualitative evaluation of the embeddings of I1** *Qualitative Evaluation der Embeddings von I1*

The UMAP embedding of stroke series I1 depicted in Figure 5.2 shows a clear dependency of the positioning of a stroke embedding in a 3-dimensional space on the number of prior process cycles. Specifically, the colormap allows for the identification of three phases ((i), (ii), and (iii)) and transitions between these.

The PCA embedding of stroke series I1 (see Figure 5.3) reinforces this finding. The number of prior process cycles substantially determines the position of the stroke embedding, suggesting that the extracted descriptive representations of the AE shearing segments undergo a systematic variation over the punch life.

**Qualitative evaluation of the embeddings of I2** *Qualitative Evaluation der Embeddings von I2*

Figure 5.4 illustrates the UMAP embedding of stroke series I2. Similar to series I1, it becomes apparent that a clear distinction between three phases can be made. Interestingly, the first phase is represented by two distinct data clusters.

Similar to the PCA embedding of series I1, the PCA embedding of series I2 exhibited in Figure 5.5 corroborates this finding. While the projection does not yield singular clusters, a clear distinction between phases (i) and phases (ii) and (iii) is displayed.

**Qualitative evaluation of the embeddings of I3** *Qualitative Evaluation der Embeddings von I3*

Stroke series I3 represents the longest stroke series recorded. The findings from these embeddings strengthen the argument made from the qualitative analysis of series I1 and I2. Both for the UMAP (see Figure 5.6) and PCA (see Figure 5.7) embedding, three distinct phases and transitions between these phases can be identified. However, it is noteworthy that the evolution within the UMAP embedding is not as clearly visible as for the PCA embedding and the embeddings of series I1 and I2. Furthermore, the formation of singular clusters can be observed within the UMAP embedding in Figure 5.6.

**Qualitative evaluation of the embeddings of I4** *Qualitative Evaluation der Embeddings von I4*

Similar to the UMAP embedding of stroke series I2, Figure 5.8 reveals that the first strokes of the stroke series I4 are represented within a singular data cluster (i). While embedded strokes for stroke series I1, I2, and I3 show a rather strong connection, the embeddings of stroke series I4 are spread farther out. Additionally, a singular data cluster from the last approximately 100 strokes formed.

This finding is emphasized by the PCA embedding of stroke series I4 depicted in Figure 5.9, which also depicts a singular data cluster (iii) for the last strokes of the stroke series.

**Qualitative evaluation of the embeddings of I5** *Qualitative Evaluation der Embeddings von I5*

Stroke series I5 represents the shortest stroke series recorded. While all the remaining stroke series depict a clear path through the embedding, determined by the number of process cycles completed, the UMAP embedding of stroke series I5 reveals a large number of singular clusters for the 1st phase of the stroke series. Additionally, no clear path can be identified in the data that depends on the number of prior process cycles. Instead, it appears that neighboring strokes are seen as similar by UMAP within different process state clusters. However, the similarity of these process state clusters to each other can not clearly be explained by the number of completed process cycles.

Figure 5.11 underlines this. While patterns of subsequent process form in the PCA embedding, a clear path that depends on the position of a stroke in the stroke series can not be identified.

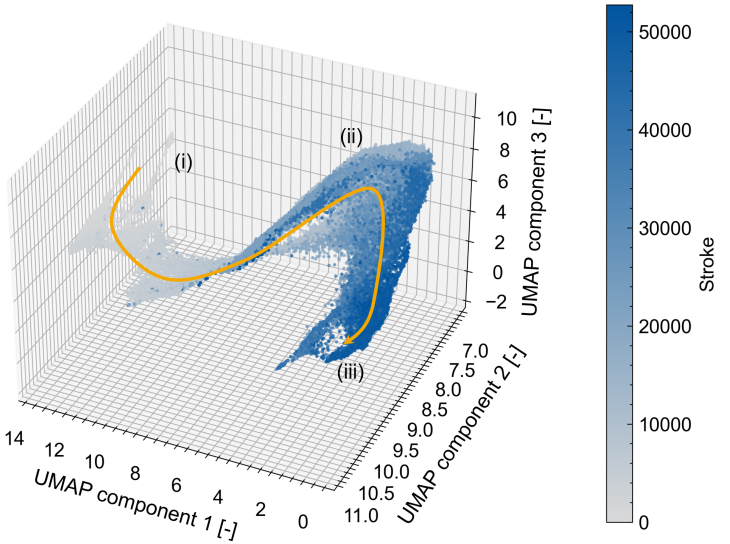



**Figure 5.2:** 3d UMAP embedding of the I1 shearing data.
*3d UMAP-Embedding der Schneidsegmente aus I1.*

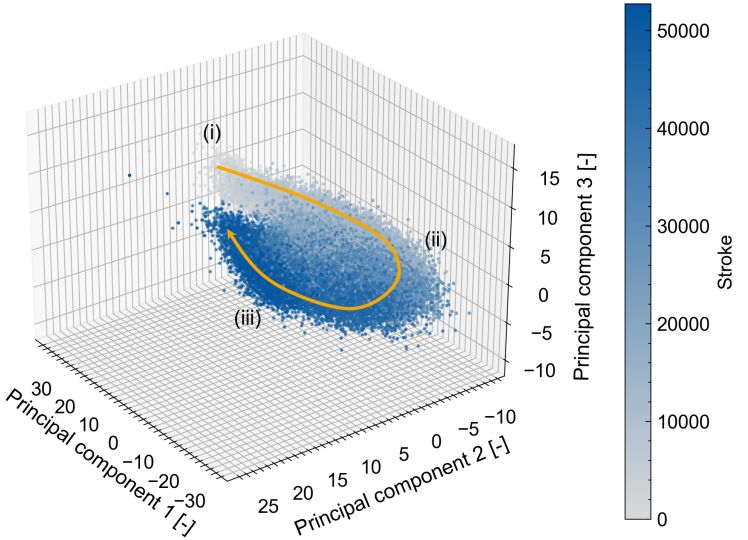



**Figure 5.3:** 3d PCA embedding of the I1 shearing data.
*3d PCA-Embedding der Schneidsegmente aus I1.*

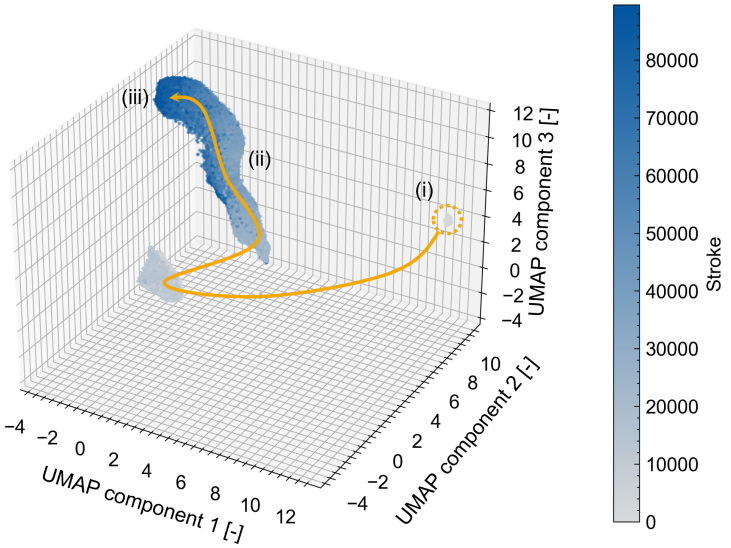



**Figure 5.4:** 3d UMAP embedding of the I2 shearing data.
*3d UMAP-Embedding der Schneidsegmente aus I2.*

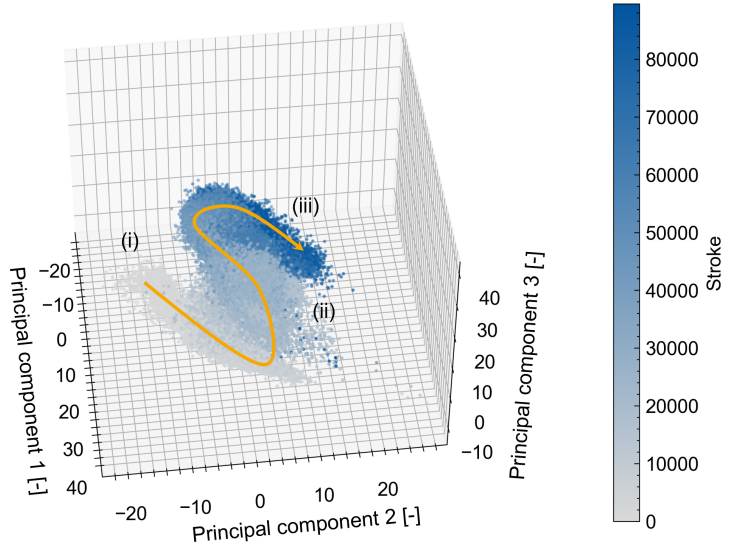



**Figure 5.5:** 3d PCA embedding of the I2 shearing data.
*3d PCA-Embedding der Schneidsegmente aus I2.*

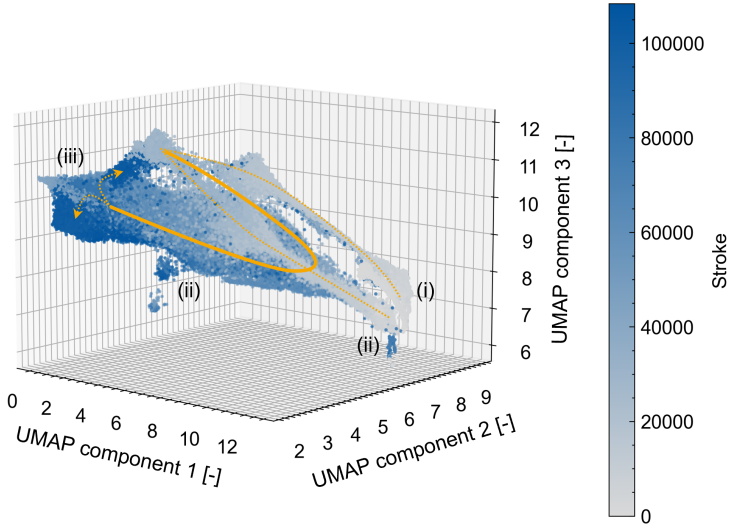



**Figure 5.6:** 3d UMAP embedding of the I3 shearing data.
*3d UMAP-Embedding der Schneidsegmente aus I3.*

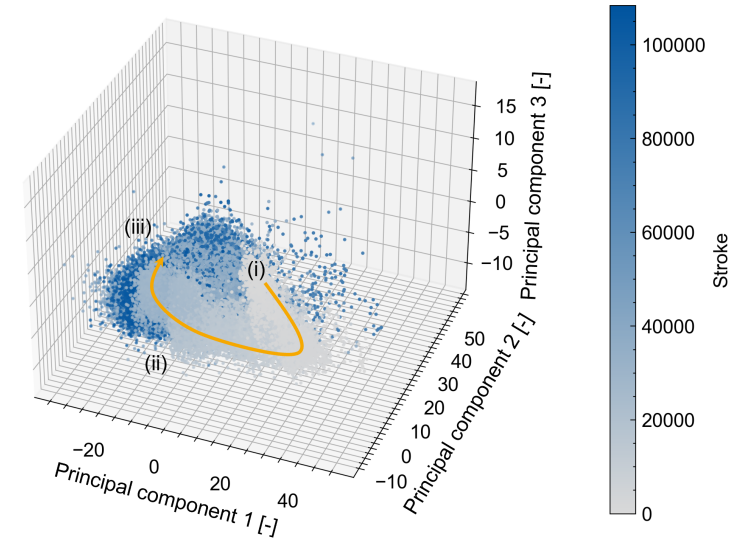



**Figure 5.7:** 3d PCA embedding of the I3 shearing data.
*3d PCA-Embedding der Schneidsegmente aus I3.*

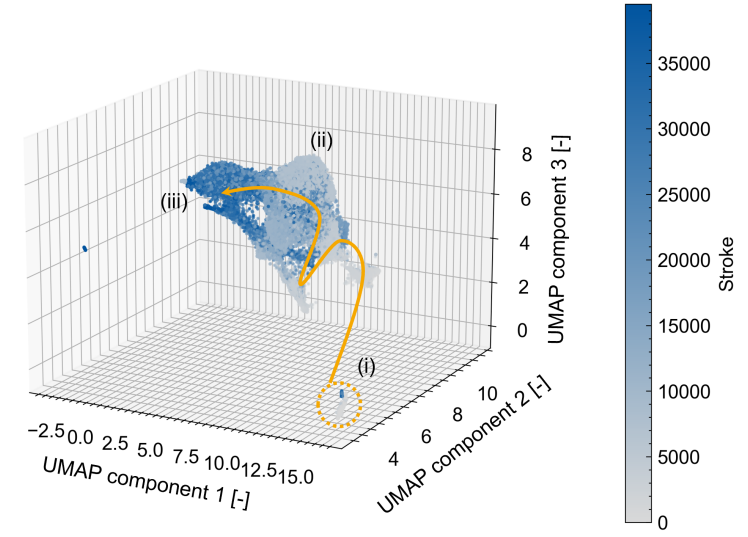



**Figure 5.8:** 3d UMAP embedding of the I4 shearing data.
*3d UMAP-Embedding der Schneidsegmente aus I4.*

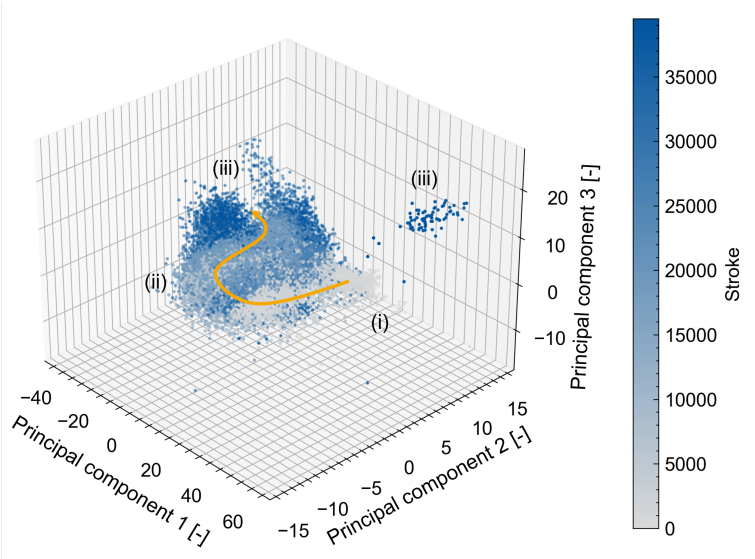



**Figure 5.9:** 3d PCA embedding of the I4 shearing data.
*3d PCA-Embedding der Schneidsegmente aus I4.*

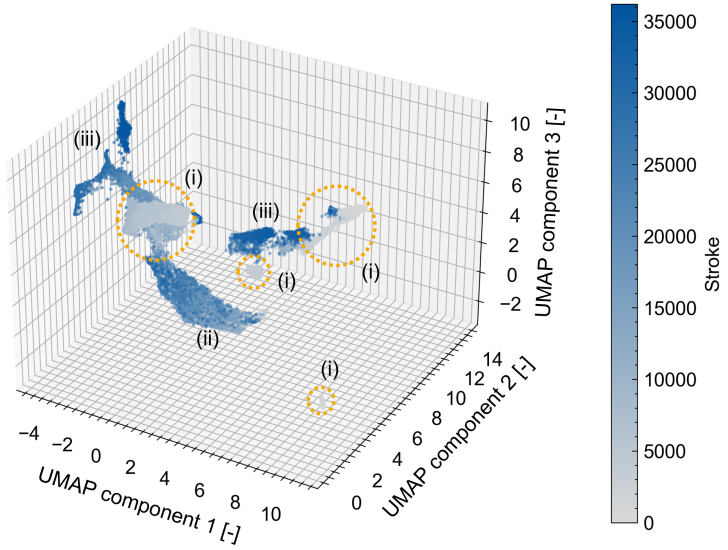



**Figure 5.10:** 3d UMAP embedding of the I5 shearing data.
*3d UMAP-Embedding der Schneidsegmente aus I5.*

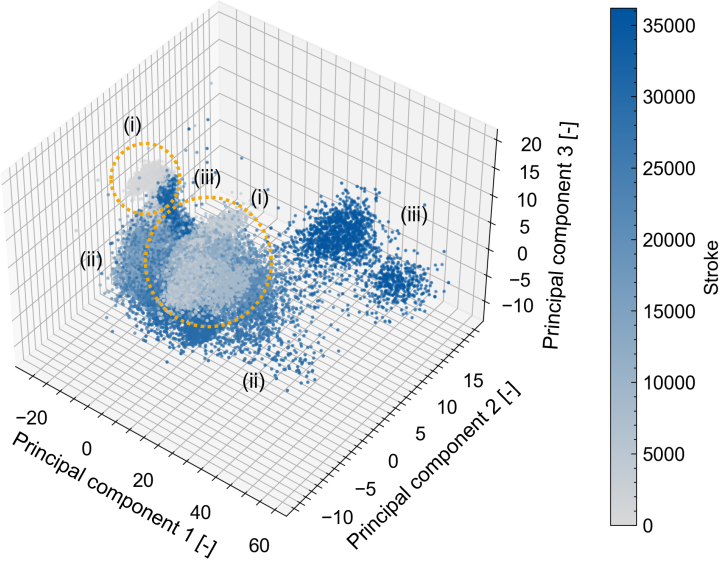



**Figure 5.11:** 3d PCA embedding of the I5 shearing data.
*3d PCA-Embedding der Schneidsegmente aus I5.*

**Table 5.1:** Overview of the results of the qualitative analysis of the 3d embeddings regarding an observable and clear evolution (that depends on the number of process cycles), the formation of singular clusters, and the observed dispersion of the data.
*Überblick über die Ergebnisse der qualitativen Analyse der 3D-Embeddings im Hinblick auf eine beobachtbare und klare Entwicklung (die von der Anzahl der Prozesszyklen abhängt), die Bildung von singulären Clustern und die beobachtete Streuung der Daten.*

| Dataset (Length) | Clear evolution | Singular clusters | Dispersion |
|---|---|---|---|
| I1 (Medium) | Yes | No singular clusters | Low |
| I2 (Long) | Yes | Singular clusters | Low |
| I3 (Long) | Partial | Singular clusters | Medium |
| I4 (Short) | Partial | Singular clusters | Medium |
| I5 (Short) | No | Many singular clusters | High |

### 5.1.3 Intermediate conclusion and unresolved challenges
*Zwischenfazit und offene Herausforderungen*

Research question **RQ2** was defined as **"What recurring patterns and differences can be qualitatively identified in the evolution of preprocessed industrial acoustic emission signals over several punch life cycles?"**. To partially answer this research question, Section 5.1 utilizes UMAP and PCA as dimensionality reduction and data visualization methods, to conduct a qualitative analysis of AE signal embeddings. These machine learning methods are crucial for handling the complex, high-dimensional characteristics of the acoustic emission data, allowing for a more nuanced understanding of wear patterns. Section 5.2 will complete the process of answering **RQ2**, as concluded in Section 5.3.

The industrial stroke series I1, I2, and I3 demonstrate a rather strong relationship between the embeddings and the number of process cycles completed. Stroke series I4 shows a weaker yet still observable relationship between the embedding and the number of completed process cycles. Since the process setup remained unchanged, these relationships are likely related to the punch wear.

Although stroke series I5 shows clusters of subsequent strokes within the embedding, no clear pattern similar to the other stroke series could be discerned. The UMAP embedding, depicted in Figure 5.10, showcases spread-out data clusters with no apparent systematic pattern. This variability could potentially be linked to excessive punch wear and other tool-related issues, but cannot be definitively concluded from the embeddings alone. While the dimensionality reduction techniques (feature extraction with TSFEL, feature filtering, and embedding with UMAP and PCA, see Figure 5.1) are instrumental in revealing data trends, they do not specify which features are particularly indicative of these trends.

Thus, in the following Section 5.2, the evolution of domain-specific features is analyzed to answer **RQ2** together with the already presented analysis in full. Domain-specific in this context means that the feature extraction is designed based on the *signal type* and the *context*. Since AE data from an industrial fineblanking process should be analyzed in large quantities, the feature extraction should i) leverage efficient algorithms, ii) be robust towards noise, iii) be able to deal with non-stationary signals, and iv) yield interpretable results.

## 5.2 Structural evolution of domain-specific acoustic emission signal features

*Strukturelle Entwicklung domänen-spezifischer Feature von Körperschallsignalen*

Section 5.2.1 lays out the methodology used to analyze the structural evolution of selected domain-specific AE signal features for the industrial stroke series. Section 5.2.2 presents the results of these analyses.

### 5.2.1 Methodology

*Methodik*

As stated in Section 2.3, the gathered signals can be considered non-stationary. Additionally, since the signals stem from an industrial setting, a substantial amount of noise within the data was expected. Hence, WELCH's method was used for each stroke of the industrial datasets to calculate the PSD and the PS.

The stroke-wise evolution of the PSD estimates within each dataset was visualized in the first step. To this end, the *scipy.signal.welch* function was used. Since identifying trends and reducing noise was the main objective in this first step, a low segment length (function parameter *nperseg*) of 128 was chosen. The function parameter *overlap* was chosen as 64 data points or 50 % of the segment length. Additionally, the depiction of all resulting PSDs was smoothed with a rolling mean, with a window size of 10. The pipeline is visualized in Figure 5.12.

In a second step, the power spectra of the strokes were calculated, again using the function from *SciPy* that implements WELCH's method to calculate the mean frequency of the power spectra. This was done with a PYTHON function developed by the author of this work.

Since the mean frequency represents an aggregated metric that averages over all frequency bins (weighted by the power of the respective bin), a higher frequency resolution was chosen with a segment length of 8,192 and a segment overlap of 4,096 or 50 % of the segment length. However, it is noteworthy that the calculated mean frequencies for the higher frequency resolution correlate strongly with the mean frequencies that result when a segment length of 128 is chosen (DC for I1: 0.984, I2: 0.959, I3: 0.953, I4: 0.969, and I5: 0.983). Hence, the choice of the segment length only has a marginal impact on the resulting mean frequency sequences, potentially since calculating the mean frequency effectively averages spectral

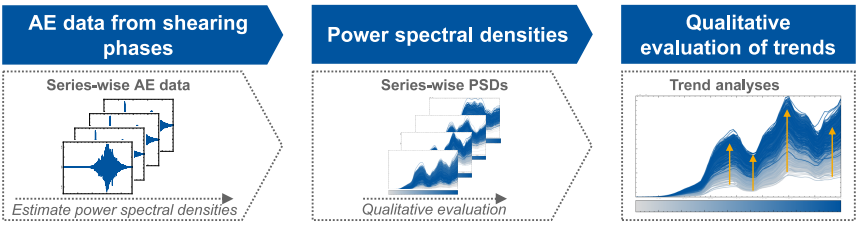



**Figure 5.12:** Developed pipeline for the qualitative analysis of the structural evolution of PSDs for the industrial stroke series.
*Entwickelte Pipeline zur qualitativen Analyse der strukturellen Entwicklung von PSDs für die industriellen Hubserien.*

components. For better visibility, a rolling mean of the mean frequency with a window size of 500 is displayed beside the raw data (scatter plot).

Additionally, machine standstills (following denoted as *events*), specifically times, durations, and reasons, were extracted from the logbooks of the company. Extracted events were labeled with reasons *Tool*, *Quality assurance (QA)*, and *Coil change (Coil)*. Subsequently, the stroke numbers were mapped to the strokes so that the last stroke prior to the event could be identified and annotated within the presented plots with vertical lines. The height of the marker on the vertical lines, e.g., in Figure 5.15, indicates the duration of the event. The full pipeline is illustrated in Figure 5.13.

Although industry professionals state that fluctuations at the level of the material have a substantial influence on the process stability and process result in fineblanking [UNTE19], systematic analyses of the influence of material fluctuations (i.e., fluctuations of the mechanical-technological properties) are still missing in the literature. Hence, an influence of the *Coil* events on the evolution of the mean frequency event was qualitatively evaluated. To this end, the evolution of the mean frequency feature is presented in an upper part containing the annotation of *Tool* and *QA* events and a lower part containing the annotation of *Coil* events.

The results of the documented approach are presented in Section 5.2.2. First, the evolution of the PSDs is visualized dataset-wise for 1,000 strokes per dataset to ensure the clarity of the figures. Following, the evolution of the mean frequency feature, including the annotated events, is presented.

### 5.2.2 Results

*Ergebnisse*

The following paragraphs present the results of the qualitative analysis of the structural evolution of the PSDs and the mean frequency feature for the industrial stroke series. The findings are summarized in Table 5.2.

**Qualitative evaluation of the evolution of domain-specific features I1** *Qualitative Evaluation der Entwicklung domänenspezifischer Feature für I1*

The evolution of the PSDs of industrial stroke series I1 is illustrated in Figure 5.14. Similar to the embeddings analyzed in Section 5.1.2, a clear trend in the number of strokes is visible. Specifically, the amplitudes of the PSDs increase throughout the stroke series. However, at the end of the stroke series, substantial decreases in the amplitudes of the

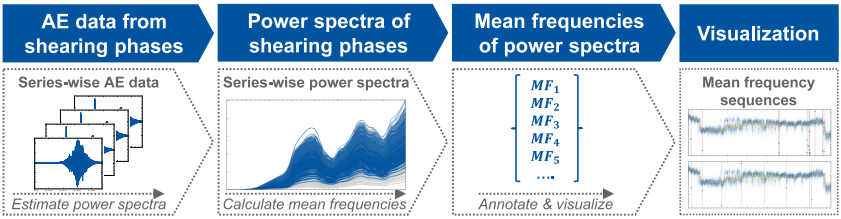



**Figure 5.13:** Developed pipeline for the qualitative analysis of the structural evolution of the mean frequency feature for the industrial shearing segments.
*Entwickelte Pipeline zur qualitativen Analyse der strukturellen Entwicklung des Features Mean Frequency für die industriellen Hubserien.*

PSDs are observable. Furthermore, a concentration of the PSDs is visible for the frequency ranges 75–115 kHz and 120–160 kHz. While the figure suggests that, judging by the PSD, frequency components larger than 200 kHz exhibit substantial power, the evolution of these frequency components can not be analyzed without aliasing due to the sampling frequency of 400 kHz.

The evolution of the mean frequency feature, as depicted in Figure 5.15, exhibits more nuances. Clearly, the mean frequency of the industrial stroke series I1 declines substantially for the first approximately 15,000 strokes. The decline can be divided into different phases: a first steep decline (until the approximately 500th stroke), a nearly constant mean frequency (approximately 500th–4,000th stroke), a second steep decline (4,000th–6,000th stroke), a second phase of nearly constant mean frequency (6,000th–8,000th stroke), and a third steep decline (8,000th–15,000th stroke). This coincides with the noticeable formation of a first phase (i) within the embeddings as visualized in Figures 5.2 and 5.3. While the mean frequency drops further after an increase that leads up to the approximately 22,500th stroke over the course of the stroke series, the steepness of the decline during the first 15,000 strokes remains unmatched. Interestingly, the final phase of the stroke series coincides with an accumulation of machine standstills due to the documented reason *Tool*. Specifically, 68 % of all standstills with reason *Tool* occur during the last 23 % of all strokes. Notably, the dispersion of the mean frequency increases towards the last approximately 5,000 strokes of the series, which coincides with an increase in the mean frequency feature.

A substantial effect of coil changes on the mean frequency feature can not be determined. In conclusion, the mean frequency feature shows a decline with clearly noticeable phases and transitions between these phases over the whole stroke series I1. Most noticeably, a steep decline at the beginning of the stroke series (first 15,000 strokes) occurs, and the evolution of the mean frequency exhibits increased dispersion towards the end of the stroke series with an increase in the mean frequency feature (last 5,000 strokes).

**Qualitative evaluation of the evolution of domain-specific features I2** *Qualitative Evaluation der Entwicklung domänenspezifischer Feature für I2*

Figure 5.16 shows the evolution of the PSDs of industrial stroke series I2. Analogous to series I1, an obvious trend can be identified since the amplitudes of the PSDs increased with an increasing amount of accumulated strokes within the series. The concentration of the PSDs is, in similarity to series I1, observable between 75–115 kHz and 120–160 kHz. In contrast to stroke series I1, the amplitudes of the PSDs do not decrease towards the end of stroke series I2.

Analyzing the evolution of the mean frequency reveals similarities but also substantial differences to stroke series I1. As illustrated in Figure 5.17, the mean frequency decreases throughout stroke series I2, with a steeper decline at the beginning of the stroke series and comparable transitions to stroke series I1.

After a steep decline of the mean frequency for the approximately first 100 strokes, the mean frequency remains approximately constant (100th–2,000th stroke). Interestingly, a considerable decline in the mean frequency occurs prior to the machine standstill with reason *Tool* at the 2,232nd stroke. Following, at the 3,756th stroke, another machine standstill with reason *Tool* occurs simultaneously with quality assurance, suggesting problems with the part quality. Until the approximately 15,000th stroke, the mean frequency remains roughly constant. Noteworthy, at the 12,889th stroke, a machine standstill with reason *Tool* occurred with a remarkably long duration of 46 min. Between the 15,455th and the 15,573rd stroke, seven machine standstills with reason *Tool* occurred during a sharp decline of the mean frequency and with a cumulated duration of 68 min. After a phase of noticeably higher dispersion but no profound change in the rolling mean of the mean frequency between

strokes 16,000 and 23,000, another steep decline and transition to a new mean frequency level occurred.

Interestingly, prior to this decline, two coil changes were carried out in quick succession, specifically a first coil change at stroke 41,716 (representing 6,450 strokes with the processed coil) and a coil change at stroke 42,633 (representing only 917 strokes). While the data from the logbook is not fully reliable, e.g., the first approximately 10,000 strokes in series I1 were seemingly carried out with a single coil. However, due to the observable coil change intervals, a failure to log a machine standstill as an event of type *Coil* is more likely. Across all industrial datasets, the mean amount of strokes between two coil changes is 5,395, with a rather high standard deviation of 1,709 strokes. The median and mode, which potentially represent the distribution of the events of type *Coil* better due to their robustness, are 5,840 strokes and 366 strokes. Major deviations from the median potentially indicate problems with the material. Two more similar occurrences can be found within Figure 5.17. Specifically, between strokes 41,716 and 47,980, four coil changes occurred. These coil changes coincide with four machine standstills with reason *QA* (strokes 42,326, 42,591, 42,634, and 46,914) and four machine standstills with reason *Tool* (strokes 46,664, 46,670, 46,717, and 46,723) that follow a steep decline in the mean frequency feature between strokes 42,000 and 46,000. The decline is followed by a rather high dispersion of the mean frequency feature that persists approximately from stroke 47,000 to stroke 58,000.

For the remainder of the stroke series, the mean frequency remains relatively constant, with a level drop at approximately stroke 77,000, followed by another machine stand still with reason *QA*. Towards the end of the stroke series, the dispersion of the mean frequency increases slightly.

**Qualitative evaluation of the evolution of domain-specific features I3** *Qualitative Evaluation der Entwicklung domänenspezifischer Feature für I3*

Similar to the PSDs of stroke series I1 and I2 visualized in Figures 5.14 and 5.16, the evolution of the PSDs of stroke series I3 (see Figure 5.18) exhibits a clear trend with an increase of the amplitudes of the PSDs over the course of the stroke series. This trend is in agreement with the clear evolution of the embeddings depicted in Figures 5.6 and 5.7. Furthermore, the concentration of the PSDs in the frequency ranges 75–115 kHz and 120–160 kHz is in line with the PSDs of I1 and I2, displayed in Figures 5.14 and 5.16.

After an initial steep decrease in the mean frequency (first approximately 100 strokes) and a machine standstill with reason *Tool* at stroke 803, the mean frequency steeply increases from thereon until stroke 4,000. Across all industrial datasets, this behavior represents a remarkable distinguishing feature of this stroke series.

From the approximately 4,000th stroke until stroke 31,000, the mean frequency decreases rather steeply and similarly to the stroke series I1 and I2, with an event of type *QA* at the 22,502nd stroke. From the 31,000th until the approximately 75,000th stroke, the mean frequency remains remarkably steady in comparison to stroke series I1. However, around the 74,000th stroke, the raw mean frequency data shows a steep increase for a considerable amount of strokes, followed by an event of type *Tool* at stroke 76,764, followed by a decrease in the mean frequency. Interestingly, three events of type *QA* occur between strokes 82,644 and 83,987, followed by two events of type *Tool* with a cumulated duration of 53 min. Until the end of the stroke series, the mean frequency declines slightly.

A major effect of the coil changes on the mean frequency can not be observed. Furthermore, the coil change intervals do not indicate an early coil change that could indicate problems with the material.

**Qualitative evaluation of the evolution of domain-specific features I4** *Qualitative Evaluation der Entwicklung domänenspezifischer Feature für I4*

The evolution of the PSDs of stroke series I4 is portrayed in Figure 5.20. While the trend generally shows agreement with the stroke series I1 (Figure 5.14), I2 (Figure 5.16), and I3 (Figure 5.18), specifically an increase in the amplitude of the PSDs with an increasing number of process cycles completed, obvious differences are noticeable. Specifically, while for the stroke series I1, I2, and I3, relatively distinct concentrations within certain frequency ranges can be observed, Figure 5.20 displays a concentration within the frequency range 75–160 kHz instead of a concentration within two frequency ranges as in stroke series I1, I2, and I3. This finding coincides with the substantially more scattered embeddings in comparison to stroke series I1, I2, and I3, as illustrated in Figures 5.8 and 5.9 and as described in Section 5.1.2.

The evolution of the mean frequency of stroke series I4 reveals a high number of machine standstills at the beginning of the stroke series. Specifically, until stroke 3,576, 15 events of type *Tool* could be observed with a cumulated duration of 90 min. This represents a distinctive feature of stroke series I4. In agreement with stroke series I1, I2, and I3, a steep decline in the mean frequency can be observed in the first phase of the stroke series. Until the approximately 11,000th stroke, the mean frequency declines substantially. Interestingly, the mean frequency increases until stroke 17,111 and an event of type *Tool* with a rather high duration of 53 min. Afterward, the mean frequency remains relatively stable until stroke 28,000.

The mean frequency drops sharply after a coil change at stroke 28,165. Until approximately stroke 36,500, the mean frequency increases steadily to then drop sharply again. Until the end of the stroke series, the mean frequency increases slightly.

Similar to stroke series I1, I2, and I3, I4 showcases that the mean frequency develops in certain levels and can change levels either abruptly or steadily. Abrupt level changes often coincide with machine standstills.

**Qualitative evaluation of the evolution of domain-specific features I5** *Qualitative Evaluation der Entwicklung domänenspezifischer Feature für I5*

As laid out in Section 5.1.2, the embeddings of stroke series I5 (see Figure 5.10 and Figure 5.11) display a high amount of dispersion in the sense that although neighboring strokes can broadly be grouped, however, no global systematic that depends on the number of completed process cycles arises from the embeddings. The PSDs' evolution over stroke series I5 emphasizes this finding, as can be inferred from Figure 5.22. While generally similarities to the evolution of the PSDs of stroke series I1, I2, and I3 can be identified, especially in the concentration of the power of the signal in the frequency ranges 75–115 kHz and 120–160 kHz, the evolution of the PSDs shows a higher amount of dispersion in comparison to stroke series I1, I2, and I3.

Figure 5.23 visualizes the evolution of the mean frequency. Clearly, the mean frequency decreases for the first approximately 2,500 strokes and followingly drops steeply. Only 2,733 strokes after the previous coil change, events of type *Coil* and *Tool* occurred at stroke 3,573. Interestingly, the mean frequency remains relatively stable after these events until stroke 8,376, when another event of type *Coil* occurred. The mean frequency shows an exceptional (in comparison to the remaining stroke series) dispersion for the following strokes until stroke 22,976, with almost periodical substantial increases and decreases. Another event of type *Tool* follows shortly after at stroke 23,075. Remarkably, after these events, the mean frequency's dispersion is lower than before the event. However, sudden drops in the mean frequency can still be observed.

Towards the end of the stroke series, an accumulation of events of type *QA* and *Tool* can be observed. Six machine standstills with reason *QA* occurred between strokes 30,371 and 36,181, indicating severe problems with the resulting part quality. This is emphasized by the two coil changes in rapid succession at strokes 30,162 and 30,857. Interestingly, the sudden drops within the mean frequency do not appear after the documented events anymore, indicating that problems with the tool were potentially solved. However, it is noteworthy that an event of type *Tool* and a rather high duration of 41 min occurred at stroke 32,165 during a rather stable period of the mean frequency evolution. At stroke 34,029, another event of type *Tool* occurred. Shortly afterward, the mean frequency drops steeply, followed by another event of type *Tool* at stroke 34,409. Until approximately stroke 34,500, the mean frequency continues to drop steeply until it remains relatively stable until the end of the stroke series. The stroke series ends with an event of type *QA*, indicating that the part quality was assessed and deemed unacceptable, finally leading to a tool change.

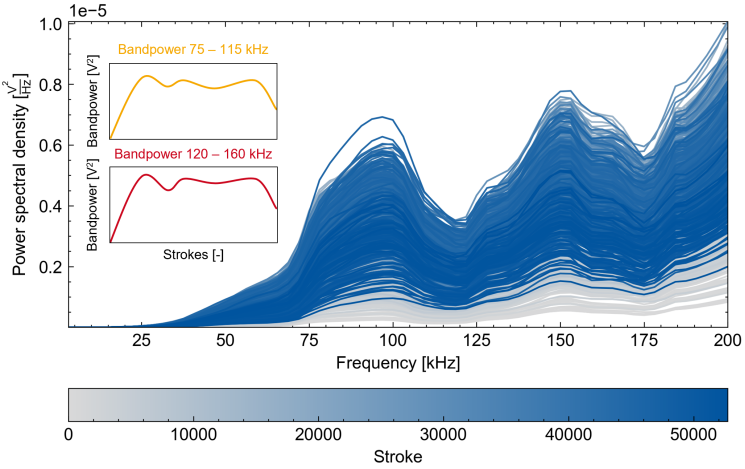



**Figure 5.14:** Visualization of the PSD evolution of the shearing phase within industrial stroke series I1 (*segment length* 128, 50 % *segment overlap*, *rolling mean* window size 10).
*Visualisierung der PSD-Entwicklung der Schneidphasen innerhalb der industriellen Hubserie I1 (Segmentlänge 128, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 10).*

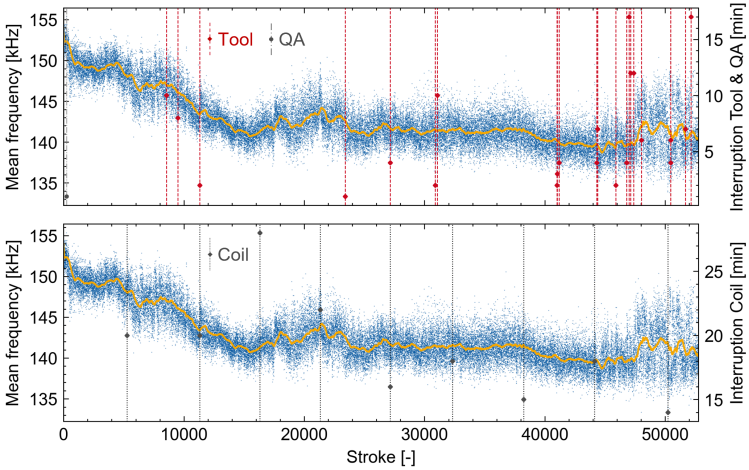



**Figure 5.15:** Visualization of the evolution of the mean frequency feature in series I1 with machine standstills (reasons *Tool*, *Coil* change, and Quality assurance (*QA*)) (*segment length* 8,192, 50 % *segment overlap*, *rolling mean* window size 500).
*Visualisierung der Entwicklung des Features Mean Frequency in der Serie I1 mit Maschinenstillständen (Gründe Werkzeug, Coilwechsel, und Qualitätssicherung (QA)) (Segmentlänge 8.192, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 500).*

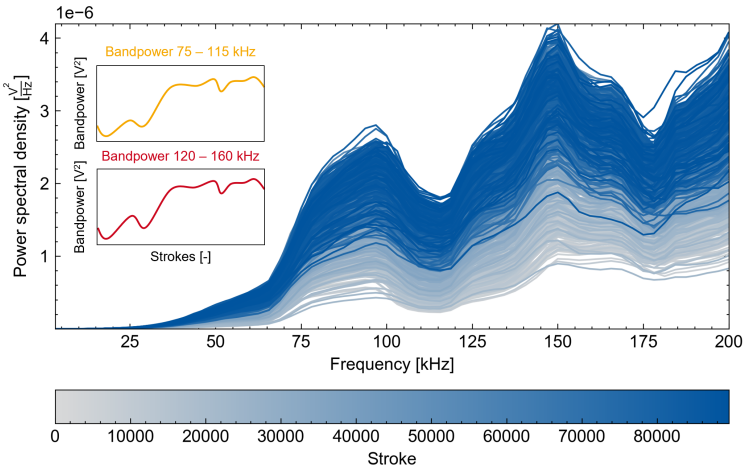



**Figure 5.16:** Visualization of the PSD evolution of the shearing phase within industrial stroke series I2 (*segment length* 128, 50 % *segment overlap*, *rolling mean* window size 10).
*Visualisierung der PSD-Entwicklung der Schneidphasen innerhalb der industriellen Hubserie I2 (Segmentlänge 128, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 10).*

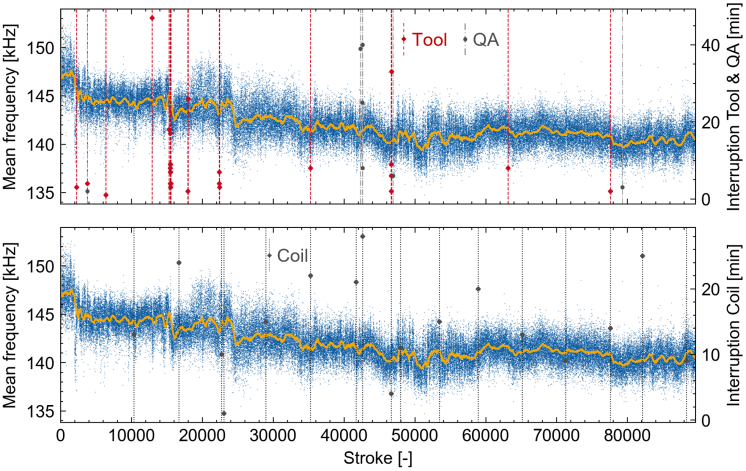



**Figure 5.17:** Visualization of the evolution of the mean frequency feature in series I2 with machine standstills (reasons *Tool*, *Coil* change, and Quality assurance (*QA*)) (*segment length* 8,192, 50 % *segment overlap*, *rolling mean* window size 500).
*Visualisierung der Entwicklung des Features Mean Frequency in der Serie I2 mit Maschinenstillständen (Gründe Werkzeug, Coilwechsel, und Qualitätssicherung (QA)) (Segmentlänge 8.192, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 500).*

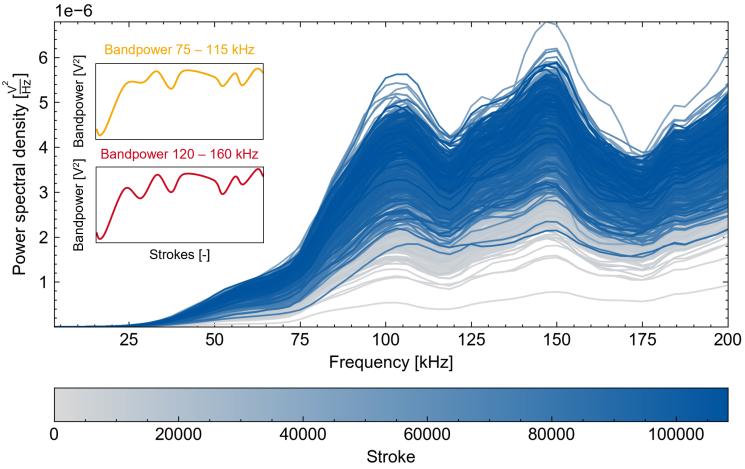



**Figure 5.18:** Visualization of the PSD evolution of the shearing phase within industrial stroke series I3 (*segment length* 128, 50 % *segment overlap*, *rolling mean* window size 10).
*Visualisierung der PSD-Entwicklung der Schneidphasen innerhalb der industriellen Hubserie I3 (Segmentlänge 128, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 10).*

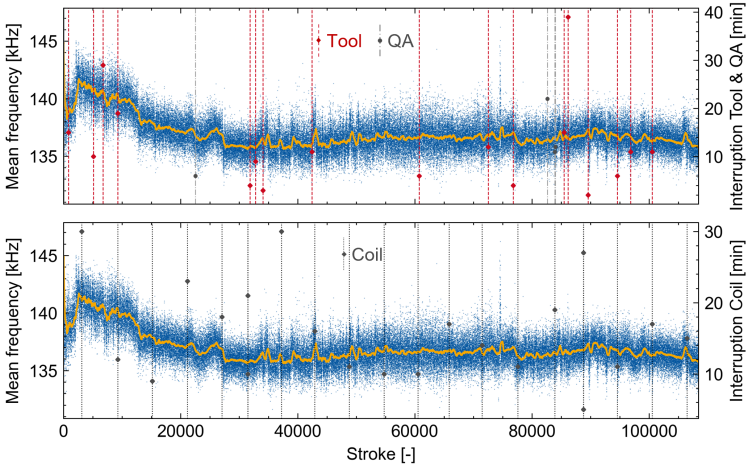



**Figure 5.19:** Visualization of the evolution of the mean frequency feature in series I3 with machine standstills (reasons *Tool*, *Coil* change, and Quality assurance (*QA*)) (*segment length* 8,192, 50 % *segment overlap*, *rolling mean* window size 500).
*Visualisierung der Entwicklung des Features Mean Frequency in der Serie I3 mit Maschinenstillständen (Gründe Werkzeug, Coilwechsel, und Qualitätssicherung (QA)) (Segmentlänge 8.192, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 500).*

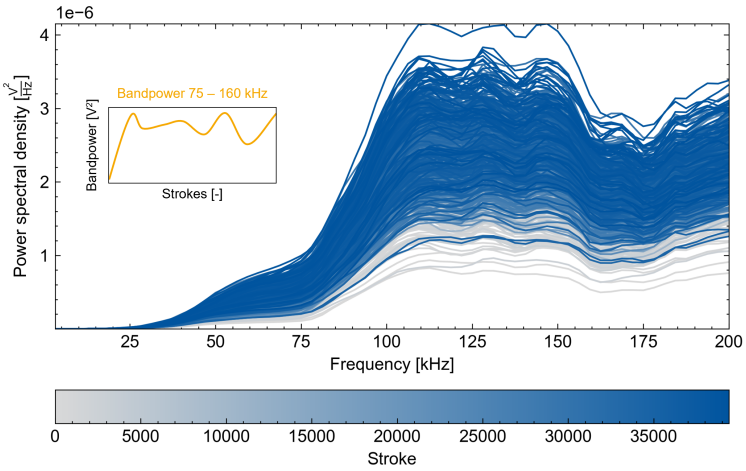



**Figure 5.20:** Visualization of the PSD evolution of the shearing phase within industrial stroke series I4 (*segment length* 128, 50 % *segment overlap*, *rolling mean* window size 10).
*Visualisierung der PSD-Entwicklung der Schneidphasen innerhalb der industriellen Hubserie I4 (Segmentlänge 128, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 10).*

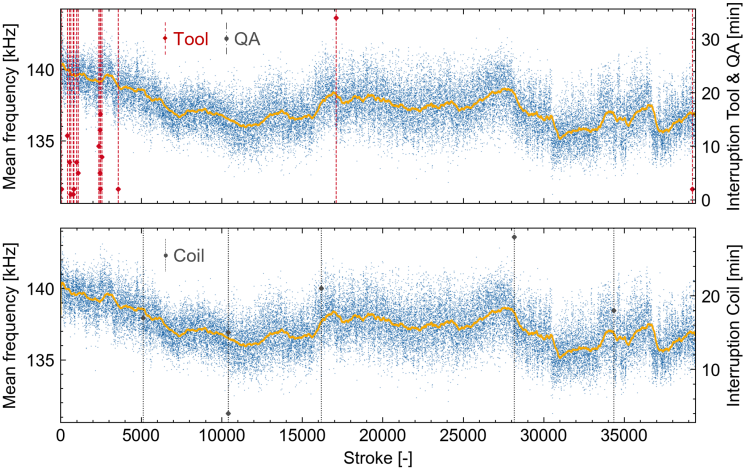



**Figure 5.21:** Visualization of the evolution of the mean frequency feature in series I4 with machine standstills (reasons *Tool*, *Coil* change, and Quality assurance (*QA*)) (*segment length* 8,192, 50 % *segment overlap*, *rolling mean* window size 500).
*Visualisierung der Entwicklung des Features Mean Frequency in der Serie I4 mit Maschinenstillständen (Gründe Werkzeug, Coilwechsel, und Qualitätssicherung (QA)) (Segmentlänge 8.192, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 500).*

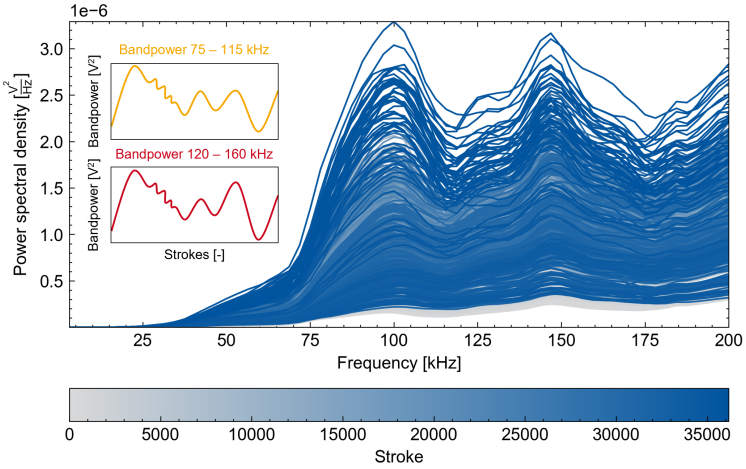



**Figure 5.22:** Visualization of the PSD evolution of the shearing phase within industrial stroke series I5 (*segment length* 128, 50 % *segment overlap*, *rolling mean* window size 10).
*Visualisierung der PSD-Entwicklung der Schneidphasen innerhalb der industriellen Hubserie I5 (Segmentlänge 128, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 10).*

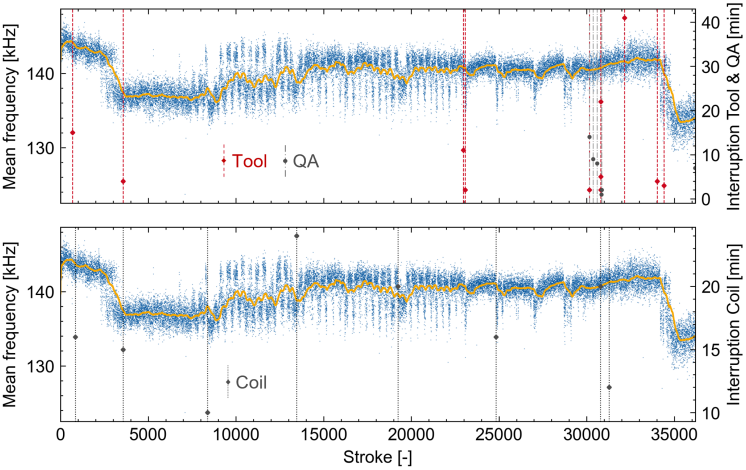



**Figure 5.23:** Visualization of the evolution of the mean frequency feature in series I5 with machine standstills (reasons *Tool*, *Coil* change, and Quality assurance (*QA*)) (*segment length* 8,192, 50 % *segment overlap*, *rolling mean* window size 500).
*Visualisierung der Entwicklung des Features Mean Frequency in der Serie I5 mit Maschinenstillständen (Gründe Werkzeug, Coilwechsel, und Qualitätssicherung (QA)) (Segmentlänge 8.192, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 500).*

**Table 5.2:** Overview of the results of the qualitative analysis of the PSD concentration (PSD conc.) and dispersion (Disp), global mean frequency (MF) trends (decrease: ↓, approximately constant: →, and variable: (var)), and observed relationships (rel) between standstills (Tool/QA/Coil) and MF evolution for the industrial stroke series.
*Überblick über die Ergebnisse der qualitativen Analyse der PSD-Konzentration (PSD conc.) und der Streuung (Disp), globaler Trends des Features Mean Frequency (MF) (Abnahme: ↓, annäherend konstant: → und variabel: (var)) und beobachtete Beziehungen (rel) zwischen Stillständen (Tool/QA/Coil) und der MF-Entwicklung für die industrielle Hubserie.*

| Dataset (Length) | PSD conc. [kHz] (Disp) | MF trend (Disp) | MF-Standstill rel |
|---|---|---|---|
| I1 (Medium) | 75–115, 120–160 (Low) | ↓ to var (Medium) | Tool |
| I2 (Long) | 75–115, 120–160 (Low) | ↓ to → (Low) | Tool/QA/Coil |
| I3 (Long) | 75–115, 120–160 (Low) | ↓ to → (Low) | Tool/QA |
| I4 (Short) | 75–160 (Medium) | ↓ to var (Medium) | Tool/Coil |
| I5 (Short) | 75–115, 120–160 (High) | ↓ to var to ↓ (High) | Tool/Coil |

## 5.3 Intermediate conclusion and unresolved challenges

*Zwischenfazit und offene Herausforderungen*

Similar to the qualitative analysis of the embeddings documented in Section 5.1.2, it is concluded that the industrial stroke series I1, I2, and I3 show a strong visually apparent relationship between the estimated PSD of a stroke and the number of process cycles carried out before said stroke. For these stroke series, a concentration of the power can be located within the frequency ranges 75–115 kHz and 120–160 kHz. The PSD plots suggest considerable power in the higher frequency ranges; however, due to the sampling rate of 400 kHz, this can not be determined with certainty.

Analyzing the evolution of the mean frequency feature across all datasets, and additionally, the annotated events reveals more nuanced insights. SHANBHAG et al. stated that the onset and severity of galling conditions could be detected in sheet-metal stamping processes by decreasing mean frequency. The authors referenced [BASS07] and argued that the reduction in the mean frequency could potentially be attributed to the number of resonant frequency components exhibited in a progressed wear state [SHAN18]. However, it is important to note that BASSIUNY i) analyzed a strain signal and ii) artificially introduced irregularities into the process, e.g., by a variation of the sheet metal thickness. Still, since such irregularities are to be expected in the industrial reality of a sheet-metal stamping process, the analysis of the mean frequency for the industrial stroke series is promising following SHANBHAG's argument.

An interesting observation across all datasets is that the mean frequency has certain *modes*, i.e., *steady phases* and *transitions between these steady phases*. These transitions are either *steep* (e.g., the steep decline in mean frequency at the beginning of the stroke series) or steadily increasing or decreasing.

SHANBHAG et al. presented mean frequency data with a steep decline in mean frequency during the initial phase of the die lifecycle ([SHAN18] and [SHAN20]) as well, analogous to the data presented within this work. Potentially, this could resemble the breaking-in phase of the tool. No tool failure occurred within the analyzed stroke series in this work, which explains the absence of a clear failure phase. However, it is noteworthy that throughout the stroke

series, the steep decline in the mean frequency feature interestingly remains a rather singular phase, except for stroke series I5. Since the severity of galling is expected to increase over the lifecycle of the die and the punch, a further decline in the mean frequency feature should be observable. Hence, the absolute values of the mean frequency do not seem promising as a tool to identify the severity of galling conditions in this process and measurement setup. Rather than the absolute values of the mean frequency feature, the dispersion of the feature could be attributed qualitatively to phases with a high number of machine standstills. Potentially, similar to the work of NIEMIETZ et al. in a recent study, a measure for the dispersion of the mean frequency feature or more generally, specific frequency ranges allows for interesting descriptive insights into what NIEMIETZ et al. coined *process stability* [NIEM23], i.e. a stable process performance under naturally varying conditions (material, wear progression, etc. [LIEW22]).

However, since the specific wear conditions for the data were unknown, causality remains unclear for this work. Still, some association between events of type *Tool*, *QA*, and *Coil*, and changes within the mean frequency and specifically the dispersion of the mean frequencies within certain stroke ranges, could be qualitatively identified and point to the spectrum of the AE signal containing important information for an indirect tool condition monitoring (TCM) and the related process stability. The analyses concluded in Sections 5.1 and 5.2 represent the answer to research question **RQ2**.

The ultimate goal of indirect TCM is to gain insights about the tool condition from process signals. Hence, to evaluate the predictive power of AE signals regarding the tool condition said tool condition must be known to evaluate model performance. Therefore, Chapter 6 introduces a proxy variable for the punch wear state and presents the analysis of the development of this proxy variable over the course of datasets L1 and L3. Furthermore, the evolution of the acquired AE signals of the laboratory datasets L1, L2, and L3 is analyzed. For the datasets L1 and L2, which both contain punch failure events, the relationship between the punch failure and the evolution of the AE signals is qualitatively evaluated.

# 6 Observation and evaluation of a punch wear proxy variable

*Beobachtung und Evaluierung einer Proxy-Variablen für den Stempelverschleiß*

Chapter 6 *proposes a proxy variable for estimating the punch wear* and tracks this proxy variable across laboratory datasets. Consequently, the results — namely the proxy variable and its observation — are pivotal for testing the research hypothesis, as they enable punch wear to be used as labels for training and testing machine learning models.

Section 6.1 lays out the *physical background and motivation* for choosing the proxy variable. Following, Section 6.2 documents the *experimental setup* for the laboratory stroke series. The *development* of the observed proxy variable and the *AE signals within the laboratory stroke series* are presented in Section 6.3.

## 6.1 Motivation and physical background

*Motivation und physikalischer Hintergrund*

As stated in Section 2.1.2, VOIGTS described that in fineblanking, the shearing edge and lateral surface of the punch and the sheared surface of the scrap web, which denotes the scrap part that results from the shearing operation, represent a tribological system during the shearing and stripping phase (see Figure 2.3) with the lubricant as an intermediate [VOIG21]. Hence, the wear of the punch has a direct influence on the quality of the sheared surface of the scrap web.

Surface roughness is a measure of the quality of a surface. Its evaluation is highly important for fundamental problems, such as positional accuracy, contact deformation, and friction. The surface roughness can be evaluated in 2d and 3d, whereas in the 2d case, *profile height deviations of a line* on the surface to a *mean line* are evaluated. In the 3-dimensional case *profile height deviations of an area* on the surface to a mean plane are evaluated [GADE02]. One of the most frequently used parameters to evaluate surface roughness in 2d is the *arithmetic average roughness* $R_\mathrm{a}$ [DECH00], which is a so-called *amplitude parameter* [GADE02]. The arithmetic average roughness represents a robust measure of the surface roughness since it describes the arithmetic mean of the profile height deviation to a mean line. Since the 2d surface roughness measurement represents an efficient approach to surface quality evaluation, it was used to assess the sheared surfaces of the scrap web segments in this work. Moreover, $R_\mathrm{a}$ was used as an evaluation parameter because of its robustness and widespread acceptance.

Surface roughness evaluation can be conducted using both optical and tactile methods. A tactile approach (*contact stylus profilometry*) was employed for the measurements within this work since it did not require disassembling the scrap web segments before measurement. For contact stylus profilometry, a stylus is moved over the surface for a defined length in a straight line. Depending on the expected surface roughness, a minimum length has to be chosen. The measurements were done perpendicular to the shearing direction. Hence, surface impairment induced by the punch can be quantified with this approach.

The content of Section 6.3.3 was in part published in [UNTE23]. Specifically, distinguishing between punch surfaces and punches in the analysis of the evolution of the proxy variable represents unpublished material. Furthermore, the analysis of the evolution of the PSDs has not been published yet. The author of this work handled the planning, material procurement, and conduct of stroke series L3. Stroke series L1 and L2 were planned and conducted by colleagues of the author. The author sampled the scrap web segments during stroke series L1 and L3 with the assistance of a technician. All material testing was carried out by the author with the assistance of a technician. CHATGPT (GPT-4, `https://chatgpt.com/`) was used in this chapter to proofread texts generated by the author and for support with LATEX formatting. DEEPL (`https://www.deepl.com/translator/`) was used for translation purposes.

Let $l > 0$ denote the length of the line on the surface and $p(l)$ the profile height of the surface. The arithmetic average roughness $R_\mathrm{a}$ is defined as [GADE02]

$$R_\mathrm{a} = \frac{1}{l} \cdot \int_0^l p(k)\, dk. \tag{49}$$

More intuitively, $R_\mathrm{a}$ can be imagined as the height of a rectangle with length $l$ and an area equal to the profile height deviations since Eq. 49 can be transformed to

$$R_\mathrm{a} \cdot l = \int_0^l p(k)\, dk \tag{50}$$

by means of multiplication with $l$. Figure 6.1 illustrates this intuitive representation.

## 6.2 Experimental setup and methodology

*Experimenteller Aufbau und Methodik*

The specific process setups for the laboratory experiments are laid out in Section 4.2.1. For clarity reasons, the setups are summarized within Table 6.1.

Since a punch breakage occurred in stroke series L2 that was attributed to a miscalibration of the tool, the series was discontinued. Thus, no scrap web segments could be collected from L2. Data acquired prior to the punch breakage is analyzed within Section 6.3.2.

After finishing the stroke series, the punches were examined with a scanning electron microscope (SEM). The most severely worn shearing edge surfaces of the punches were

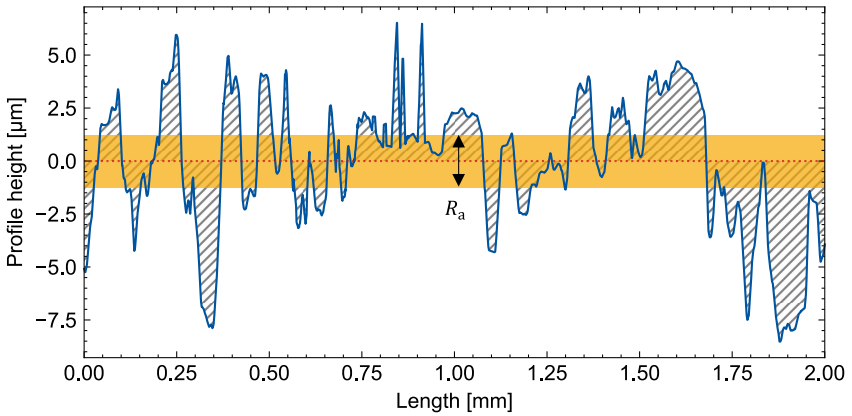



**Figure 6.1:** Representation of the arithmetic average roughness $R_\mathrm{a}$ as the height of a rectangle with length 2 and area equal to the integral of the profile height deviations over the length (based on [KEYE23]).
*Darstellung der arithmetischen mittleren Rauheit $R_\mathrm{a}$ als Höhe eines Rechtecks mit der Länge 2 und der Fläche gleich dem Integral der Profilhöhenabweichungen über die Länge.*

**Table 6.1:** Overview of the specifics of the laboratory datasets.
*Überblick über die Gegebenheiten der Datensätze aus der Laborumgebung.*

| Name | Material/Thickness | Stroke frequency | Num. of strokes | Segment sampling |
|---|---|---|---|---|
| L1 | 100Cr6/6 mm | 50 $min^{-1}$ | 30,284 | Every 2,500th |
| L2 | 42CrMo4/6 mm | 50 $min^{-1}$ | 970 | Every 2,500th |
| L3 | 58CrV4/5 mm | 50 $min^{-1}$ | 40,450 | Every 200th |

identified qualitatively to determine suitable sheared surfaces of the scrap web segments to examine. The roughnesses of the identified sheared surfaces were subsequently tested with a contact stylus profilometer of type JENOPTIK Waveline W920RC. For each surface, three measurements were done. The measurements started at the midpoint of the sheet metal thickness (in $y$-direction, see Figure 6.3 (b)), with a distance of 2.4 mm to the edge radius of the sheared surface in $x$-direction, representing the measurement length. A vertical offset of 0.8 mm was maintained perpendicular to the shearing direction for each measurement. The selected sheared surfaces of the scrap web segments and the corresponding punch surfaces are illustrated in Figure 6.2 for the stroke series L1 and L3. Figure 6.3 presents the (a) SEM images, (b) selected sheared surfaces of a sample scrap web segment, and (c) the testing procedure for the stroke series L3.

Figure 6.4 visualizes one measurement of the profile height deviations for stroke series L3 for the first and the last stroke of the stroke series, showcasing a clear degradation of the sheared surface quality. To robustify the results from the measurements, the mean of all measurements for a sheared surface was calculated and used as a representative roughness for this scrap web segment. The resulting roughness values were used as labels for the supervised ML models trained and evaluated within Chapter 7. In total, 78 roughness tests were carried out for stroke series L1. For stroke series L3, 2,436 roughness tests were conducted.

While this approach represents a viable solution in a controlled setting, implementing in-line supervision of the sheared surface roughness of scrap web segments in an industrial mass production setting presents major challenges. The main barrier is the financial aspect since

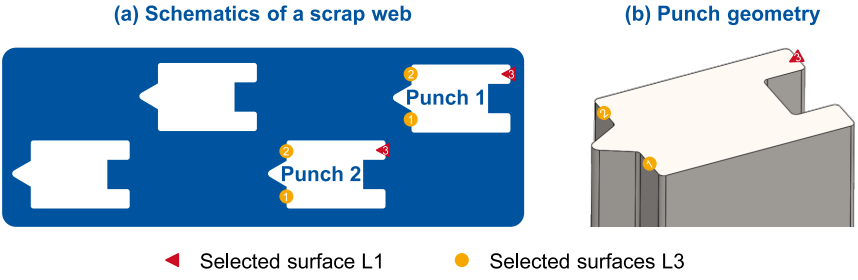



**Figure 6.2:** Depiction of (a) selected sheared surfaces of the scrap web and (b) corresponding surfaces on the punch.
*Darstellung (a) ausgewählter Schnittflächen des Stanzgitters und (b) entsprechende Oberflächen am Stempel.*

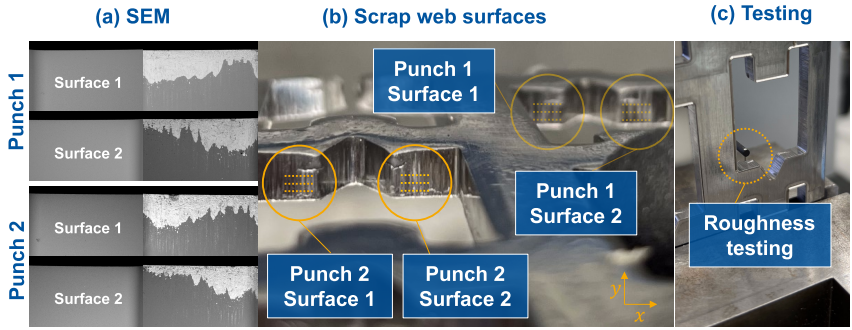



**Figure 6.3:** (a) SEM images of the punch surfaces with the most severe wear (40-fold magnification) (L3) before / after, (b) selected scrap web sheared surfaces, and (c) roughness testing procedure (see [UNTE23]).
*(a) REM-Aufnahmen der Stellen am Stempel mit dem stärksten Verschleiß (40-fache Vergrößerung) (L3) vorher/nachher, (b) ausgewählte Stanzgitter-Schnittflächen und (c) Darstellung der Rauheitsprüfungen.*

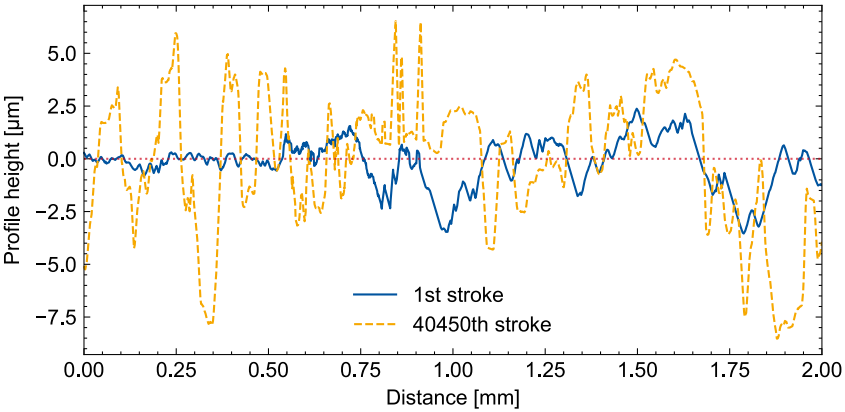



**Figure 6.4:** Comparison between profile height deviation from scrap web sheared surfaces at the beginning and the end of the punch life.
*Vergleich der Profilhöhenabweichung von Stanzgitter-Schnittflächen zu Beginn und am Ende der Lebensdauer eines Stempels.*

the initial setup and ongoing maintenance cost is far too high for practical application since sufficient robustness towards rough industrial conditions needs to be ensured.

Following, the results of the measurements are presented. Additionally, the PSDs of the laboratory stroke series L1, L2, and L3 are presented analogous to Section 5.2.1.

## 6.3 Results

*Ergebnisse*

Following, the results of the described experiments are presented. Section 6.3.1 presents the results for stroke series L1. The results for stroke series L2 are laid out in Section 6.3.2. Section 6.3.3 concludes Section 6.3 with the results for stroke series L3.

### 6.3.1 Results of the preparative experiment L1

*Ergebnisse der Voruntersuchungen L1*

The punch position in stroke series L1 was switched after 12,800, 15,881, and 25,675 strokes to account for cavity effects. Furthermore, chipping occurred after 15,881 strokes (see Figure 6.5) on one of the punches. Hence, the punch was switched to an analogous punch in pristine condition. The switches in position and replacing the punch complicate the analysis of the roughness of the scrap web segments.

Hence, Figure 6.6 (a) visualizes the results of the roughness measurements punch-wise. Punch 2 was replaced with Punch 3 from stroke 15,882 on. Clearly, the roughness of the sheared surfaces that resulted from interaction with Punch 2 increased substantially prior to the chipping on Punch 2. Interestingly, the roughness of the sheared surfaces that resulted from interaction with Punch 1 increased considerably prior to the chipping on Punch 2 as well. Notably, with Punch 3 replacing Punch 2, this trend stops until the roughness increases exponentially towards the end of stroke series L1, concluding with a punch screw breakage. The roughness of the sheared surfaces resulting from interaction with Punch 3 increases approximately linearly.

Since the AE signal is the result of the simultaneous interaction of both punches with the sheet metal, it is appropriate to average the roughness sample-wise, as illustrated in Figure 6.6 (b). While the averaged roughness measured for the scrap web segment sampled at stroke 15,000 clearly shows a substantial increase prior to chipping, the averaged roughnesses resemble the wear that would be expected in a tribological system as described in Section 2.1 and as depicted in Figure 2.6. The replacement of Punch 2 with a punch in pristine condition has a substantial impact on the averaged roughness, and only one surface was examined. However, since the roughness shows a profound increase prior to chipping and the punch screw breakage, the preparative experiment demonstrates promising results.

To use the sheared surface roughnesses of the scrap web segments as labels for ML models, the sampling interval of the scrap web segments seems too low. Important information, e.g., the roughness evolution shortly prior to chipping or the punch screw breakage, is missing. Thus, sampling intervals for the main experiment were chosen with a higher frequency,

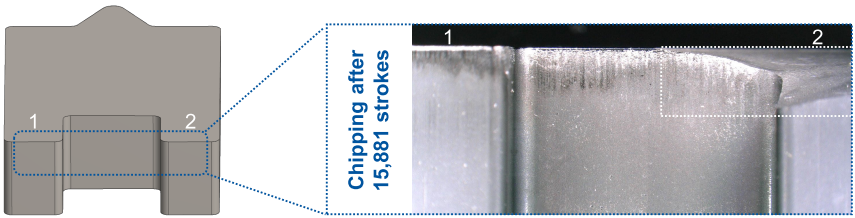



**Figure 6.5:** Visualization of chipping on Punch 2 during stroke series L1.
*Visualisierung des Chippings an Stempel 2 in Hubserie L1.*

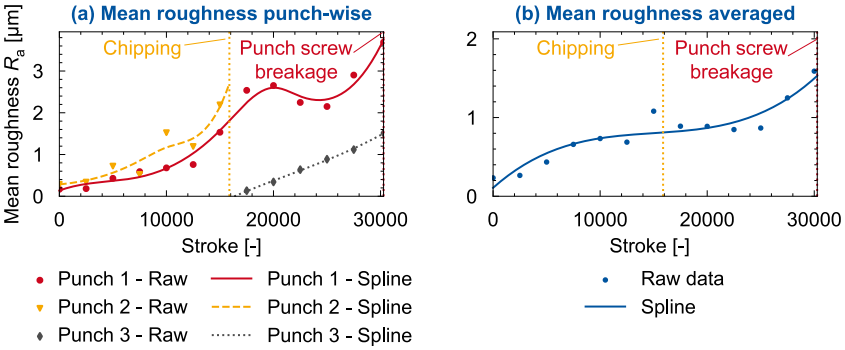



**Figure 6.6:** Visualization of the evolution of the sheared surface roughness for the scrap web segment samples in L1 (a) punch-wise and (b) averaged.
*Visualisierung der Entwicklung der Schnittflächen-Rauheit für die Stanzgitter aus L1 (a) aufgelöst nach Stempeln und (b) gemittelt.*

specifically every 200th stroke. The results of this experiment are documented in Section 6.3.3.

Following, the evolution of the PSDs is presented for stroke series L1. To enable a meaningful interpretation using the provided colorbar, the dataset is divided into the AE signals recorded prior to the occurrence of chipping and AE signals after chipping.

**Evolution of the PSDs of L1** *Entwicklung der spektralen Leistungsdichten von L1*

Figure 6.7 reveals that the power substantially increases prior to chipping for the shearing segments. While this increase in power is spread across all frequencies, it is especially apparent for the frequency range 75–260 kHz. This effect is even more visible for the stripping segments, see Figure 6.9, with an even higher emphasis on the frequency range 75–160 kHz.

Remarkably, after chipping and after substituting the damaged punch with a punch in pristine condition, Figure 6.8 and Figure 6.10 reveal that the PSDs exhibit even higher amplitudes than prior to the occurrence of chipping. However, over the course of the remainder of the stroke series, this effect recedes. The increased power could be the result of a combination of the break-in phase of the pristine punch and increased wear accumulation of Punch 1, see Figure 6.6.

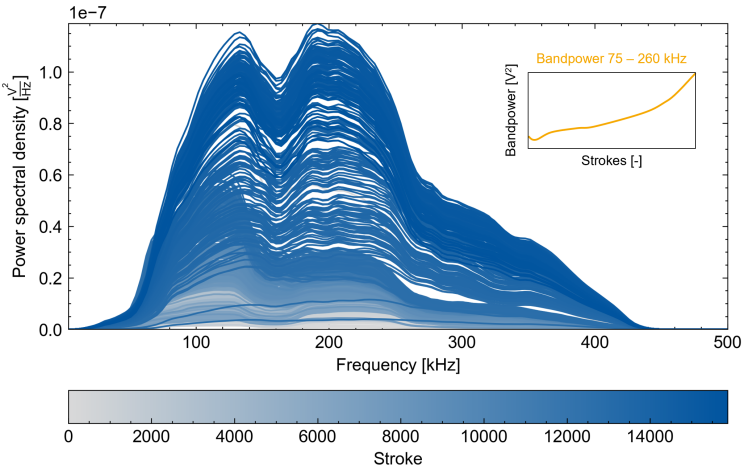



**Figure 6.7:** Evolution of the PSDs of the shearing AE data from stroke series L1 acquired prior to chipping (*segment length* 256, 50 % *segment overlap*, *rolling mean* window size 20).
*Visualisierung der PSD-Entwicklung der Schneidphasen innerhalb der Hubserie L1 vor dem Auftreten von Chipping (Segmentlänge 256, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 20).*

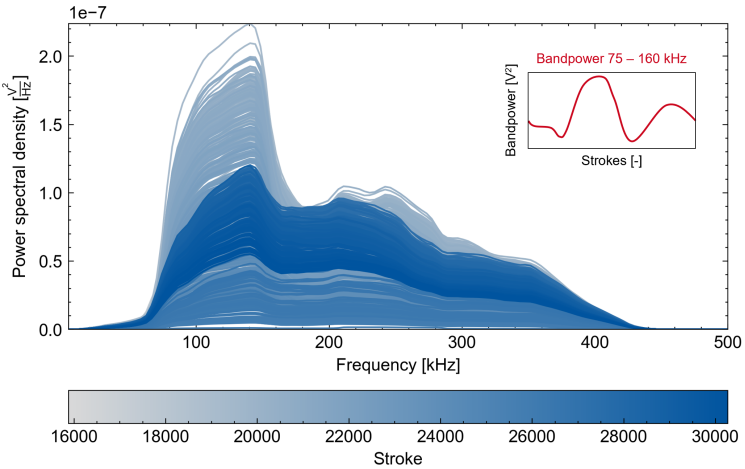



**Figure 6.8:** Evolution of the PSDs of the shearing AE data from stroke series L1 acquired after the occurrence of chipping (*segment length* 256, 50 % *segment overlap*, *rolling mean* window size 20).
*Visualisierung der PSD-Entwicklung der Schneidphasen innerhalb der Hubserie L1 nach dem Auftreten von Chipping (Segmentlänge 256, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 20).*

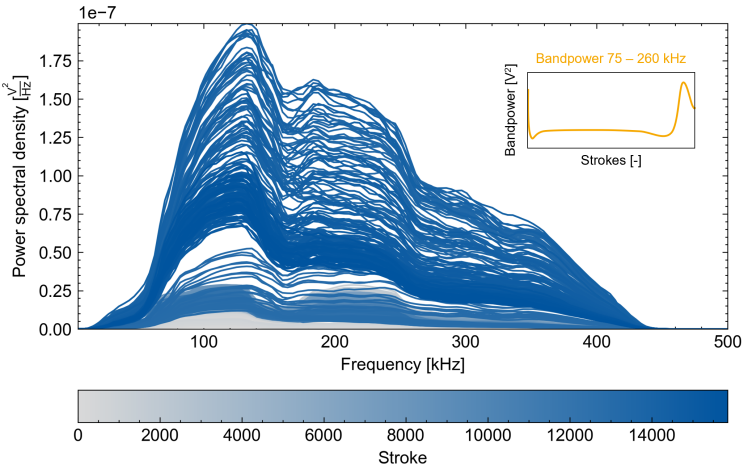



**Figure 6.9:** Evolution of the PSDs of the stripping AE data from stroke series L1 acquired prior to the occurrence of chipping (*segment length* 256, 50 % *segment overlap*, *rolling mean* window size 20).
*Visualisierung der PSD-Entwicklung der Abstreifphasen innerhalb der Hubserie L1 vor dem Auftreten von Chipping (Segmentlänge 256, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 20).*

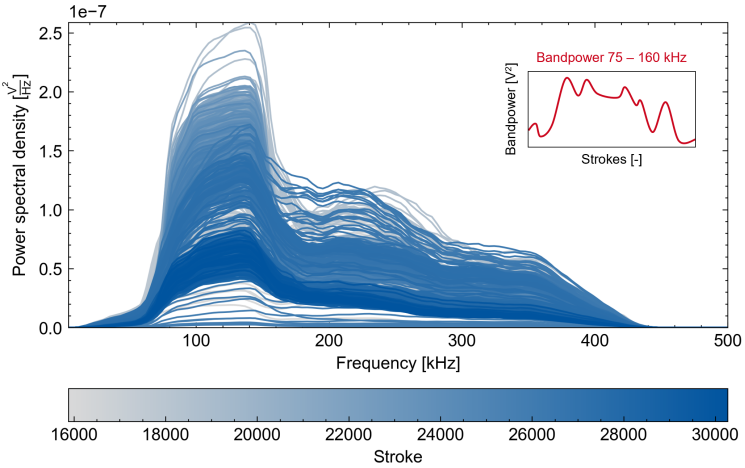



**Figure 6.10:** Evolution of the PSDs of the stripping AE data from stroke series L1 acquired after the occurrence of chipping (*segment length* 256, 50 % *segment overlap*, *rolling mean* window size 20).
*Visualisierung der PSD-Entwicklung der Abstreifphasen innerhalb der Hubserie L1 nach dem Auftreten von Chipping (Segmentlänge 256, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 20).*

### 6.3.2 Results of the preparative experiment L2

*Ergebnisse der Voruntersuchungen L2*

Stroke series L2 represents the dataset with the least amount of recorded process cycles. During the stripping phase of the last stroke of the stroke series, a punch breakage occurred. Since this breakage occurred prior to the 2,500th stroke, no scrap web segments were sampled from the process since an ex-post assignment of scrap web segments to stroke numbers is impossible. Inspection of the punch after breakage revealed traces of severe galling between the punch and the punch guide, see Figure 6.11, which significantly distincts the observed damage from the punch damage observed in stroke series L1 and L3. Several potential reasons were identified that led to the observed failure:

(i) If the *punch was not properly aligned during its assembly*, this could have led to uneven stresses being exerted on it during operation.

(ii) The *presence of metal shavings or debris being dragged into the guide clearance* can increase friction and possibly create a jamming effect, which can contribute to punch breakage.

(iii) If the *punch and its punch guide were not properly matched* or the *guide clearance was uneven at certain locations*, this could lead to increased friction during operation. Increased friction, in turn, potentially caused severe galling between punch and punch guide, ultimately leading to mechanical overloading.

Given the potential issues identified, an explorative analysis of the signal evolution seems promising. While causal attribution is impossible retrospectively, it is to be expected that the identified potential reasons had an impact on the evolution of the AE signal.

**Evolution of the PSDs of L2** *Entwicklung der spektralen Leistungsdichten von L2*

The ABP (see Eq. 47) within the frequency range of 75–160 kHz is visualized in Figure 6.12. Figure 6.13 illustrates the PSDs of every stroke within stroke series L2, ultimately leading up to the punch breakage. The analogous PSD data for the stripping segments is visualized in Figure 6.14.

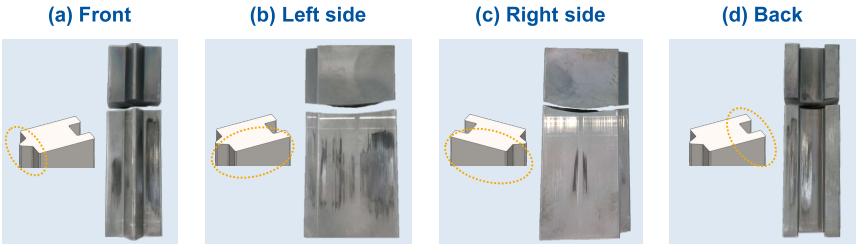



**Figure 6.11:** Fotographs of the broken punch from stroke series L2 from (a) the front, (b) left, (c) right, and (d) back side. The punch shows severe galling between the punch and the punch guide.
*Fotografien des gebrochenen Stempels aus der Hubserie L2 von (a) vorne, (b) links, (c) rechts und (d) hinten. Die Spuren am Stempel deuten auf starkes Fressen zwischen dem Stempel und der Stempelführung hin.*

In comparison to the evolution of the PSDs of the industrial stroke series presented in Section 5.2.2 and the evolution of the PSDs of laboratory stroke series L1 displayed in Section 6.3.1, the evolution of the PSDs of the laboratory stroke series L2 exhibits substantial differences.

Specifically, although the process setup of stroke series L2 is comparable to the process setup of stroke series L1, especially during the first phase of the stroke series, L2 exhibits considerably higher amplitudes, especially in the frequency range 75–160 kHz, both for shearing and stripping. Except for the last stripping segment (visible in Figure 6.14, which represents the signal captured during punch breakage), it is apparent that the first roughly 100 strokes of the stroke series exceed the amplitudes of the remaining strokes substantially).

Calculating the mean of the ABP within the frequency range 75–160 kHz for the shearing segments of L2 yields $4.378 \cdot 10^{-5}\,V^2$, which represents 21-times the ABP of the first 970 shearing segments from stroke series L1 ($2.073 \cdot 10^{-6}\,V^2$). While not quite as pronounced, the comparison between the mean ABP of the stripping signals also reveals a profoundly increased mean bandpower in L2 compared to L1 ($4.117 \cdot 10^{-5}\,V^2$ vs. $6.756 \cdot 10^{-6}\,V^2$). Closer inspection of the ABP in the frequency range 75–160 kHz visualized within Figure 6.12 (note that the $y$-axis is scaled logarithmically for better visibility of changes within the data) reveals approximately five phases ((i) to (v)) of subsequent strokes within stroke series L2 with substantial changes in the bandpower.

Interestingly, the phases (i), (ii), and (iii) are initiated within the stripping phase of the process, while the ABP of the shearing segments shows a lagging effect. However, the phases (iv) and (v) are mainly noticeable in the shearing segments. The actual punch breakage event remarkably undercuts phase (i) in ABP, emphasizing the severity of wear that likely accumulated within phase (i).

The findings reveal that *excessive galling between the punch and the punch guide* occurred during stroke series L2. It is highly likely that the *substantial concentration of power within the frequency range 75–160 kHz* is a result of galling between the punch and the punch guide, both for shearing and for stripping, since the process setup remained nearly unchanged to stroke series L1.

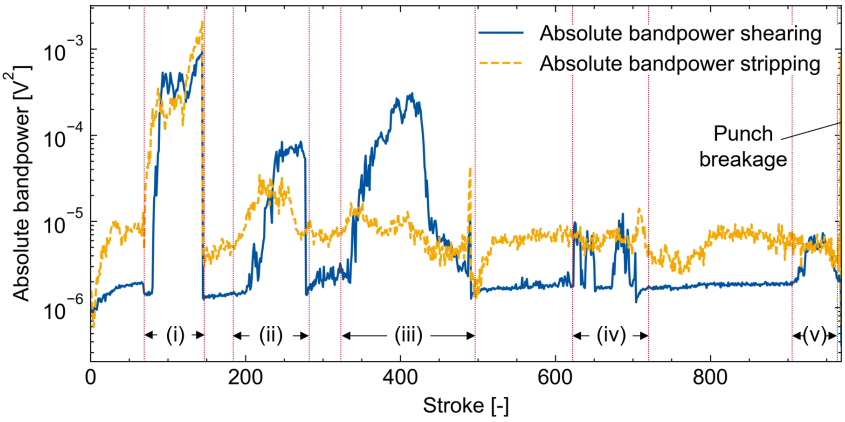



**Figure 6.12:** Visualizaion of the ABP within the frequency range 75–160 kHz of stroke series L2, both for the shearing and the stripping phases.
*Visualisierung der absoluten Bandleistung im Frequenzbereich 75–160 kHz der Hubserie L2, sowohl für die Schneid- als auch für die Abstreifphase.*

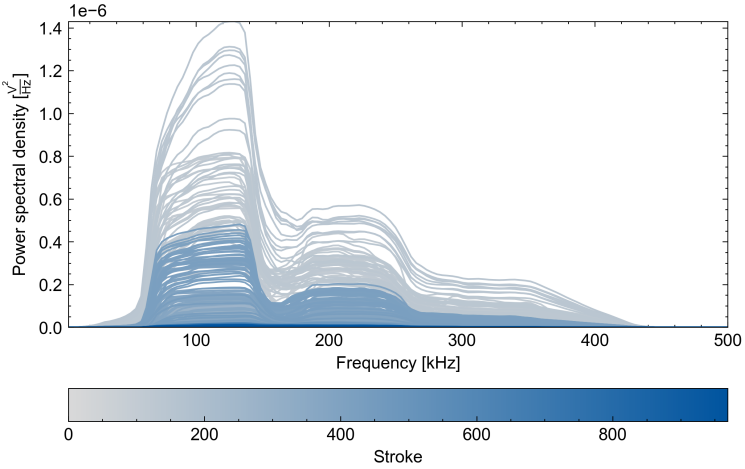



**Figure 6.13:** Evolution of the PSDs of the shearing AE data from stroke series L2 acquired prior to the punch breakage (*segment length* 256, 50 % *segment overlap*, *rolling mean* window size 20).
*Visualisierung der PSD-Entwicklung der Schneidphasen innerhalb der Hubserie L2 vor dem Stempelbruch (Segmentlänge 256, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 20).*

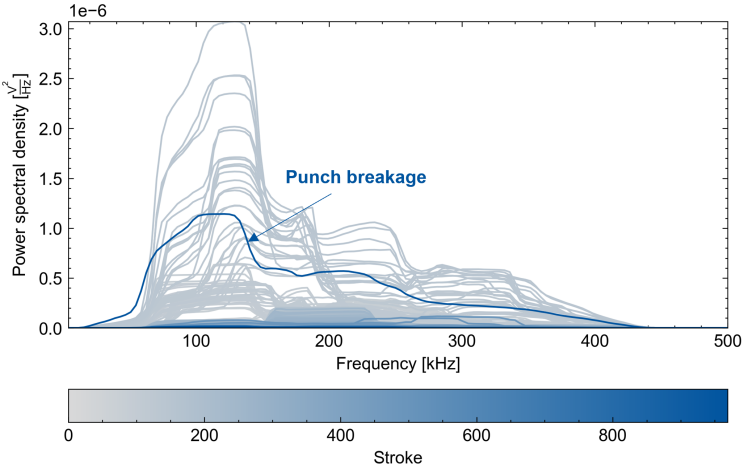



**Figure 6.14:** Evolution of the PSDs of the stripping AE data from stroke series L2 acquired prior to the punch breakage (*segment length* 256, 50 % *segment overlap*, *rolling mean* window size 20).
*Visualisierung der PSD-Entwicklung der Abstreifphasen innerhalb der Hubserie L2 vor dem Stempelbruch (Segmentlänge 256, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 20).*

### 6.3.3 Results of the main experiment L3

*Ergebnisse der Hauptuntersuchungen L3*

Stroke series L3 represents the laboratory stroke series (i) with the *highest amount of metadata* since 203 scrap web segments were sampled and tested, (ii) with the *highest amount of strokes*, and (iii) with the most steady process conditions since no interventions regarding the tool occurred (such as disassembly and subsequent assembly). Figure 6.15 shows the mean roughness values, calculated as outlined in Section 6.2 for all sampled scrap web segments. A visualization of the results distinguished punch-wise is given in Figure 6.16.

Up to roughly the 3,400th stroke, an upward trajectory in the mean surface roughness of the analyzed scrap web surfaces is observed, potentially mirroring the break-in phase of the punch, as illustrated in Figure 2.6. While Figure 6.16 shows that for both punches, the resulting sheared surface roughness of the scrap web segments exhibits a sharp increase, closer inspection of the punch-wise resulting sheared surface roughness and comparison with the averaged mean roughness in Figure 6.15 reveals nuanced both punch- and surface-dependant differences.

Both punches caused a sharp increase in the resulting sheared surface roughness until the approximately 4,000th stroke. Surface 1 Punch 1 continues to rise after a short plateau up until the 8,000th stroke, while Surface 2 Punch 1 even decreases in mean roughness until stroke 8,000. Interestingly, after the initial sharp increase of roughness until stroke 4,000, the surface roughness appears to almost show anti-cyclic behavior surface-wise, i.e., a local maximum is reached for Surface 1 Punch 1 at stroke 9,000. In contrast, a local minimum is reached at stroke 9,000 for Surface 2 Punch 1. The same can be stated at strokes 12,000, 20,000, 25,000, and 30,000. However, towards the end of the stroke series, both surfaces display a steep increase in mean roughness. Since a high amount of metal

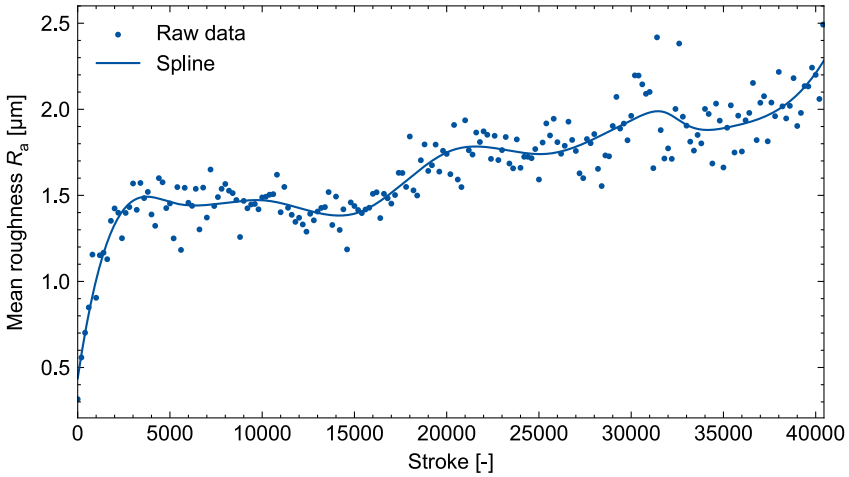



**Figure 6.15:** Visualization of averaged mean scrap web sheared surface roughness and spline interpolation.
*Visualisierung der gemittelten mittleren Oberflächenrauheit der Stanzgitter-Schnittflächen und Spline-Interpolation.*

shavings was observed within the tool, the experiment was ended to prevent damage to the tool components.

The sheared surfaces corresponding to Punch 2 globally exhibit more persistent phases, i.e., not as many transitions as the surfaces that correspond to Punch 1. The trajectories of Surface 1 Punch 2 and Surface 2 Punch 2 between stroke 4,000 and stroke 15,000 are especially interesting. While the mean roughness of Surface 1 shows a slight and stable increase, the mean roughness of Surface 2 decreases substantially until stroke 15,000, which could point to welded adhesions loosening because of the high pressure and temperature. Following, the mean roughness of Surface 2 increases drastically, whereas the mean roughness of Surface 1 remains constant. This increase can also be observed for Surface 1 Punch 1 from stroke approximately 15,000 on. Similar to the surfaces corresponding to Punch 1, the mean roughness of both surfaces corresponding to Punch 2 increases substantially towards the end of the stroke series. A visualization of the initial sheared surfaces of the first and the last sampled scrap web segments is provided in Figure 6.17, clearly depicting the impairment of the sheared surfaces at the end of the stroke series. A more fine-granular view of the evolution of the sheared surface quality is given in Figure A.1 in the Appendix A.2.

Curiously, closer inspection of the AE data reveals that the shearing phase of the 15,401st stroke represents an extreme value. During the tool movement towards the TDC, a high-energy event within the AE signal can identified (see Figure 6.18).

The energy (the sum of all squared signal values within the time domain) registered for this particular shearing phase amounts to 7,997.83 $V^2$, whereas the second highest energy recorded during shearing amounts to an energy of 457.55 $V^2$ at the 39,845th stroke. Given the standard deviation across all 40,450 strokes stands at 53.12 $V^2$, the energy disparity between these two shearing phases amounts to an almost 142-fold standard deviation from the mean. It is plausible that this high-energy event may be the result of a physical and wear-related event, such as crack formation and delamination or minor shearing edge

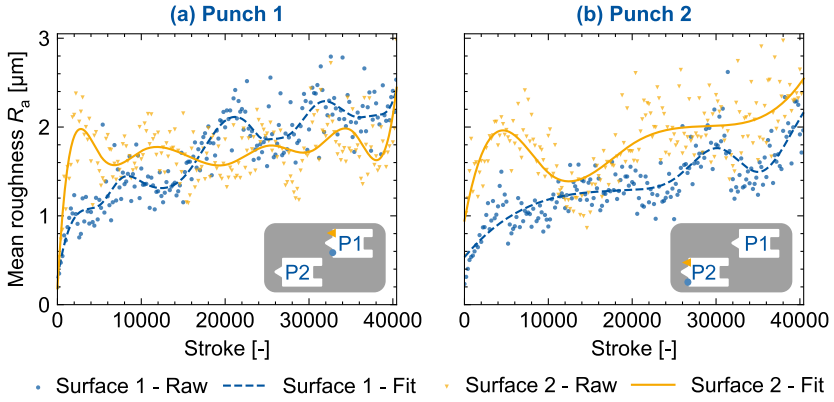



**Figure 6.16:** Depiction of scrap web sheared surface roughness and spline interpolation for (a) Punch 1 and (b) Punch 2.
*Darstellung der Oberflächenrauheit der Stanzgitter-Schnittflächen und Spline-Interpolation für (a) Stempel 1 und (b) Stempel 2.*

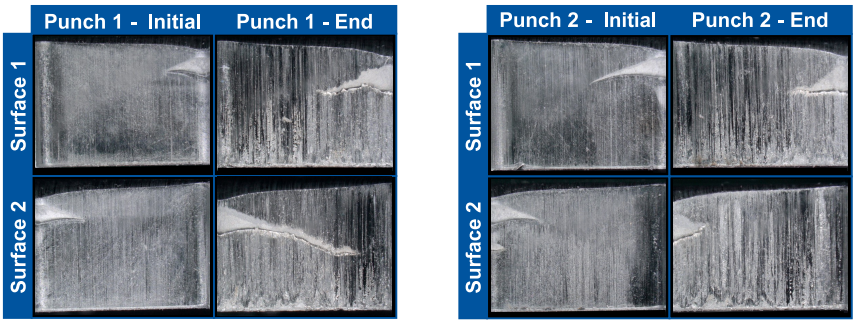



**Figure 6.17:** Comparison of light microscopic images of the scrap web surfaces from the first and the last stroke of the series.
*Vergleich lichtmikroskopischer Aufnahmen der Oberflächen der Stanzgitter-Schnittflächen zum ersten und letzten Hub der Hubserie.*

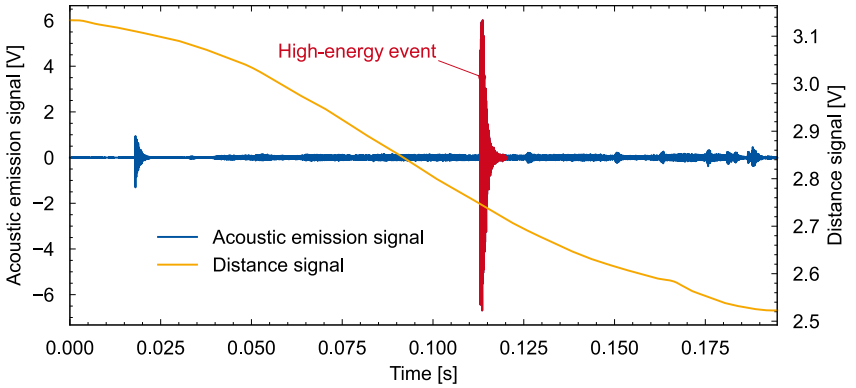



**Figure 6.18:** AE signal for the 15,401st stroke with an observed high energy event during the shearing phase.
*AE-Signal des 15.401. Hubs mit beobachtetem hochenergetischen Ereignis während der Schneidphase.*

breakage, and that this event represents the cause for the subsequent rise in the mean surface roughness by resulting material adhesion.

Qualitative analysis of SEM images revealed severe adhesive wear. Figure 6.19 shows the results of performed *Energy Dispersive X-Ray Analysis* (EDX), verifying adhesive wear through the presence of Vanadium and Chromium, originating from the sheet metal on the surface of Punch 1.

*In summary*, all four examined surfaces exhibit an initial sharp increase in mean surface roughness up to the 4,000th stroke. Surface 1 and Surface 2 of the scrap web segments corresponding to Punch 1 display nearly anti-cyclic behavior in certain intervals, such as around strokes 9,000, 12,000, 20,000, 25,000, and 30,000. Both surfaces conclude with a steep increase in mean roughness. In contrast, the surfaces corresponding to Punch 2 show steadier trends, with Surface 1 gradually increasing and Surface 2 decreasing until around

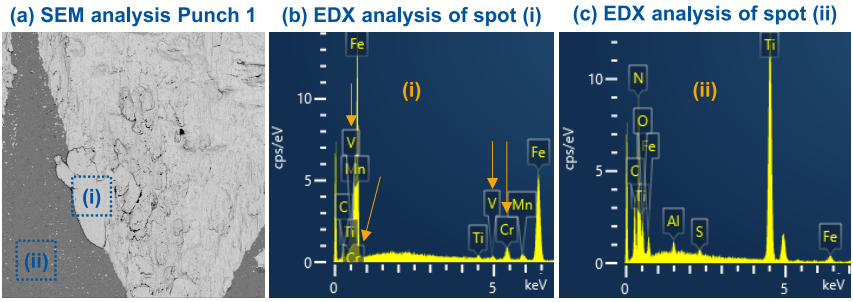



**Figure 6.19:** Depiction of (a) SEM analysis of a section of Surface 1 - Punch 1 and EDX analysis of (b) spot (i) and (c) spot (ii), which reveals material adhesion on the punch surface.
*Darstellung der (a) REM-Analyse eines Ausschnitts von Oberfläche 1 - Stempel 1 und der anschließenden EDX-Analyse von (b) Stelle (i) und (c) Stelle (ii), die Materialanhaftungen auf der Stempeloberfläche erkennen lässt.*

the 15,000th stroke, after which Surface 2 experiences a sharp rise, a pattern also observed for Surface 1 Punch 1. This transition is potentially related to a high-energy event that can be identified at the 15,401st stroke within the shearing segments. While the described trends in the data are clearly observable, the raw data shows stochastic behavior, potentially due to stochasticity in the underlying phenomenon causing the impairment of the sheared surfaces and measurement errors.

As stated in Section 6.1, it is hypothesized that the observed roughness is a direct result of punch wear due to adhesive wear on the punch and resulting galling between the punch and the scrap web. The anti-cyclic behavior observed for the surfaces corresponding to Punch 1 indicates variations in wear patterns, whereas the pronounced increase in roughness at the later stages is indicative of accelerated wear. Similar to the findings documented in Section 6.3.1, the resulting averaged mean roughness generated for stroke series L3, as depicted in Figure 6.15, closely resembles the expected punch wear.

**Evolution of the PSDs of L3** *Entwicklung der spektralen Leistungsdichten von L3*

Although the process setups of stroke series L1 and L3 are comparable, the visualization of the evolution of the PSDs reveals substantial differences between the two stroke series. With an increasing amount of completed process cycles, the power concentrates roughly between 120 and 260 kHz both for shearing and stripping, with an increasing dispersion towards the end of the stroke series, see Figures 6.20 and 6.21. Both lower and higher frequency bands (50–120 kHz and 260–400 kHz) shift towards 120–260 kHz approaching the end of the stroke series. In their work, SHANBHAG et al. identified the frequency range 100–200 kHz as containing the most relevant information about the onset and severity of galling wear [SHAN20]. The shift of the power to 120–260 kHz with increasing sheared surface roughness of the scrap web segments could potentially be attributed to galling between the punch and the scrap web segments.

From the qualitative analysis, it remains unclear *which frequency ranges exhibit the strongest correlation* with the mean sheared surface roughness of the scrap web segments and whether this differs between the shearing and stripping phases. Hence, modeling approaches are needed to examine the correlation between the acquired signals and the developed proxy.

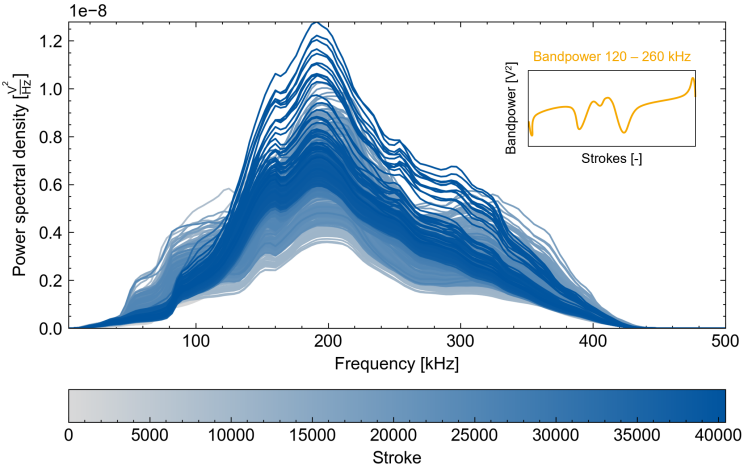



**Figure 6.20:** Evolution of the PSDs of the shearing AE data from stroke series L3 (*segment length* 256, 50 % *segment overlap*, *rolling mean* window size 20).
*Entwicklung der PSDs der Schneidphasen innerhalb der Hubserie L3 (Segmentlänge 256, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 20).*

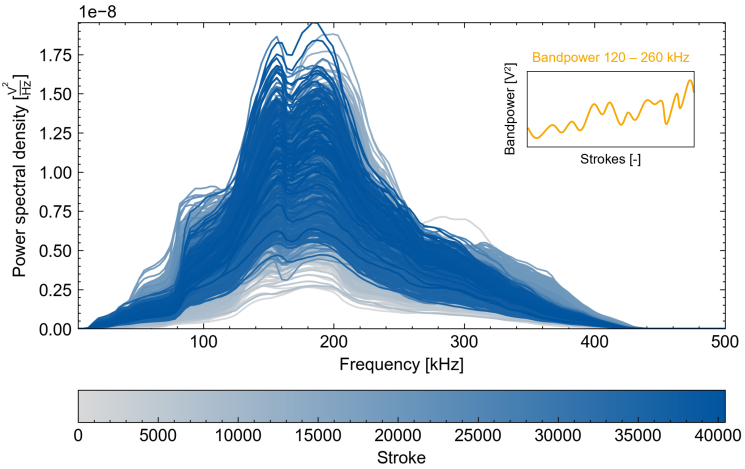



**Figure 6.21:** Evolution of the PSDs of the stripping AE data from stroke series L3 (*segment length* 256, 50 % *segment overlap*, *rolling mean* window size 20).
*Entwicklung der PSDs der Abstreifphasen innerhalb der Hubserie L3 (Segmentlänge 256, 50 % Segmentüberlappung, gleitender Mittelwert mit Fenstergröße 20).*

## 6.4 Intermediate conclusion and unresolved challenges

*Zwischenfazit und offene Herausforderungen*

Research question **RQ3** was constructed as **"How can punch wear information be made available as a proxy variable?"**. Chapter 6 documents the approach to answer this research question.

Since the shearing edges and lateral surface of the punch and the sheared surface of the scrap web represent a tribological system with the lubricant as an intermediate during the shearing and stripping phase [VOIG21], the quality of the sheared surface of the scrap web represents a suitable proxy for the punch wear from a physical point of view. Hence, three different experimental series L1, L2, and L3, were carried out, and scrap web segments were sampled during these experimental series. Subsequently, the roughness of the sampled scrap web segments was evaluated. Additionally, the acquired AE signals during the shearing and stripping phases of the experimental series were analyzed regarding the evolution of their power spectral densities (PSDs) during the shearing and stripping phases. Furthermore, the evolution prior to and after the occurrence of chipping in stroke series L1 and the punch breakage in L2 was analyzed. The punch breakage in stroke series L2 resulted from an unclear issue, causing severe galling between the punch and punch guide, resulting in mechanical overloading. Hence, the punch damage observed in L2 is not comparable to that in L1 and L3. Since the punch breakage in series L2 occurred prior to sampling the first scrap web segment, no roughness data was available for this series.

Stroke series L1 showed an increase in power within the frequency range 75–260 kHz prior to chipping, both within the shearing and stripping phases. After the occurrence of chipping and following punch substitution, an increase in power was observed within the frequency range 75–160 kHz. This effect receded over the course of the remaining stroke series, potentially because of the breaking-in phase of the pristine punch and subsequent receding difference in punch wear states. However, the power of the signals stayed concentrated within this range, both for stripping and shearing.

Interestingly, analysis of the evolution of the PSDs of L2 revealed a high concentration of power within the frequency range 75–160 kHz that receded throughout the stroke series. However, although the process setups are comparable, L2 showed substantially higher amplitude levels than L1, ultimately leading to a remarkably early punch failure caused by excessive galling between the punch and the punch guide. Analysis of the absolute bandpower within the frequency range 75–160 kHz for L2 reveals five phases of substantially elevated absolute bandpower within the shearing segments. Interestingly, the first three of these phases are initiated by the stripping segments, which also show elevated bandpower levels within the first three phases. The actual punch breakage occurred during the stripping phase, as indicated by a substantial increase in bandpower at the last stroke.

The mean roughness of the sheared surfaces of scrap web segments sampled from stroke series L1 and L3 resembled the observed punch wear. Prior to chipping in stroke series L1, a substantial increase in mean roughness is observed, indicating the validity of the mean roughness as a proxy variable for the punch wear. Furthermore, a high energy event during the shearing phase of stroke 15,401 coincided with a following steep increase in mean roughness, possibly caused by a microfracture, delamination, subsequent excessive adhesive wear, and resulting galling between the punch and the scrap web segments. The evolution of the mean roughness within L1 and L3 and punch-wise differentiation shows a punch-dependant evolution and transitory evolution, i.e., phases of increasing roughness are also followed by phases with decreasing roughness or stagnancy. Since material adhesions can also be detached during the shearing and stripping operation and adhesion

was detected using EDX analysis (see Figure 6.19), galling represents a likely explanation for the observed roughness development. The observed concentration of power further strengthens this argument since SHANBHAG et al. found the frequency range 100–200 kHz to carry the most useful information about adhesive wear and subsequent galling between the die and the workpiece during a sheet-metal stamping process.

In contrast to stroke series L1 and L2, the PSDs of L3 showed a more spread-out distribution of power. Towards the end of the stroke series, a concentration of the power within the range 120–260 kHz was observed. Although the same geometry was used and a material grade comparable in strength to series L1 and L2 was processed, the differences could potentially be explained by a different tool being used within series L3 than in series L1 and L2.

Research question **RQ3** was answered by tracking the evolution of the developed proxy variable throughout series L1 and L3. Prior to chipping in series L1, the proxy variable showed a substantial increase. Globally, the proxy variable resembles the expected and observed punch wear, indicating the validity of the approach. Furthermore, the qualitatively identified frequency ranges point to adhesive wear and resulting galling between the punch and the scrap web segments.

Although clear, visually identifiable trends within the PSD evolution and the mean roughnesses of the sheared surfaces of the scrap web segments could be qualitatively identified, the correlation between the mean roughness and the AE signals has yet to be evaluated quantitatively. Hence, Chapter 7 introduces a modeling approach based on machine learning with a subsequent feature importance analysis to analyze potential correlations. The derived knowledge is then applied to the industrial stroke series in Chapter 8.

# 7 Modeling the punch wear proxy variable from acoustic emission data

*Modellierung der Proxy-Variablen für den Stempelverscheiß anhand von Körperschallsignalen*

Chapter 7 presents the modeling approach designed to *predict the punch wear proxy* introduced in Chapter 6, utilizing AE data from stroke series L3. It also aims to *evaluate feature importances*, providing an explainable AI perspective. The modeling and evaluation processes will be collectively referred to as the *data experiment*. This chapter is central to testing the research hypothesis since it involves predicting the punch wear state from AE signals. The insights obtained from the feature importance analysis will subsequently be used for the explorative analysis of the industrial datasets in Chapter 8.

Section 7.1 lays out the *setup for the data experiment* including the *model evaluation approach*. The *results of the data experiment* are documented in Section 7.2.

## 7.1 Data experiment setup

*Experimenteller Aufbau des Datenexperiments*

Following, the aspects of *handling missing data*, *model performance evaluation*, and *feature importance analysis* for the data experiments are presented.

**Handling missing data** *Umgang mit fehlenden Daten*

Sections 6.3.1 and 6.3.3 revealed that the developed proxy variable for the punch wear state develops within certain stages throughout a stroke series. Since it is unknown both *when* and *why* transitions between punch wear states occur *at a certain time*, it is unknown how frequently the proxy variable for the punch wear state needs to be observed to document the punch wear states and transitions between the states. It was decided to choose a sampling interval of 200 strokes for the scrap web segments and subsequently observe the proxy variable since it represents an acceptable trade-off between resolution and the number of measurements. The data was then interpolated (the data points that resulted from this interpolation are followingly referred to as *synthetic data* or *synthetic labels* in the context of supervised ML models).

Two key assumptions were made to justify interpolation between these data points of the proxy variable:

(i) The roughness evolution of the sheared surfaces of the scrap web segments has an *underlying physical causation* and is correlated with punch wear phenomena. Thus, it is not completely stochastic. Subsequently, the missing data can be imputed with suitable models at the cost of introducing further uncertainty.

(ii) The mean of the *actual roughness measurements* for a scrap web segment is, even in the presence of measurement errors, *the most reliable proxy for this physical causation* at a certain stroke available in this work. Therefore, the actual roughness measurements should be present in all models.

---

The content of Chapter 7 was published in [UNTE23]. The author of this work handled the planning, code implementation, setup of the computational resources, and data generation of this work by himself. The coauthors thankfully contributed with impulses towards the structure, presentation of the experimental setup, and presentation of the results within the work. CHATGPT (GPT-4, `https://chatgpt.com/`) was used in this chapter to proofread texts generated by the author and for support with LATEX formatting. DEEPL (`https://www.deepl.com/translator/`) was used for translation purposes.

Different imputation approaches were carried out to reduce the total amount of bias introduced into the subsequently trained models with the selection of an imputation model and to account for stochastic effects in the underlying physical process. The synthetic data was generated with interpolation methods with ascending model complexity:

(i) *Data repetition* (assign roughness measured for scrap web segment 1 to segments 1–200, roughness for segment 200 to segments 201–400, etc.)

(ii) *Linear interpolation*

(iii) *Cubic spline interpolation* (smoothing factor chosen as 0, thus interpolating through each data point)

(iv) *Gaussian process regression*

This approach is illustrated in Figure 7.1. The kernel for the Gaussian process regression (GPR) was chosen as a sum-kernel consisting of a *white noise kernel* and a *Matérn kernel* to introduce measurement noise and model the underlying physical process. A Gaussian process was fit to the roughness measurements with the *GaussianProcessRegressor* class from the sklearn library. The synthetic data was then sampled from the Gaussian process for each model run and used as synthetic labels.

Multiple regression models were trained with a combination of the 203 averaged sample roughnesses for each model run and randomly drawn $s \in \{0; 1{,}000; 2{,}000; 4{,}000; 8{,}000; 16{,}000; 32{,}000; 40{,}247\}$ synthetic labels.

An overview of the full data acquisition, preprocessing, and analysis pipeline is given in the activity diagram depicted in Figure 7.2. Following, the details of this pipeline are discussed.

**Model performance evaluation approach** *Vorgehen zur Evaluation der Modellgüten*

The task of modeling the roughness measurements and the synthetic roughness data from the AE features was treated as a regression problem, specifically to not only model potential punch wear states but also transitions between punch wear states. Prior to selecting the modeling approaches, a list of three criteria for the selection was compiled:

(i) To reduce a potential bias introduced by the choice of a model class, both a *linear* and a *non-linear modeling approach* should be chosen.

(ii) Since no prior feature selection step was performed, both modeling approaches should utilize *regularization* to avoid overfitting.

(iii) To allow a high number of model runs for large datasets, both modeling approaches should be *highly scalable*.

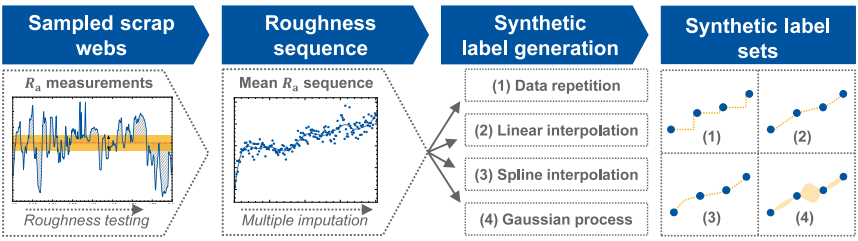



**Figure 7.1:** Developed pipeline for the (synthetic) label generation for the stroke series L3.
*Entwickelte Pipeline für die (synthetische) Label-Erzeugung für die Hubserie L3.*

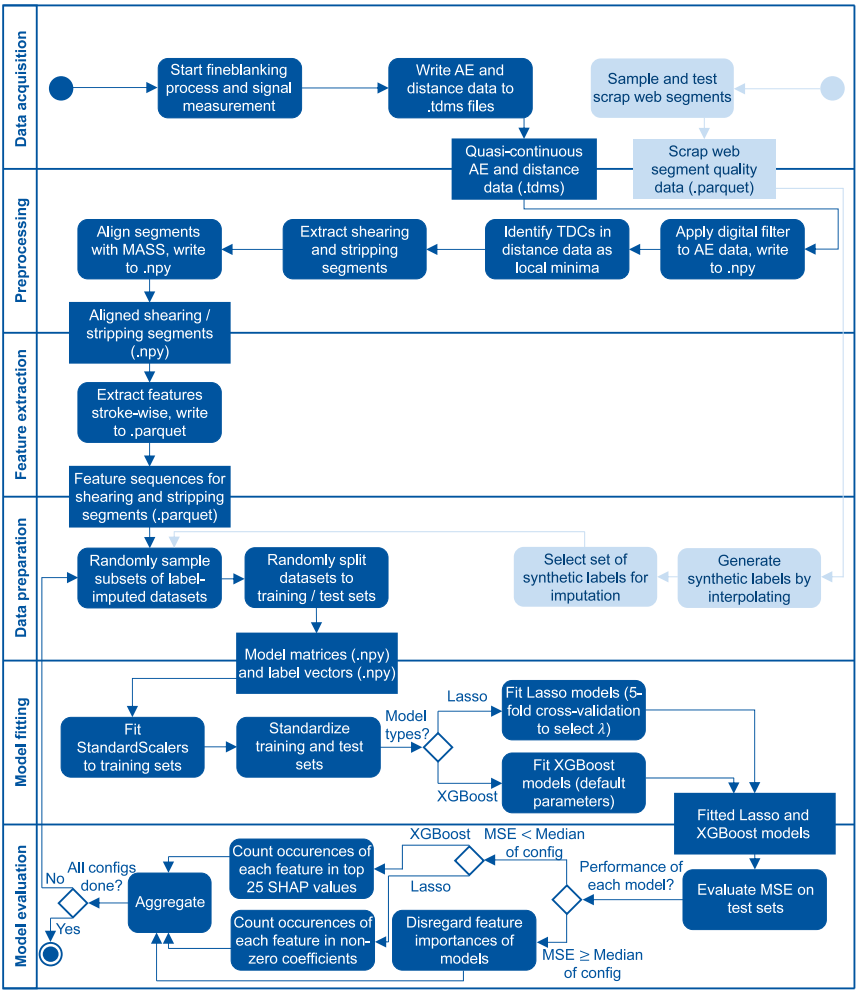



**Figure 7.2:** Activity diagram for the developed pipeline for stroke series L3.
*Aktivitätsdiagramm für die entwickelte Pipeline für Hubserie L3.*

Based on these criteria, two modeling approaches were chosen, namely XGBoost and Lasso regression (refer to Section 2.2 for the theoretical background of these regression approaches).

For every

(i) *dataset size* (e.g., 1,203 labels, consisting of 1,000 synthetic labels and 203 averaged actual measurements),

(ii) *process phase* (e.g., shearing), and

(iii) *imputation method* (e.g., cubic spline),

100 random train/test splits (75 %/25 %) were drawn from 100 randomly sampled subsets of the respective label-imputed datasets. The synthetic labels from the GPR were sampled repeatedly for every model run from the Gaussian process.

Subsequently, the model input was standardized for each model run. A *StandardScaler* object was fit to and then used to transform (standardize) the training set and subsequently transform (without another fitting procedure) the test set to prevent data leakage from the test to the training set. To reduce computation time, the default library parameters were used as documented in [XGBO23] without further hyperparameter optimization. For Lasso regressions, 5-fold cross-validation was carried out with 25 different regularization parameters $\lambda$ between 0.02 and 0.5 for each model run with the *LassoCV* module from sklearn. This range has been determined empirically. The models were trained on the drawn train set and subsequently evaluated for the respective test set regarding the MSE (see Eq. 29). In total, 12,800 models were trained and subsequently evaluated on 6,400 randomly drawn train/test splits.

**Model feature importance evaluation approach** *Vorgehen zur Evaluation der Feature Importance der Modelle*

For Lasso regressions, the feature importance was directly derived by identifying non-zero coefficients and inferring the important features from the trained model for each model run. For the resulting XGBoost models, SHAP values were used (refer to Section 2.2 for the foundations of SHAP).

The following approach was taken for each model run $i \in \{1; 2; 3; \ldots; 100\}$ for a fixed process phase, dataset size, imputation method, and model configuration with $M(\mathrm{MSE}_{\mathrm{model}})$ denoting the median MSE for all model runs within the configuration:

(i) If $\mathrm{MSE}_{\mathrm{XGB}_i} < M(\mathrm{MSE}_{\mathrm{XGB}})$: Extract feature names with the highest 25 SHAP values of the XGBoost model, evaluated on the test set. *Each extracted feature name scores one point.*

(ii) If $\mathrm{MSE}_{\mathrm{Lasso}_i} < M(\mathrm{MSE}_{\mathrm{Lasso}})$: Extract all feature names with non-zero coefficients from the Lasso model. *Each extracted feature name scores one point.*

The scores were calculated across all configurations and used to calculate the relative frequency of each feature that occurs in the most important features. The following Section 7.2 presents the results of the described performance and feature importance evaluation approaches.

## 7.2 Results

*Ergebnisse*

The model performances are presented and discussed in Section 7.2.1, followed by an evaluation of the feature importances in Section 7.2.2.

### 7.2.1 Results of the model performance evaluation

*Ergebnisse der Evaluierung der Modellgüten*

Figure 7.3 visualizes the distribution of MSEs for the shearing segments and the respective examined configurations. Interestingly, the dispersion is exceptionally high for Lasso models trained and evaluated on small datasets for the shearing segments, with multiple training and test set configurations yielding an MSE of more than 10. Therefore, the distribution of MSEs was additionally visualized with a cutoff MSE of 0.05 in Figure 7.5. The distribution of MSEs for the stripping segments is visualized in Figure 7.4.

The dispersion for Lasso models trained and evaluated on smaller stripping datasets is lower than for the shearing segments, with an on-par performance of Lasso with XGBoost for training and test sets sampled from the 203 averaged roughness measurements (number of synthetic data points 0). In contrast, for the stripping segments, the linear models appear to be robust towards the selection of the training and dataset, even for small datasets.

Comparing the absolute MSE values and the trend with increasing dataset sizes, it becomes apparent that XGBoost models outperform Lasso models across all imputation methods. This observation is particularly noteworthy given that no hyperparameter tuning was conducted for the XGBoost models, unlike the Lasso models which were optimized through cross-validation. While the errors evaluated on the test sets of XGBoost models decrease with increasing dataset sizes (and thus both the training set and the test set sizes), the same does not hold for Lasso models. This is further emphasized by comparing the statistical quantities (sample) mean, median standard deviation (SD) for the MSEs across all model runs within a fixed model and imputation method configuration as presented in Table 7.1. These statistical quantities confirm the qualitative findings derived from Figures 7.3 and 7.4: While both modeling approaches appear to be able to predict the test set labels well from the given features, XGBoost outperforms Lasso both in terms of mean and median of the MSE. Furthermore, comparing the mean and median of Lasso models for the shearing segments confirms the identified dispersion in Figure 7.3 since, depending on the imputation method, the mean MSE is 6–8 times higher than the median. This does not hold for the XGBoost models trained and evaluated on the shearing segments nor for the XGBoost and Lasso models trained and evaluated on the stripping segments.

As documented in Section 6.3.3, stroke 15,401 showed exceptionally high signal energy during the shearing phase. Since this stroke was included in every dataset either as a training or test sample (see Section 7.1) and its features showed, upon further investigation, substantially different values from all other strokes, its impact on the model performances was additionally examined by excluding the shearing and stripping phase of this stroke from the datasets and evaluating the resulting model performances.

For shearing, the resulting Lasso models yielded a mean MSE of 0.032 with an SD of 0.006, which represents a substantially better performance than for the datasets that included the shearing segment of stroke 15,401 (mean MSE 0.229, SD 0.952). In contrast, the performance of the XGBoost models for the shearing segments did not change substantially (mean MSE 0.018, SD 0.009 without stroke 15,401 vs. mean MSE 0.019, SD 0.010 with stroke 15,401). Excluding the stripping segment of stroke 15,401 did not influence the performance of Lasso and XGBoost models, indicating that only the shearing segment of stroke 15,401 represents an extreme value. Further investigation revealed that the high MSEs of Lasso models (see Figure 7.3) were caused by the presence of stroke 15,401 in the test sets. Other than the tree-based XGBoost, Lasso models failed to predict within the usual label value range when encountering the substantially different feature values of stroke 15,401 due to their linear structure, which explains the unusually high MSEs in test sets containing this stroke.

From a physical standpoint, while the AE signal measured during the stripping phase only contains information about the friction between the punch and the scrap web, the shearing segments contain, besides the friction between the punch and the scrap web, information about the elastic and plastic deformation of the material. While XGBoost models outperformed Lasso models, both modeling approaches were capable of modeling the relationship between the AE signal features and the scrap web sheared surface roughnesses.

### 7.2.2 Results of the feature importance evaluation

*Ergebnisse der Evaluierung der Feature Importances*

Figure 7.6 presents the 25 features that show the highest relative frequencies (as described in Section 7.1, further denoted as *feature importances*) for the shearing segments across all dataset sizes, interpolation methods, and models that resulted in a better-than-median performance (as in lower-than-median MSE) of a respective configuration. The analogous feature importances for the stripping segments are presented in Figure 7.7. The feature importances for the shearing segments, excluding stroke 15,401, are shown in Figure 7.8. Since the feature importances for the stripping segments did not change after removing stroke 15,401, a further visualization is omitted. The features labeled *FFT ... kHz* represent the *FFT mean coefficient* features from TSFEL.

The identified feature importances suggest that across all model runs, similar features contribute to model decisions with a better-than-median performance, indicating a strong correlation of the contained information with the measured and synthetic roughness values of the examined scrap web surfaces.

Evidently, the information contained in certain FFT mean coefficient features contributed frequently to the predictions of the better-than-median performing models, which interestingly

**Table 7.1:** Mean, median, and standard deviation (SD) of the MSEs across all model runs and dataset sizes for the respective configurations.
*Mittelwert, Median und Standardabweichung (SD) der MSEs über alle Modellläufe und Datensatzgrößen für die jeweiligen Konfigurationen.*

| | Shearing | | | | | |
|---|---|---|---|---|---|---|
| | XGBoost | | | Lasso | | |
| Interpolation | Mean | Median | SD | Mean | Median | SD |
| Repetition | 0.019 | 0.016 | 0.010 | 0.230 | 0.035 | 1.005 |
| Linear | 0.014 | 0.011 | 0.011 | 0.240 | 0.030 | 0.992 |
| Cubic spline | 0.016 | 0.013 | 0.010 | 0.223 | 0.032 | 0.921 |
| GPR | 0.027 | 0.025 | 0.009 | 0.222 | 0.036 | 0.883 |
| | Stripping | | | | | |
| | XGBoost | | | Lasso | | |
| Interpolation | Mean | Median | SD | Mean | Median | SD |
| Repetition | 0.023 | 0.020 | 0.012 | 0.039 | 0.039 | 0.007 |
| Linear | 0.018 | 0.014 | 0.013 | 0.034 | 0.033 | 0.006 |
| Cubic spline | 0.020 | 0.016 | 0.011 | 0.036 | 0.035 | 0.005 |
| GPR | 0.028 | 0.025 | 0.010 | 0.040 | 0.039 | 0.005 |

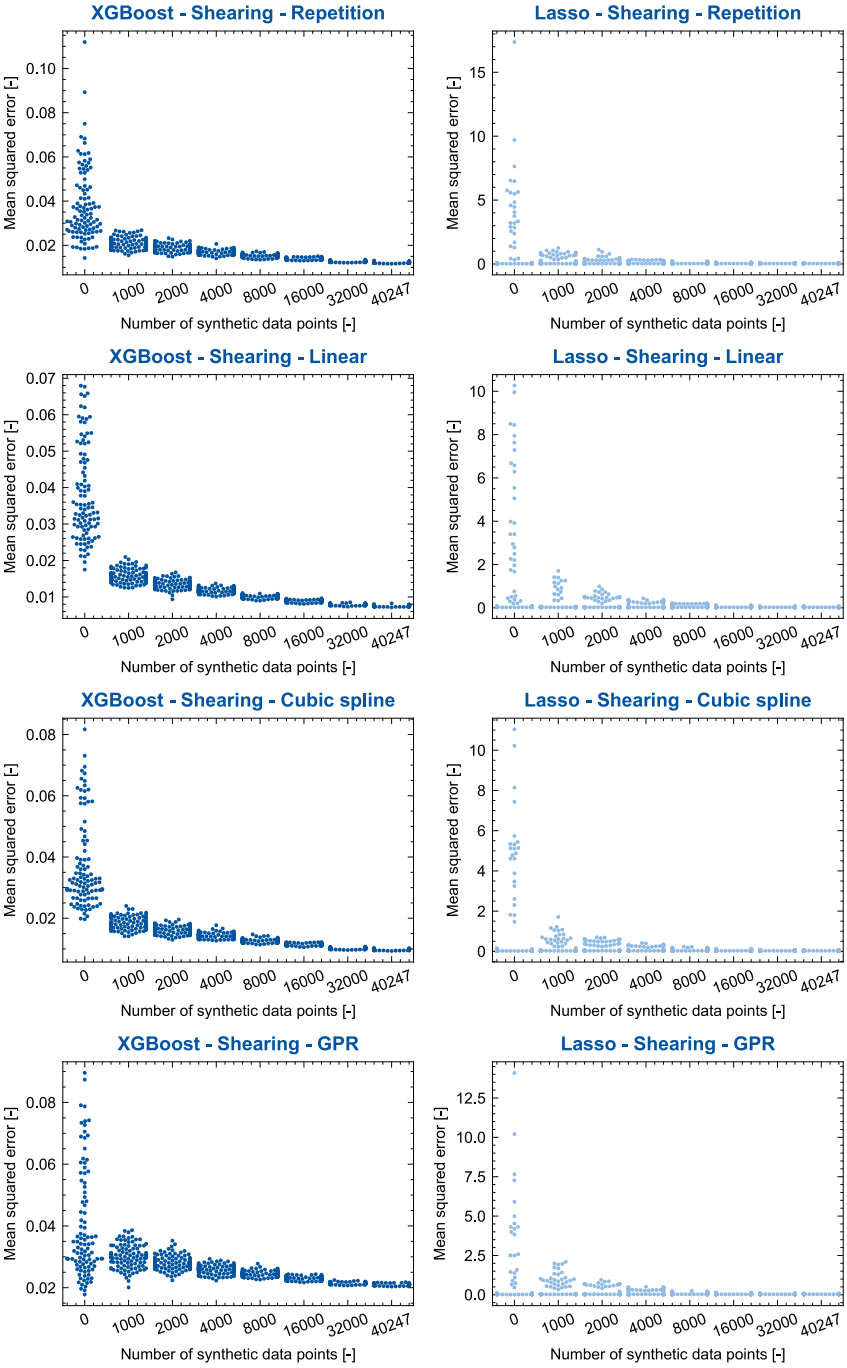



**Figure 7.3:** Comparison MSE shearing across all model runs ([UNTE23]).
*Vergleich des MSE für die Schneidphasen über alle Modell-Durchgänge.*

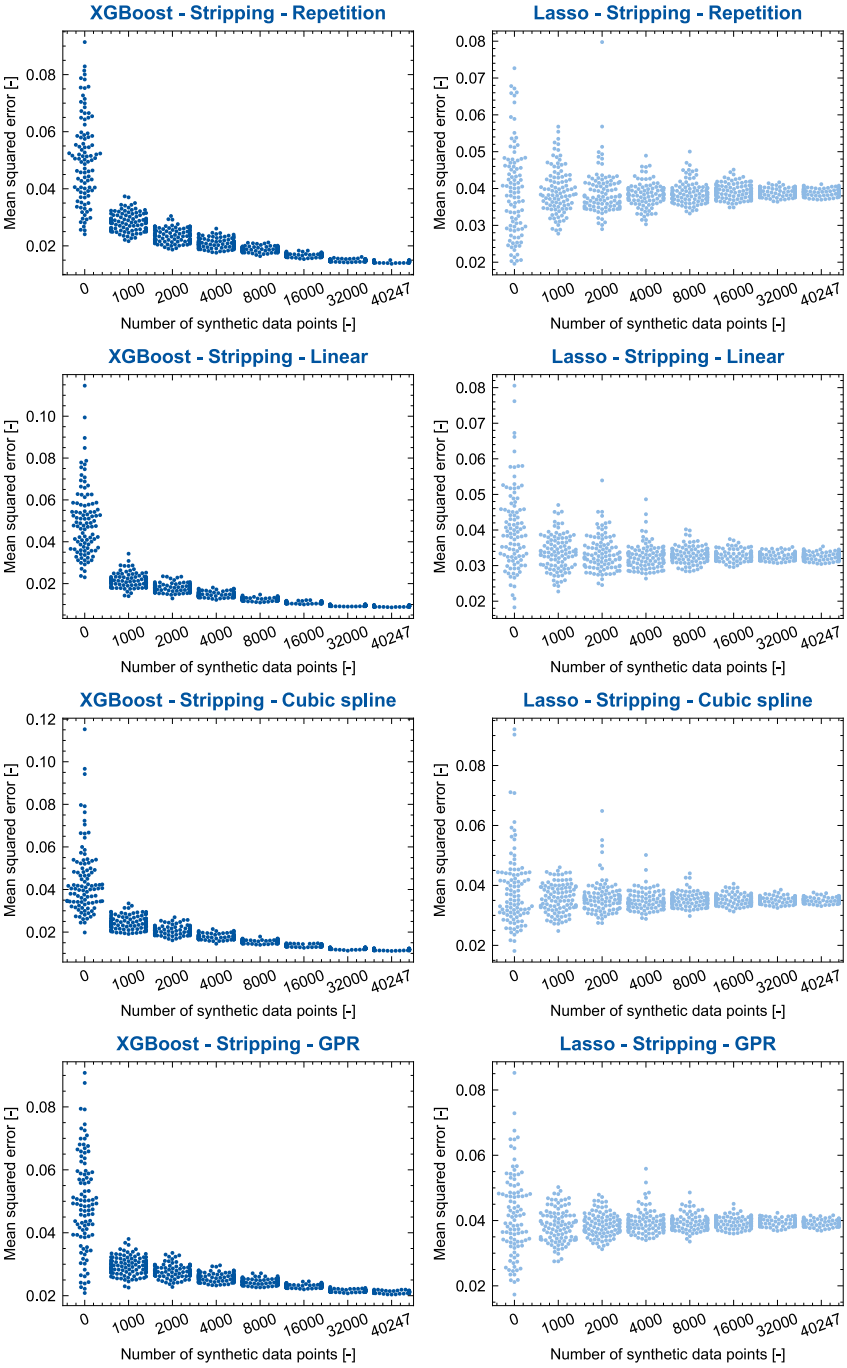



**Figure 7.4:** Comparison MSE stripping across all model runs ([UNTE23]).
*Vergleich des MSE für die Abstreifphasen über alle Modell-Durchgänge.*

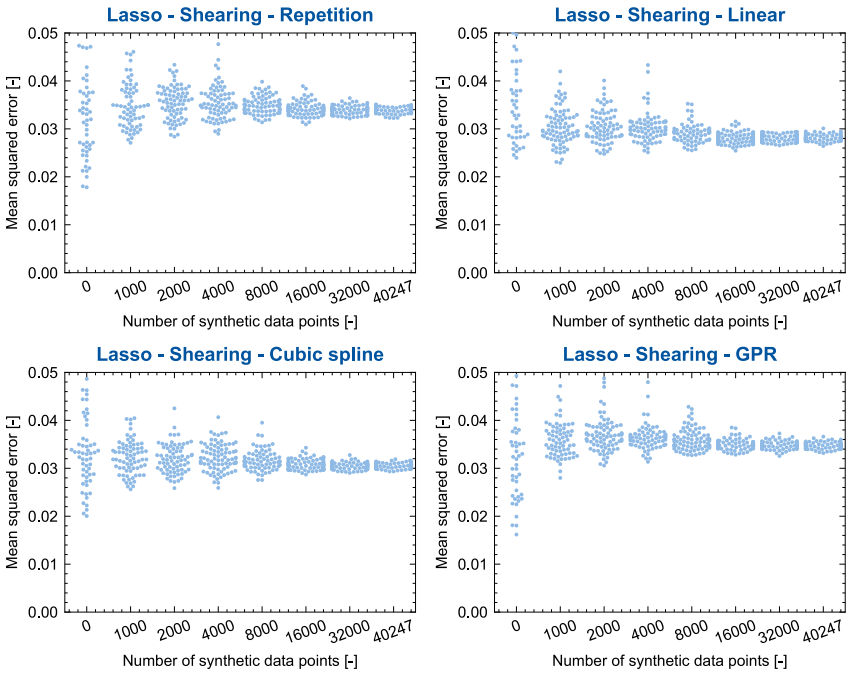



**Figure 7.5:** Limited view Lasso MSE shearing across all model runs ([UNTE23]).
*Eingeschränkte Ansicht der Lasso MSE für die Schneidphasen über alle Modell-Durchgänge.*

holds for both process phases. Therefore, the source code from the TSFEL library for the computation of these features is described here:

(i) Compute a spectrogram of the time series utilizing SciPy's *spectogram* function, yielding a matrix with rows corresponding to frequency ranges and columns corresponding to time ranges and

(ii) calculate the arithmetic mean with NumPy's *mean* function along the columns, i.e., the time ranges, for a given frequency range (matrix row). In total, the frequencies were divided into 256 ranges within this work, yielding 256 FFT mean coefficient features.

While both for shearing and stripping segments frequency ranges such as 64–66 kHz or 257–259 kHz appear, the findings suggest that for both shearing and stripping, the frequency range 103–144 kHz carries the most relevant information to enable model predictions (with a higher-than-median performance) of the scrap web roughnesses. For the datasets that were sampled from all strokes, the FFT mean coefficient for the frequency range 107–109 kHz represents the feature with the highest feature importance across all model runs. Removing stroke 15,401 from the datasets slightly changes the feature importances for the shearing segments. The FFT mean coefficient for the frequency range 107–109 kHz places 5th, whereas the FFT mean coefficients for the frequency ranges 119–121 kHz and 117–119 kHz represent the features with the highest and second highest feature importances.

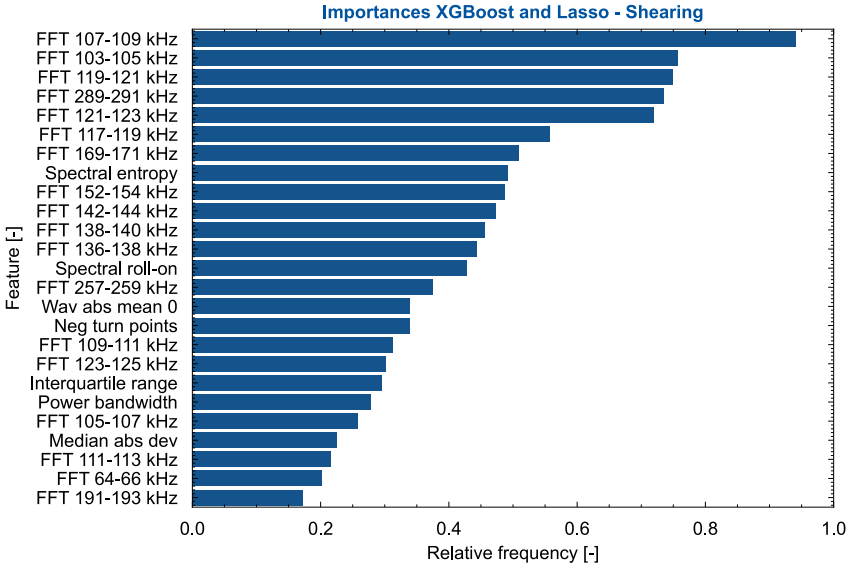



**Figure 7.6:** Global feature importances for the shearing segments ([UNTE23]).
*Globale Feature-Importances für die Schneidphasen.*

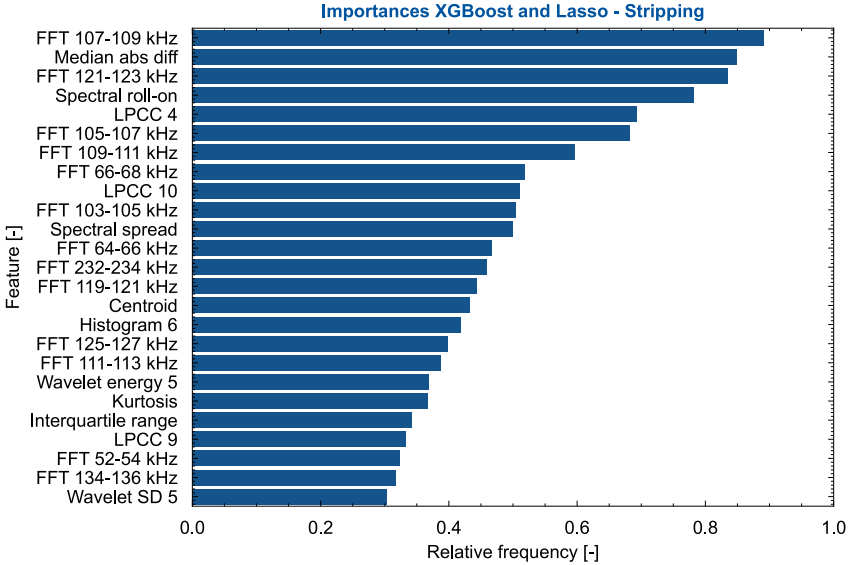



**Figure 7.7:** Global feature importances for the stripping segments ([UNTE23]).
*Globale Feature-Importances für die Abstreifphasen.*

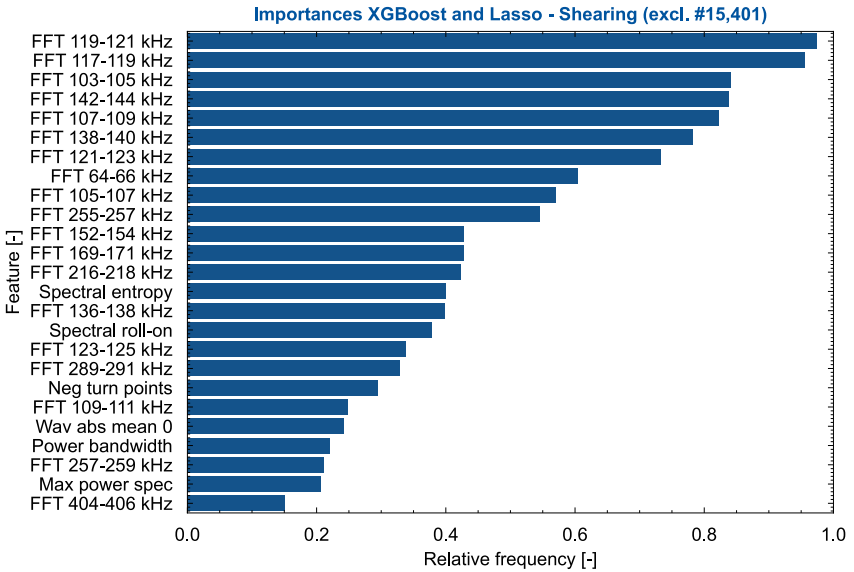



**Figure 7.8:** Global feature importances for the shearing segments (stroke 15,401 removed, ([UNTE23])).
*Globale Feature-Importances für die Schneidphasen (ohne Hub 15.401).*

As laid out in Section 2.4, SHANBHAG et al. experimentally identified the frequency range 100–200 kHz as the frequency range that carries the most important information about the onset and severity of galling between the die and the sheet metal. The identified important frequency range 103–144 kHz points to adhesive wear on the punch and resulting galling as a root cause for the development of the sheared surface roughness of the scrap web segments. This implies substantial adhesive wear on the punch during the presumed break-in phase, see Figure 6.15, and strengthens the argument for the sheared surface roughness of the scrap web segments as a suitable proxy variable for the punch wear state.

While similarities in the feature importance between the shearing and stripping phases can be observed, substantial differences occur, e.g., for the *median absolute difference* (MAD), a measure of statistical dispersion of a signal. The feature MAD displays a high feature importance for the stripping segments. However, it does not appear in the top 25 for the shearing segments.

The importance of the frequency-domain features becomes apparent when evaluating the top 10 features. For stripping, 9 of the 10 most important features are frequency-domain features. For shearing, all 10 of the 10 most important features are frequency-domain features. These findings further emphasize the importance of frequency-domain representations for AE signals.

## 7.3 Intermediate conclusion and unresolved challenges

*Zwischenfazit und offene Herausforderungen*

Research question **RQ4** was postulated as **"How can machine learning models be used to predict the punch wear in sheet-metal stamping from acoustic emission signals and evaluated from an explainable AI perspective to derive feature importances?"**. To answer this research question, a data experiment was conducted to predict the mean roughness of selected sheared surfaces of scrap web segments as a proxy variable for the punch wear state from AE signal features for the shearing and stripping process phases with suitable machine learning models.

In total, 40,450 strokes were completed with two fineblanking punches in initially pristine condition. Scrap web segments were sampled every 200th stroke, and the roughnesses of selected sheared surfaces of these scrap web segments were tested (2,436 cumulated roughness measurements). Different imputation methods in ascending complexity (repetition, linear interpolation, cubic spline interpolation, and Gaussian process regression) were used to account for missing data between two subsequent scrap web segment samples. The AE signals were segmented into shearing and stripping phases, and a comprehensive set of features was extracted for both process phases of each stroke.

Altogether, 12,800 XGBoost and Lasso regression models were trained on randomly sampled and differently sized training sets and evaluated regarding the MSEs of their test set predictions. Both XGBoost and Lasso were able to model the scrap web surface roughnesses from the features of the acoustic emission signals. XGBoost outperformed Lasso for both process phases and across all dataset sizes and imputation methods, even in the absence of hyperparameter tuning for XGBoost, whereas Lasso models were optimized with cross-validation. Lasso models exhibited high MSEs for test sets containing the shearing phase of a stroke with extreme feature values compared to the other strokes (with a label in the usual range) due to their linear structure. Removing said stroke (stroke 15,401) resulted in substantially improved Lasso model performances for the shearing segments, especially for smaller dataset sizes.

Evaluation of feature importances with SHAP for XGBoost models and coefficient examination for Lasso models revealed that information contained in the frequency range 103–144 kHz was most relevant for successful predictions (models with lower-than-median error) for both the shearing and stripping phases. This finding suggests that the development of the sheared surface roughness of the scrap web segments was mainly caused by adhesive punch wear progression and manifested in the acoustic emission signal during both the shearing and stripping phases. It also corresponds to the frequency range that was qualitatively identified to show a substantial evolution within the datasets L1 (Section 6.3.1), L2 (Section 6.3.2), and L3 (Section 6.3.3).

Research question **RQ4** was answered by developing the pipeline for and evaluation of the data experiment. While both the physical background and the identified frequency range point to adhesive wear and subsequent galling between the punch and the scrap web segments as a root cause for the quality impairment of the sheared surfaces, the experimental data presented within this work does not validate this finding with certainty. Future research will need to analyze the punch wear state during the process, i.e., with direct tool condition monitoring, to examine a potential causal relationship between the sheared surface impairment of the scrap web segments and the punch wear state with a suitable experimental and methodological design. Furthermore, an important topic for future research is to examine i) whether there are delayed (lagged) causal effects of the punch wear state on the workpiece quality and if these exist, ii) whether a systematic pattern can be discerned in the size of these lags.

The findings of Chapter 7 will be verified within Chapter 8 with a transfer of the results on the industrial datasets. Specifically, the findings will be utilized to develop and evaluate a scalar indicator for the punch condition.

# 8 Transfer of experimental results to industrial datasets

*Transfer der Ergebnisse auf die Industriedatensätze*

Chapter 8 builds on the findings from Chapters 4, 5, 6, and 7 to develop and evaluate a scalar indicator specifically for assessing the punch condition in the industrial setting. Within the framework of the research hypothesis, Chapter 8 employs both quantitative and qualitative analyses to investigate the state of the punch wear in the industrial datasets.

Specifically, Section 8.1 lays out the *motivation* and the *methodology*. Section 8.2 provides the *results of the applied methodology*. The *intermediate conclusion* is given in Section 8.3.

## 8.1 Motivation and methodology

*Motivation und Methodik*

Ideally, models such as the ones developed in Chapter 7 will be employed in industrial TCM endeavors in the future. However, the models require initial label generation and are potentially not robust to changes in the process settings, hence requiring maintenance, e.g., by retraining. Thus, Chapter 8 focuses on identifying a suitable scalar indicator for the punch wear in the laboratory setting and assessing the evolution of this scalar indicator in the industrial setting.

Scalar indicators for TCM are highly valued in the industry due to their ease of interpretation and comparability with historical data. However, identifying an appropriate scalar indicator for TCM, specifically in the context of sheet-metal stamping, remains a challenge yet to be fully addressed. In light of the absence of a widely adopted scalar TCM indicator it is to be expected that constructing a meaningful indicator requires an aggregation of various information.

While no information on the punch condition was available for the analysis of the industrial datasets, the maintenance planning and capacity planning of the fineblanking press was known to the author and is not a part of the NDA:

(i) The press was used *exclusively for manufacturing a single product*.

(ii) *Two identical tools* were used on the same fineblanking press with *identical punch geometries*.

(iii) If either *a certain stroke threshold* was reached or if *problems with the process and/or the part quality* occurred, the tools were switched and subsequently maintained.

Based on these facts, *two central assumptions* were derived:

(i) Since only one product was manufactured on the press, a tool switch was *either caused by reaching the stroke threshold* or by *worker decision based on problems with the process and/or the part quality*.

(ii) Due to the assumption mentioned above, *stroke series I1, I2, I4, and I5 were all ended because of problems with the process and/or the part quality*. Since series I3 represents the longest stroke series, it could represent the only stroke series that reached the stroke threshold.

---

The results and the methodology presented in Chapter 8 represent unpublished material. CHATGPT (GPT-4, `https://chatgpt.com/`) was used in this chapter to proofread texts generated by the author and for support with LATEX formatting. DEEPL (`https://www.deepl.com/translator/`) was used for translation purposes.

Hence, the stroke series length is indicative of *how problematic or stable a stroke series* was. According to maintenance experts from the company, the tool condition, especially the punch wear, represents a frequent root cause for problematic stroke series. However, the punch wear i) is not the *only* factor for the tool condition and ii) was *not quantitatively evaluated* for the industrial datasets. Still, in light of the lack of metadata about the true tool condition for the industrial data, the findings from Chapter 6 and Chapter 7 were utilized for further explorative analysis.

A selection of 10 potential scalar TCM indicators was chosen based upon the findings of Chapter 5 and Chapter 7. For the features based on PCA, the first principal component was selected with the assumption that relevant information about the wear state explains the largest part of the variance contained in the selected features. The selection is presented in Table 8.1.

Figure 8.1 visualizes the developed pipeline for the evaluation of the TCM indicators. The indicators were analyzed regarding their correlation (specifically their DC, see Section 2.2) with the proxy variable for the punch wear (see Chapter 6).

Since multiple interpolation approaches were carried out as documented in Section 7.1 to account for missing data, the DC was evaluated for every interpolation approach, and the mean of all four interpolated sequences was calculated. The TCM indicator with the highest mean DC was calculated for all industrial stroke series, visualized, and qualitatively evaluated. Furthermore, the dispersion of the TCM indicator was evaluated quantitatively for all industrial stroke series by calculating a cumulated rolling variance. The findings are documented in the following Section 8.2.

**Table 8.1:** Overview of chosen scalar TCM indicators, their definition, and the reference within this work.
*Überblick über die ausgewählten skalaren TCM-Indikatoren, ihre Definition, und Verweise innerhalb dieser Arbeit.*

| TCM indicator | Definition and reference |
|---|---|
| Mean frequency | $\frac{\sum_{k=0}^{k_{\text{Nyquist}}} f_k P_{XX}(k)}{\sum_{k=0}^{k_{\text{Nyquist}}} P_{XX}(k)}$, see Eq. 45 |
| Total bandpower (TBP) | $\sum_{k=0}^{k_{\text{Nyquist}}} P_{XX}(k)$, see Eq. 46 |
| Absolute bandpower (ABP) | $\sum_{k=k_{\text{min}}}^{k_{\text{max}}} P_{XX}(k)$, see Eq. 47 |
| Relative bandpower (RBP) | $\frac{\sum_{k=k_{\text{min}}}^{k_{\text{max}}} P_{XX}(k)}{\sum_{k=0}^{k_{\text{Nyquist}}} P_{XX}(k)}$, see Eq. 48 |
| PCA all | First principal component (PC) of all TSFEL features |
| UMAP all | 1d UMAP embedding of all TSFEL features |
| PCA selected 25 | First PC of top 25 TSFEL features (Figure 7.8) |
| PCA selected 5 | First PC of top 5 TSFEL features (Figure 7.8) |
| UMAP selected 25 | 1d UMAP embedding of top 25 TSFEL features (Figure 7.8) |
| UMAP selected 5 | 1d UMAP embedding of top 5 TSFEL features (Figure 7.8) |

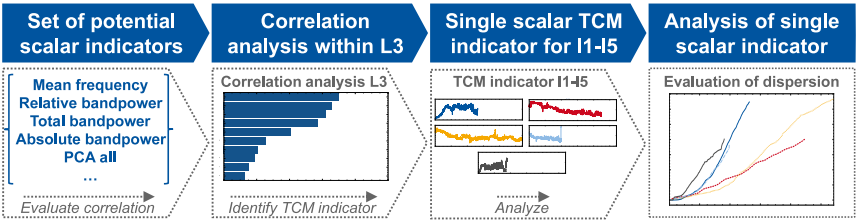



**Figure 8.1:** Depiction of the established pipeline for the development and evaluation of a scalar TCM indicator for the industrial datasets.
*Darstellung der erarbeiteten Pipeline für die Entwicklung und Bewertung eines skalaren TCM-Indikators für die industriellen Datensätze.*

## 8.2 Results
*Ergebnisse*

Figure 8.2 presents the pairwise mean DC of the selected scalar TCM indicators for the shearing segments of laboratory stroke series L3.

Noteworthy, the two TCM indicators with the highest mean DC with the punch wear proxy variable are PCA-based indicators, similar to the findings of NIEMIETZ et al., who compared multiple dimensionality reduction techniques regarding the resulting embeddings of force-displacement signals acquired during a fineblanking process and the observed punch wear [NIEM22].

Embedding only subsets of the features that were identified as most important in Section 7.2.2 substantially increased the mean DC compared to the embeddings of all 390 available features. Interestingly, without prior feature selection, UMAP clearly outperformed PCA, strengthening the argument for UMAP as a strong tool for exploratory data analysis. The

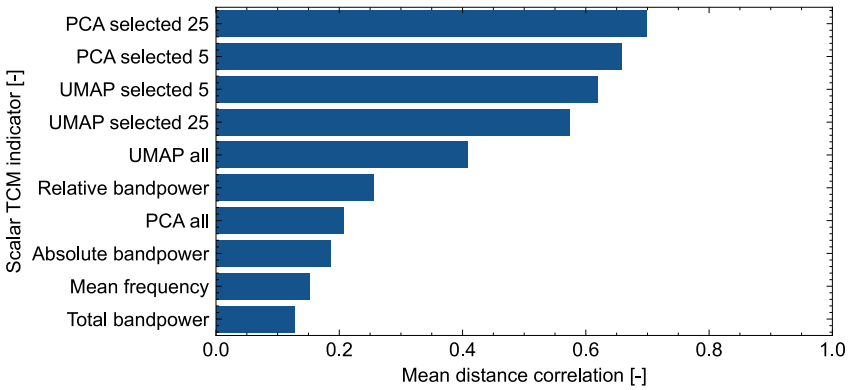



**Figure 8.2:** Pairwise mean DC of selected scalar TCM indicators for the shearing segments of L3 with the interpolated roughness sequences, averaged over all interpolation methods.
*Paarweise mittlere DC ausgewählter skalarer TCM-Indikatoren für die Schneidsegmente aus L3 mit den interpolierten Rauheitssequenzen, gemittelt über alle Interpolationsmethoden.*

fact that UMAP-based indicators underperformed PCA-based indicators with a prior feature selection step could potentially be due to the lack of hyperparameter tuning (hyperparameter tuning was omitted to reduce computational effort). The features RBP and ABP show a substantially lower mean DC with the punch wear proxy variable than the embeddings of the feature subsets, although the information about the important frequency range that was identified in Section 7.2.2 was used for their calculation. However, the lowest correlation is exhibited by the features that utilize no information from the prior analyses, namely mean frequency and TBP. The development of the RBP, however, reveals an interesting dynamic and is visualized in the Appendix A.3 in Figure A.2. The RBP increases for the stroke series I1, I2, I3, and I5 rapidly in the initial stage of the punch lifecycles until a certain saturation level. Interestingly, I4 remains an exception to this and shows a steep decrease and a following steep increase. The steep increases in the RBP for the frequency range 103–144 kHz explains the initial steep decreases of the mean frequency feature as documented in Section 5.2. While this represents an interesting finding, the correlation analysis suggests that a more complex aggregation of information is necessary to model the punch wear dynamics within a scalar indicator.

Since the AE data from the industrial stroke series were acquired with a different sampling rate than the AE data from the laboratory stroke series, five features (features that require a sampling rate above 400 kHz, e.g., *FFT 255-257 kHz*) could not be extracted. Hence, they were omitted for further analysis. Omitting said features only slightly changed the mean DC and did not change the ranking of the indicators. The remaining features were standardized for every industrial dataset and subsequently PCA-transformed. The first principal component (PC1) was extracted and normalized for visualization and comparison purposes. Figure 8.3 depicts the evolution of the PC1 across stroke series I1–I5.

The resulting sequences reveal several aspects. The PC1 evolutions of stroke series I2 and I3, the longest and hence potentially most unproblematic stroke series, show a high degree of similarity and stability, i.e., a low amount of dispersion and a rather clear global trend. This is further highlighted by Table 8.2, which presents the pairwise DC of the PC1 between the stroke series, cropped to the shortest pairwise length.

Stroke series I2 and I3 exhibit the highest pairwise DC with 0.637, followed by stroke series I1 and I3 with 0.566. Stroke series I3 and I5 have the lowest pairwise DC with 0.238.

Figure 8.3 (a), (d), and (e) show a rather high degree of dispersion in stroke series I1, I4, and I5, with abrupt changes at the end of the stroke series. In contrast, stroke series I2 and I3 (Figure 8.3 (b) and (c)) exhibit a stable trend.

Figure 8.4 displays the cumulated rolling variance (window size 100) of the PC1 (not normalized) for the industrial stroke series and confirms the findings mentioned above.

**Table 8.2:** Pairwise DC of PC1 evolution of the 25 highest ranked features across series I1 - I5 (cropped to shortest pairwise length).
*Paarweise DC der Entwicklung der 1. Hauptkomponente der 25 höchstrangigen Features in den Serien I1 - I5 (auf die kürzeste paarweise Länge zugeschnitten).*

| | PC1 series I1 | PC1 series I2 | PC1 series I3 | PC1 series I4 |
|---|---|---|---|---|
| PC1 series I2 | 0.347 | | | |
| PC1 series I3 | 0.566 | 0.637 | | |
| PC1 series I4 | 0.457 | 0.400 | 0.478 | |
| PC1 series I5 | 0.264 | 0.246 | 0.238 | 0.276 |

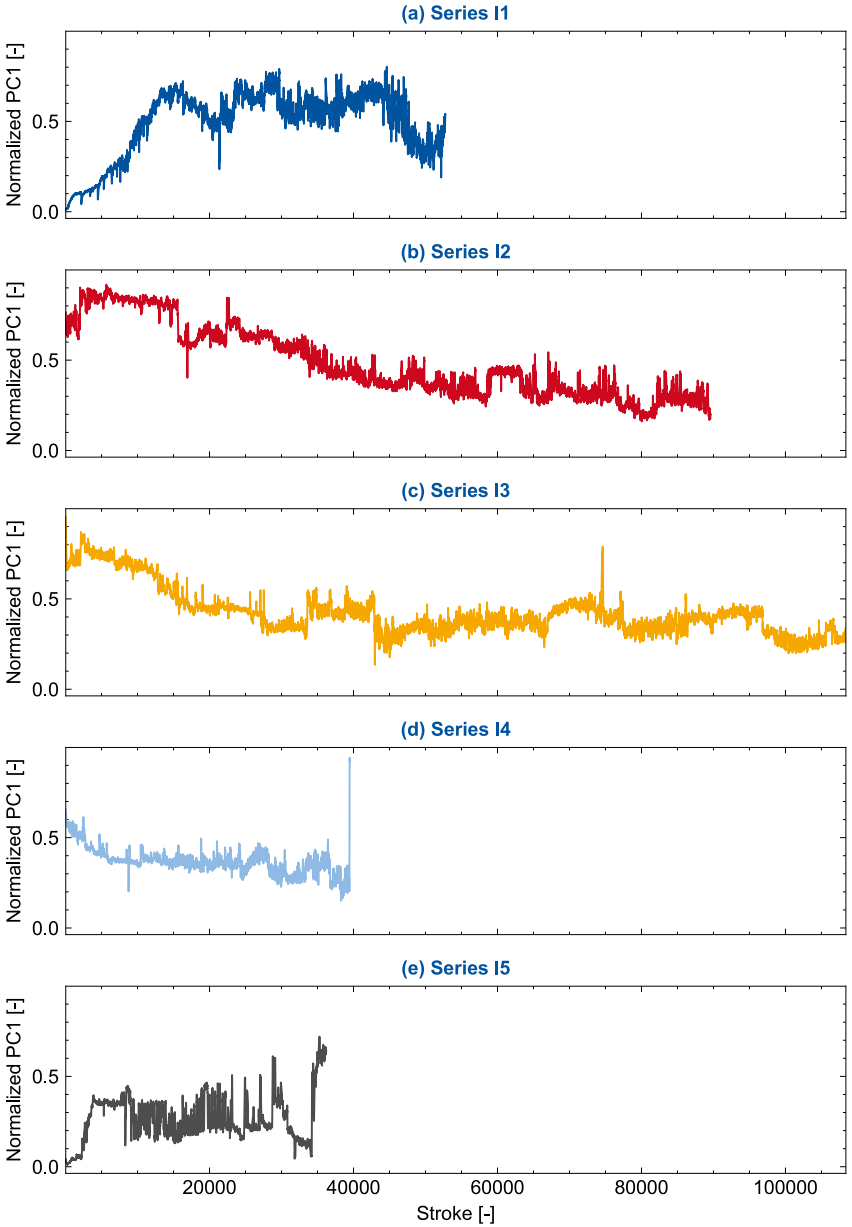



**Figure 8.3:** Rolling mean (window size 10) of the series-wise normalized PC1 of the identified top 25 features.
*Gleitender Mittelwert (Fenstergröße 10) der serienweise normalisierten ersten Hauptkomponente der identifizierten Top-25-Features.*

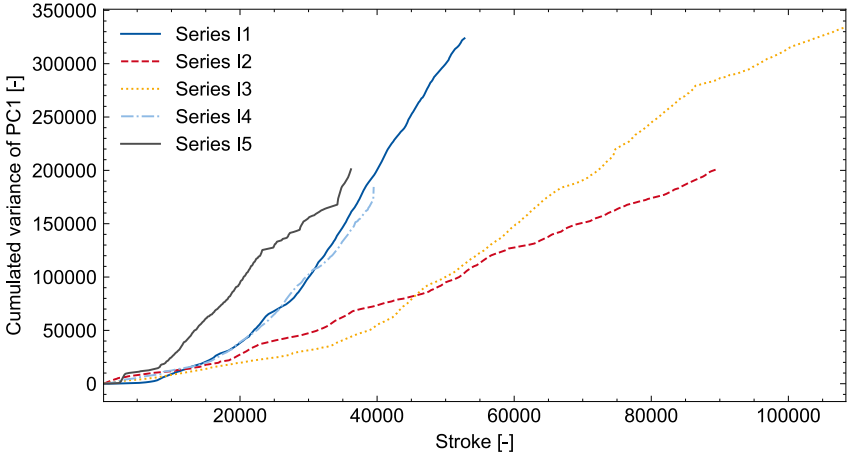



**Figure 8.4:** Cumulated rolling variance (window size 100) of the first principal component of the identified top 25 features.
*Kumulierte gleitende Varianz (Fenstergröße 100) der ersten Hauptkomponente der identifizierten Top-25-Features.*

While the three shortest stroke series (I1, I4, and I5) exhibit notably higher cumulated variance relative to their length, it's important to note that the relationship between stroke series length and cumulated variance is more of a general trend rather than a strict hierarchy. As illustrated in Figure 8.4, while an exact one-to-one correspondence between stroke series length and cumulated variance is not observed, the data does enable a clear differentiation between the potentially problematic and unstable stroke series (I1, I4, and I5), characterized by their high cumulated variance and short lengths, and the more stable stroke series (I2 and I3), which exhibit lower cumulated variance in comparison.

## 8.3 Intermediate conclusion and unresolved challenges

*Zwischenfazit und offene Herausforderungen*

The research question **RQ5** was formulated as **"How can the findings from RQ1–RQ4 be used to develop a scalar indicator for indirect tool condition monitoring in industrial sheet-metal stamping processes?"**. Chapter 8 outlines the steps and methodologies used to answer this research question.

Initially, insights from Chapter 5 (RQ2) and Chapter 7 (RQ4) were leveraged to identify a set of potentially relevant scalar TCM indicators. These indicators were evaluated and ranked based on their mean DC with the four interpolated representations of the proxy variable (refer to Section 7.1, RQ3). The first principal component of the top 25 features, as ranked in Section 7.2.2 (PC1), demonstrated the highest mean DC. Research question RQ1 was indirectly instrumental to this process, as the solutions for this question provided the datasets employed in the subsequent research questions.

A comprehensive visualization of the PC1 evolution within the industrial stroke series unveiled differing patterns among the stroke series of varying lengths. Specifically, the lengthier stroke series I2 and I3 presented a globally rather stable trend, while the shorter stroke series I1, I4,

and I5 displayed substantial dispersion and abrupt shifts towards the latter part of the series. This observation was further accentuated through a visualization of the cumulated rolling variance, which markedly differentiated the longer series from the shorter ones. Assuming the length of a stroke series serves as an indicator of the stability of the series, these results offer promising implications since the features used to calculate PC1 are low in computational cost and could potentially serve as the basis for an online monitoring system.

The successful development of a scalar indicator and the subsequent analysis of its evolution across the five industrial stroke series facilitated a comprehensive answer to **RQ5**. This was achieved by leveraging the findings from RQ1–RQ4. However, it is pertinent to underscore that the selection of the scalar indicator was performed according to its correlation with a proxy for punch wear within a laboratory context, thereby introducing an element of uncertainty. In addition, while the punch wear represents an important part of the tool condition, it is not the sole determining factor, thus necessitating a cautious interpretation of the results. This underscores the essential role of contextual information in future research, such as specific metadata relating to the actual tool condition, in the evaluation of TCM indicators and TCM modeling approaches.

Optimally, industrial contexts would harness machine learning models similar to those developed in Chapter 7, enabling comprehensive real-time tool condition monitoring. Yet, the feasibility of such models hinges on a substantial infrastructure capable of comprehensive process monitoring. Until this infrastructure is sufficiently mature, the scalar indicator proposed in this research presents an interesting preliminary approach to tool condition monitoring in sheet-metal stamping processes. Nonetheless, the long-term objective should be creating and deploying advanced predictive models for superior accuracy and completeness. Therefore, while this scalar indicator offers valuable initial insights and constitutes an important step forward, the process of refining tool condition monitoring solutions is ongoing and should be focused on in future research.

# 9 Summary and outlook

*Zusammenfassung und Ausblick*

**Summary** Sheet-metal stamping processes are among the industry's most frequently used manufacturing processes for the mass production of identical parts due to their resource efficiency. The increasing usage of higher-strength steels leads to increased process forces, friction, and resulting wear of sheet-metal stamping tool components. Excessive wear, in turn, leads to reject production and unplanned downtimes. Since the wear state of critical tool components, such as the punch, is not observable without disassembling the tool, a tool condition monitoring strategy needs to leverage secondary information sources, such as accessible and suitable process signals, to infer the wear state of the tool (indirect tool condition monitoring). Both research and industry approaches to indirect monitoring of the punch condition in sheet-metal stamping under industrial conditions are notably absent.

This work seeks to fill this gap by developing a data-driven strategy for data preprocessing and data analysis of acoustic emission signals acquired during fineblanking processes under industrial conditions to monitor the punch wear by leveraging data mining and machine learning techniques. The work done is based on the research hypothesis that **"With a combination of both qualitative and quantitative analysis of acoustic emission signals, it is possible to model and predict the punch wear state in sheet-metal stamping using suitable machine learning methods."**.

Five research questions were formulated and answered in five successive steps. The work done in the first of five steps focused on answering the research question **"What is the process for designing a domain-specific preprocessing pipeline to isolate relevant process phases from acoustic emission signals in the context of punch wear monitoring in sheet-metal stamping?"**. Based on the physical process principles, the acoustic emission signal emitted during the process phases shearing and stripping was identified as containing relevant information regarding the punch wear. Therefore, a segmentation pipeline was developed to extract these process phases from acoustic emission data generated in a laboratory yet quasi-industrial setting. This was done by i) using a secondary process signal to identify the top dead center of the tool and ii) subsequently aligning resulting segments by leveraging characteristic features of the acoustic emission signal during the stripping phase. The segmentation pipeline was extended for datasets from an industrial production line by introducing an algorithmic error detection mechanism to sort out misegmentations that resulted from a lower signal quality compared to the laboratory setting. To increase the signal-to-noise ratio, the resulting segments were digitally filtered to the frequency response of the used acoustic emission sensor. In total, 398,047 strokes (one opening and closing cycle of the tool) were preprocessed and analyzed for this work. The extracted segments represent the datasets for the following investigations.

Due to the typical sparsity of context information in industrial settings, an explorative analysis of structural changes within the acoustic emission signals was carried out next. The research question **"What recurring patterns and differences can be qualitatively identified in the evolution of preprocessed industrial acoustic emission signals over several punch life cycles?"** was answered by qualitative analysis of

(i) 3-dimensional embeddings of the acoustic emission data,

(ii) the evolution of spectral components,

(iii) and the evolution of the mean frequency feature as a feature previously proposed in research for wear detection in sheet-metal stamping processes

DEEPL (https://www.deepl.com/translator/) was used for translation purposes in this chapter.

throughout several punch lifecycles. It was found that the 3-dimensional embeddings show a clear association between the position of a specific data point in the embedding and the number of strokes carried out before this data point, suggesting a relationship between the acoustic emission signal and the punch wear. Analysis of the spectral components confirmed this finding and revealed a concentration of the power within the frequency ranges 75–115 kHz and 120–160 kHz in the late wear stages. However, it is noteworthy that the dispersion within the power of the spectral components showed substantial differences between the analyzed datasets. Finally, machine standstill information was extracted from a machine logbook, and its association with changes within the mean frequency feature was qualitatively analyzed since this feature has been proposed in the literature to identify (adhesive) wear onset and severity. While all datasets showed an initial decrease in the mean frequency feature, no clear association between increases or decreases in the feature and machine standstills could be identified, and no clear trend in the mean frequency within each dataset could be identified. An association between the dispersion of the feature, which represents a dispersion in the spectral components, and certain machine standstills was, however, presumed.

To overcome the lack of context information, specifically the lack of labels for supervised machine learning models for punch wear prediction, a proxy variable for the punch wear was proposed and evaluated in a laboratory setting in the third step. The research question **"How can punch wear information be made available as a proxy variable?"** was answered by analyzing the evolution of this proxy variable throughout two laboratory experiments and its association with the evolution of the frequency components of the acoustic emission signals. Additionally, acoustic emission data from a third experiment, terminated due to potential tool misconfiguration and punch failure, was evaluated. The evolution of the proxy variable matched the expected and observed punch wear. A clear association was found between the proxy variable and both a chipping event on the punch and the breakage of a punch screw. Additionally, an association between a concentration of power in the frequency components 75–160 kHz and the evolution of the proxy variable was observed. Furthermore, punch failures (breakage and chipping) could be identified prior to their occurrences by a concentration of power in this frequency range.

In the fourth step, a quantitative approach to predict the punch wear (estimated through the proxy variable) from acoustic emission signal features was developed for a dataset. The most extensive investigation from a i) model performance and ii) explainable AI perspective known to the author was conducted to answer the main research question of this work: **"How can machine learning models be used to predict the punch wear in sheet-metal stamping from acoustic emission signals and evaluated from an explainable AI perspective to derive feature importances?"**. By training and evaluating a total of 12,800 machine learning models with different dataset compositions, it was found that both the shearing and the stripping phases contain information that can be used to model the punch wear. Specifically, across all model runs, the components within the frequency range 103–144 kHz were identified as most relevant for accurate predictions. This quantitative finding coincides with the qualitative findings from the prior research questions for both laboratory and industrial datasets. From a physical standpoint and taking prior research into account, the observed evolution of the proxy variable likely resulted from adhesive wear on the punch and resulting galling between the punch and the scrap web.

The findings derived from research conducted to answer the prior research questions were aggregated to develop a scalar tool condition monitoring indicator for the industrial datasets. To this end, the research question **"How can the findings from RQ1–RQ4 be used to develop a scalar indicator for indirect tool condition monitoring in industrial sheet-metal stamping processes?"** was answered by conducting a correlation analysis

for the dataset developed in the fourth step, ranking 10 different scalar indicators. The first principal component of the 25 most relevant acoustic emission signal features of the shearing segments for the punch wear prediction from the third step was identified as having the strongest correlative association with the proxy variable. An analysis of the evolution of the derived indicator revealed that for shorter and, hence, more problematic industrial stroke series, the cumulated rolling variance was substantially higher than for longer and, hence, less problematic stroke series. The findings suggest that the indicator could be used to identify a stroke series as problematic, i.e., identify instabilities before they manifest in excessive machine standstills and unplanned downtimes as a result of early tool changes. This predictive capability would allow companies to plan for expected process instabilities ahead of time, which represents a substantial improvement compared to the current modus operandi.

The research objective of achieving **"[...] a data-driven and explainable model to estimate the punch wear in sheet-metal stamping processes with acoustic emission signals."** was achieved.

**Outlook** Sheet-metal stamping, especially with the increasing use of higher-strength steels, represents a complex and demanding manufacturing process. The innovations and advancements described in this work enable a more intelligent and responsive approach to tool wear monitoring using acoustic emission signals. To pave the way for industrial applicability, the segmentation strategy developed needs to be further optimized, either by using

(i) higher-quality distance sensors,

(ii) alternative secondary sensor signals,

(iii) or a different segmentation approach solely based on the acoustic emission signal

to identify the tool's top dead center and extract the shearing and stripping process phases from the raw acoustic emission signals. These three approaches are not mutually exclusive. A viable strategy could be to train a supervised machine learning model on a sufficient amount of extracted segments derived from a secondary sensor source to identify segments in raw acoustic emission data automatically. While the shapes of the acoustic emission signals differ substantially between process setups (press type, sheet metal thickness, geometry, progressive die, etc.), a sufficiently diverse dataset could potentially overcome these differences. The problem of time series segmentation has been tackled in several domains already, e.g., in manufacturing [GAUG23], activity recognition [XIA20], and sleep staging [PERS19]. Hence, existing architectures should be used and tested on sufficiently diverse datasets to develop reliable segmentation models.

The development of the proxy variable for the punch wear is based on physical principles. The observation of the proxy variable was limited to the most severely worn sections of the examined punches. Future research should focus on examining the causal relationship between punch wear mechanisms and the evolution of the proxy variable, e.g., with suitable methods from the domain of causal inference. To this end, planned machine standstills and direct condition monitoring of the punch, which requires frequent tool disassembly, need to be conducted. Furthermore, a larger portion of the sheared surfaces of the scrap web segments should be examined to generate a more extensive and differentiated representation of the punch wear state.

It is well-known that the mechanical-technological parameters of the processed sheet-metal material vary. Since sheet-metal stamping process parameters are usually set up once and then left unchanged once the process runs stable, i.e., uninterruptedly, this fluctuation of mechanical-technological parameters with constant process parameters influences the

wear development of critical tool components. A combination of approaches that use non-destructive testing of sheet-metal material to infer mechanical-technological material parameters with the developed indirect condition monitoring methodology is promising. A potential future dataflow between material supplier (non-destructive testing) and material processor (acoustic emission data) represents a cooperation that promises substantial economic and ecological advantages over the current treatment of material and wear as black-boxes.

Future research should explore the integration of the developed approach into broader Industry 4.0 frameworks, coupling tool wear insights with other production parameters for a more holistic optimization strategy, and planning downtimes accordingly with intelligent resource allocation. The developed methodology is not limited to sheet-metal stamping processes and might be extended to other clocked sheet-metal forming processes, enhancing the universality of the approach.

# Zusammenfassung und Ausblick

**Zusammenfassung** Stanzverfahren gehören aufgrund ihrer Ressourceneffizienz zu den in der Industrie am häufigsten eingesetzten Fertigungsverfahren für die Massenproduktion von Blechteilen. Der zunehmende Einsatz von höherfesten Stählen führt zu erhöhten Prozesskräften, Reibung, und daraus resultierendem Verschleiß von Stanzwerkzeugkomponenten. Übermäßiger Verschleiß wiederum führt zu Ausschussproduktion und ungeplanten Stillstandszeiten. Da der Verschleißzustand kritischer Werkzeugkomponenten, wie z. B. des Stempels, ohne Demontage des Werkzeugs nicht beobachtet werden kann, muss eine Strategie zur Überwachung des Werkzeugzustands sekundäre Informationsquellen nutzen, wie z. B. geeignete Prozesssignale, um auf den Verschleißzustand des Werkzeugs zu schließen (indirekte Werkzeugzustandsüberwachung). Sowohl in der Forschung als auch in der Industrie gibt es bisher keine Ansätze zur indirekten Überwachung des Stempelzustands unter Industriebedingungen beim Stanzen.

Die vorliegende Arbeit versucht diese Lücke zu schließen, indem sie eine auf Data Mining und maschinellem Lernen basierende datengetriebene Strategie zur Datenvorverarbeitung und Datenanalyse von Körperschallsignalen entwickelt, die während Feinschneidprozessen unter industriellen Bedingungen zur Überwachung des Stempelverschleißes erfasst wurden. Die Arbeit basiert auf der Forschungshypothese: **"Mit einer Kombination aus qualitativer und quantitativer Analyse von Körperschallsignalen ist es möglich, den Zustand des Stempelverschleißes beim Stanzen mit Hilfe geeigneter maschineller Lernverfahren zu modellieren und vorherzusagen."**.

Es wurden fünf Forschungsfragen formuliert und in fünf aufeinanderfolgenden Schritten beantwortet. Die Arbeit im ersten von fünf Schritten konzentrierte sich auf die Beantwortung der Forschungsfrage **"Was sind die Schritte bei der Entwicklung und Implementierung einer systematischen, domänenspezifischen Vorverarbeitungsstrategie für Körperschall-Signale, um relevante Prozessphasen im Zusammenhang mit der Überwachung des Stempelverschleißes beim Stanzen zu extrahieren?"**. Basierend auf den physikalischen Prozessprinzipien wurde postuliert, dass das akustische Emissionssignal, das während der Prozessphasen Schneiden und Abstreifen entsteht, relevante Informationen über den Stempelverschleiß enthält. Daher wurde eine Segmentierungspipeline entwickelt, um diese Prozessphasen aus Körperschallsignalen zu extrahieren, die in einer quasi-industriellen Laborumgebung erzeugt wurden. Dazu wurde i) ein sekundäres Prozesssignal verwendet, um den oberen Totpunkt des Werkzeugs zu identifizieren, und ii) die sich daraus ergebenden

Segmente durch Ausnutzung charakteristischer Merkmale des Körperschallsignals während der Abstreifphase zueinander ausgerichtet. Die Segmentierungspipeline wurde für Datensätze aus einer industriellen Produktionslinie erweitert, indem ein algorithmischer Fehlererkennungs-Mechanismus zum Aussortieren von Fehlsegmentierungen eingeführt wurde, die aus einer geringeren Signalqualität im Vergleich zur Laborumgebung resultieren. Um das Signal-Rausch-Verhältnis zu erhöhen, wurden die resultierenden Segmente digital auf den Frequenzgang des verwendeten Körperschall-Sensors gefiltert. Insgesamt wurden für diese Arbeit 398.047 Hübe (ein Öffnungs- und Schließzyklus des Werkzeugs) vorverarbeitet und analysiert. Die extrahierten Segmente stellen die Datensätze für die folgenden Untersuchungen dar.

Aufgrund der typischerweise kaum vorhandenen Kontextinformationen in industriellen Umgebungen wurde als nächstes eine explorative Analyse der strukturellen Veränderungen innerhalb der akustischen Emissionssignale durchgeführt. Die Forschungsfrage **"Welche wiederkehrenden Muster und Unterschiede lassen sich qualitativ in der Entwicklung der vorverarbeiteten industriellen Körperschall-Signale über mehrere Stempel-Lebenszyklen erkennen?"** wurde durch eine qualitative Analyse von

(i) 3-dimensionale Embeddings der Körperschallsignale,

(ii) die Entwicklung der spektralen Komponenten,

(iii) und die Entwicklung des Features Mean Frequency als ein Feature, das zuvor in der Forschung zur Verschleißerkennung bei Stanzprozessen untersucht wurde

über den Verlauf mehrerer Stempel-Lebenszyklen beantwortet. Es wurde festgestellt, dass die 3-dimensionalen Embeddings einen eindeutigen Zusammenhang zwischen der Position eines bestimmten Datenpunktes im Embedding und der Anzahl der vor diesem Datenpunkt durchgeführten Hübe aufweisen, was auf eine Beziehung zwischen dem Körperschallsignal und dem Stempelverschleiß schließen lässt. Die Analyse der spektralen Komponenten bestätigte dieses Ergebnis und zeigte eine Konzentration der Leistung in den Frequenzbereichen 75–115 kHz und 120–160 kHz in den späten Verschleißphasen. Interessanterweise zeigte sich, dass die Streuung innerhalb der Leistung der Spektralkomponenten erhebliche Unterschiede zwischen den analysierten Datensätzen aufwies. Schließlich wurden Informationen über Maschinenstillstände aus einem Maschinenlogbuch extrahiert und ihr Zusammenhang mit Veränderungen innerhalb des Features Mean Frequency qualitativ analysiert, da dieses Feature in der Literatur zur Identifizierung von (adhäsivem) Verschleißbeginn und -schweregrad untersucht und vorgeschlagen wurde. Während alle Datensätze einen anfänglichen Abfall im Mean Frequency Feature zeigten, konnte kein eindeutiger Zusammenhang zwischen einem Anstieg oder Abfall des Features und Maschinenstillständen festgestellt werden, und es konnte kein eindeutiger Trend bei dem Feature innerhalb jedes Datensatzes ermittelt werden. Allerdings wurde ein Zusammenhang zwischen der Streuung des Features (Streuung innerhalb der Spektralkomponenten) und bestimmten Maschinenstillständen vermutet.

Um den Mangel an Kontextinformationen, insbesondere den Mangel an Labels für überwachte maschinelle Lernmodelle zur Prädiktion des Stempelverschleißes, zu überwinden, wurde im dritten Schritt eine Proxy-Variable für den Stempelverschleiß vorgeschlagen und in einer Laborumgebung evaluiert. Die Forschungsfrage **"Wie können Informationen über den Stempelverschleiß als Proxy-Variable verfügbar gemacht werden?"** wurde beantwortet, indem die Entwicklung dieser Proxy-Variable im Verlauf von zwei Laborexperimenten und ihre Verbindung mit der Entwicklung der Frequenzkomponenten der Körperschallsignale analysiert wurde. Zusätzlich wurden die Körperschallsignale eines dritten Laborexperimentes ausgewertet, das aufgrund einer möglichen Fehlkonfiguration des Werkzeugs und eines

daraus resultierenden Stempelversagens abgebrochen wurde. Es wurde festgestellt, dass die Entwicklung der Proxy-Variablen dem erwarteten und beobachteten Stempelverschleiß entspricht. Es wurde auch ein eindeutiger Zusammenhang mit aufgetretenem Chipping am Stempel und einem Bruch der Stempelschraube festgestellt. Darüber hinaus wurde ein Zusammenhang zwischen einer Konzentration der Leistung in den Frequenzkomponenten 75–160 kHz und der Entwicklung der Proxy-Variablen festgestellt. Außerdem konnten Stempelversagen (Bruch und Chipping) durch eine Leistungskonzentration in diesem Frequenzbereich vor ihrem Auftreten erkannt werden.

Im vierten Schritt wurde ein quantitativer Ansatz zur Prädiktion des Stempelverschleißes (geschätzt durch die Proxy-Variable) aus Features von Körperschallsignalen für einen Datensatz entwickelt. Zur Beantwortung der Hauptforschungsfrage **"Wie können maschinelle Lernmodelle zur Prädiktion des Stempelverschleißes beim Stanzen aus Körperschall-Signalen verwendet und aus einer erklärbaren KI-Perspektive ausgewertet werden, um Feature Importances abzuleiten?"** wurde die umfangreichste dem Autor bekannte Untersuchung aus der Perspektive i) der Modellgüte und ii) der erklärbaren KI durchgeführt. Durch Training und Auswertung von insgesamt 12.800 maschinellen Lernmodellen mit unterschiedlichen Datensatzzusammensetzungen wurde festgestellt, dass sowohl die Körperschallsignale der Schneid- als auch der Abstreifphasen Informationen enthalten, die zur Modellierung des Stempelverschleißes verwendet werden können. Insbesondere wurden über alle evaluierten Modelle hinweg die Komponenten im Frequenzbereich von 103-144 kHz als am relevantesten für genaue Prädiktionen identifiziert. Dieses Ergebnis deckt sich mit den qualitativen Erkenntnissen aus den vorangegangenen Forschungsfragen sowohl für Labor- als auch für Industriedatensätze. Aus physikalischer Sicht und unter Berücksichtigung früherer Forschungsarbeiten ist die beobachtete Entwicklung der Proxy-Variablen wahrscheinlich auf die Adhäsivverschleiß am Stempel und daraus resultierendes Fressen zwischen Stempel und Stanzgitter zurückzuführen.

Die Ergebnisse der zur Beantwortung der vorherigen Forschungsfragen durchgeführten Untersuchungen wurden aggregiert, um einen skalaren Indikator zur Überwachung des Werkzeugzustands für die industriellen Datensätze zu entwickeln. Zu diesem Zweck wurde die Forschungsfrage **"Wie können die Erkenntnisse aus den Forschungsfragen RQ1–RQ4 genutzt werden, um einen skalaren Indikator für die indirekte Überwachung des Werkzeugzustands in industriellen Stanzprozessen zu entwickeln?"** durch die Durchführung einer Korrelationsanalyse für den im vierten Schritt entwickelten Datensatz beantwortet, in der 10 verschiedene skalare Indikatoren verglichen wurden. Die erste Hauptkomponente der 25 relevantesten Körperschallsignal-Features der Schneidsegmente für die Stempelverschleißvorhersage aus dem dritten Schritt zeigte den stärksten korrelativen Zusammenhang mit der Proxy-Variable. Eine Analyse der Entwicklung dieses Indikators für die industriellen Hubserien ergab, dass für kürzere und damit problematischere Hubserien die kumulierte (gleitende) Varianz wesentlich höher war als für längere und damit weniger problematische Hubserien. Die Ergebnisse deuten darauf hin, dass der Indikator dazu verwendet werden kann, eine Hubserie als problematisch zu identifizieren, d. h. Prozessinstabilitäten zu identifizieren, bevor sie sich in übermäßigen Maschinenstillständen und ungeplanten Ausfallzeiten infolge eines frühzeitigen Werkzeugwechsels manifestieren. Dies würde es den Unternehmen ermöglichen, zu erwartende Prozessinstabilitäten zu antizipieren, was eine wesentliche Verbesserung gegenüber der derzeitigen Arbeitsweise darstellt.

Das Forschungsziel **"[...] ein datengesteuertes und erklärbares Modell zur Abschätzung des Stempelverschleißes in Blechstanzprozessen mit Schallemissionssignalen."** wurde erreicht.

**Ausblick** Die Blechumformung ist, insbesondere mit dem zunehmenden Einsatz höherfester Stähle, ein komplexer und anspruchsvoller Fertigungsprozess. Die in dieser Arbeit beschrie-

benen Innovationen ermöglichen einen intelligenteren und reaktionsfähigeren Ansatz zur Überwachung des Werkzeugverschleißes unter Verwendung von Körperschallsignalen. Um den Weg für die industrielle Anwendbarkeit zu ebnen, muss die entwickelte Segmentierungsstrategie weiter optimiert werden, entweder durch den Einsatz von

(i) höherwertige Abstandssensoren,

(ii) alternative sekundäre Sensorsignale,

(iii) oder einen anderen Segmentierungsansatz, der ausschließlich auf den Körperschallsignalen basiert

um den oberen Totpunkt des Werkzeugs zu identifizieren und die Phasen des Schneid- und Abstreifprozesses aus den Rohsignalen zu extrahieren. Diese drei Ansätze schließen sich nicht gegenseitig aus. Eine praktikable Strategie könnte darin bestehen, ein überwachtes maschinelles Lernmodell auf einer ausreichend heterogenen Menge extrahierter Segmente zu trainieren, die von einer sekundären Sensorquelle stammen, um Segmente in den Körperschallsignalen automatisiert zu identifizieren. Körperschallsignale unterscheiden sich zwar erheblich zwischen den verschiedenen Prozesskonfigurationen (Pressentyp, Blechdicke, Geometrie, Folgeverbundwerkzeug usw.), aber ein ausreichend heterogener Datensatz könnte diese Unterschiede möglicherweise überwinden. Das Problem der Segmentierung von Zeitreihen wurde bereits in mehreren wissenschaftlichen Domänen bearbeitet, z. B. in der Fertigung [GAUG23], bei der Aktivitätserkennung [XIA20] und bei der Schlafstadienbestimmung [PERS19]. Diese Architekturen sollten genutzt und an ausreichend heterogenen Datensätzen getestet werden, um zuverlässige Segmentierungsmodelle zu entwickeln.

Die Entwicklung der Proxy-Variable für den Stempelverschleiß basiert auf physikalischen Prinzipien. Die Beobachtung der Proxy-Variable war auf die am stärksten verschlissenen Flächen der untersuchten Stempel beschränkt. Zukünftige Arbeiten sollten sich auf die Untersuchung des kausalen Zusammenhangs zwischen Stempelverschleißmechanismen und der Entwicklung der Proxy-Variable konzentrieren, zum Beispiel mit Methoden der kausalen Inferenz. Zu diesem Zweck müssen geplante Maschinenstillstände und eine direkte Zustandsüberwachung des Stempels, die eine häufige Demontage des Werkzeugs erfordert, durchgeführt werden. Weiterhin sollte ein größerer Teil der Schnittflächen der Stanzgitter-Segmente untersucht werden, um eine umfangreichere und differenziertere Darstellung des Stempelverschleißzustandes zu generieren.

Es ist bekannt, dass die mechanisch-technologischen Parameter von Blechwerkstoffen schwanken. Da die Prozessparameter beim Stanzen in der Regel einmalig eingestellt werden und bei stabilem Prozessverlauf unverändert bleiben, beeinflusst diese Schwankung der mechanisch-technologischen Parameter bei konstanten Prozessparametern die Verschleißentwicklung kritischer Werkzeugkomponenten. Eine Kombination von Ansätzen, die aus der zerstörungsfreien Prüfung von Blechwerkstoffen auf mechanisch-technologische Werkstoffparameter schließen, mit der entwickelten Methodik der indirekten Zustandsüberwachung ist vielversprechend. Ein möglicher zukünftiger Datenfluss zwischen Materiallieferant (zerstörungsfreie Prüfung) und Materialverarbeiter (Körperschallsignalen) stellt eine ökonomisch und ökologisch sinnvolle und aussichtsreiche Kooperation dar.

Zukünftige Forschungsarbeiten sollten die Integration des entwickelten Ansatzes in breitere Industrie 4.0-Rahmenwerke untersuchen, indem Erkenntnisse über den Werkzeugverschleiß mit anderen Produktionsparametern für eine ganzheitlichere Optimierungsstrategie gekoppelt werden und Stillstandszeiten mit intelligenter Ressourcenzuweisung entsprechend geplant werden. Die entwickelte Methodik ist nicht auf Stanzprozesse beschränkt und könnte auf andere getaktete Blechumformprozesse ausgeweitet werden, um die Universalität des Ansatzes zu erhöhen.

# Bibliography

*Literaturverzeichnis*


[ALLA20] ALLAOUI, M., M. L. KHERFI, and A. CHERIET. "Considerably improving clustering algorithms using UMAP dimensionality reduction technique: a comparative study". In: *International Conference on Image and Signal Processing*. Springer. 2020, pp. 317–325. DOI: https://doi.org/10.1007/978-3-030-51935-3_34.

[ARIK21] ARIK, S. Ö. and T. PFISTER. "Tabnet: Attentive interpretable tabular learning". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. 8. 2021, pp. 6679–6687. DOI: https://doi.org/10.1609/aaai.v35i8.16826.

[ASAH21] ASAHI, S., C. KARADOGAN, S. TAMURA, S. HAYAMIZU, and M. LIEWALD. "Process data based estimation of tool wear on punching machines using TCN-Autoencoder from raw time-series information". In: *IOP Conference Series: Materials Science and Engineering* 1157(1) (June 2021), p. 012078. DOI: https://doi.org/10.1088/1757-899X/1157/1/012078.

[BACC15] BACCAR, D. and D. SÖFFKER. "Wear detection by means of wavelet-based acoustic emission analysis". In: *Mechanical Systems and Signal Processing* 60-61 (2015), pp. 198–207. DOI: https://doi.org/10.1016/j.ymssp.2015.02.012.

[BARA20] BARANDAS, M., D. FOLGADO, L. FERNANDES, S. SANTOS, M. ABREU, P. BOTA, H. LIU, T. SCHULTZ, and H. GAMBOA. "TSFEL: Time Series Feature Extraction Library". In: *SoftwareX* 11 (2020), p. 100456. DOI: https://doi.org/10.1016/j.softx.2020.100456.

[BARR20] BARREDO ARRIETA, A., N. DÍAZ-RODRÍGUEZ, J. DEL SER, A. BENNETOT, S. TABIK, A. BARBADO, S. GARCIA, S. GIL-LOPEZ, D. MOLINA, R. BENJAMINS, R. CHATILA, and F. HERRERA. "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI". In: *Information Fusion* 58 (2020), pp. 82–115. DOI: https://doi.org/10.1016/j.inffus.2019.12.012.

[BASS07] BASSIUNY, A., X. LI, and R. DU. "Fault diagnosis of stamping process based on empirical mode decomposition and learning vector quantization". In: *International Journal of Machine Tools and Manufacture* 47(15) (2007), pp. 2298–2306. URL: https://doi.org/10.1016/j.ijmachtools.2007.06.006.

[BECH19] BECHT, E., L. MCINNES, J. HEALY, C.-A. DUTERTRE, I. W. KWOK, L. G. NG, F. GINHOUX, and E. W. NEWELL. "Dimensionality reduction for visualizing single-cell data using UMAP". In: *Nature Biotechnology* 37(1) (2019), pp. 38–44. DOI: https://doi.org/10.1038/nbt.4314.

[BEYE99] BEYER, K., J. GOLDSTEIN, R. RAMAKRISHNAN, and U. SHAFT. "When is "Nearest neighbor" meaningful?" In: *Proceedings of the 7th International Conference on Database Theory (ICDT '99)*. Springer, 1999, pp. 217–235. DOI: https://doi.org/10.1007/3-540-49257-7_15.

[BISH06] BISHOP, C. M. and N. M. NASRABADI. *Pattern recognition and machine learning*. Springer, 2006. ISBN: 978-0387310732.

[BLOO00] BLOOMFIELD, P. *Fourier analysis of time series: An introduction*. John Wiley & Sons, 2000. ISBN: 978-0471889489.

[BOX15] BOX, G. E., G. M. JENKINS, G. C. REINSEL, and G. M. LJUNG. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015. ISBN: 978-1118675021.

[BREI01] BREIMAN, L. "Random forests". In: *Machine Learning* 45(1) (2001), pp. 5–32. DOI: https://doi.org/10.1023/A:1010933404324.

[BRIL01] BRILLINGER, D. R. *Time series: Data analysis and theory*. SIAM, 2001. ISBN: 978-0030769757.

[BUTT30] BUTTERWORTH, S. et al. "On the theory of filter amplifiers". In: *Wireless Engineer* 7(6) (1930), pp. 536–541.

[CARV19] CARVALHO, T. P., F. A. SOARES, R. VITA, R. D. P. FRANCISCO, J. P. BASTO, and S. G. ALCALÁ. "A systematic literature review of machine learning methods applied to predictive maintenance". In: *Computers & Industrial Engineering* 137 (2019), p. 106024. DOI: https://doi.org/10.1016/j.cie.2019.106024.

[CHEN16] CHEN, T. and C. GUESTRIN. "XGBoost: A scalable tree boosting system". In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2016, pp. 785–794. DOI: https://doi.org/10.1145/2939672.2939785.

[COOL65] COOLEY, J. W. and J. W. TUKEY. "An algorithm for the machine calculation of complex Fourier series". In: *Mathematics of Computation* 19(90) (1965), pp. 297–301. URL: https://community.ams.org/journals/mcom/1965-19-090/S0025-5718-1965-0178586-1/S0025-5718-1965-0178586-1.pdf.

[CZIC10] CZICHOS, H. and K.-H. HABIG. *Tribologie-Handbuch: Tribometrie, Tribomaterialien, Tribotechnik*. Springer, 2010. ISBN: 978-3834800176.

[DAMR21] DAMRICH, S. and F. A. HAMPRECHT. "On UMAP's true loss function". In: *Advances in Neural Information Processing Systems* 34 (2021), pp. 5798–5809. URL: https://proceedings.neurips.cc/paper_files/paper/2021/file/2de5d16682c3c35007e4e92982f1a2ba-Paper.pdf.

[DECH00] DECHIFFRE, L., P. LONARDO, H. TRUMPOLD, D. LUCCA, G. GOCH, C. BROWN, J. RAJA, and H. N. HANSEN. "Quantitative characterisation of surface texture". In: *CIRP Annals* 49(2) (2000), pp. 635–652. DOI: https://doi.org/10.1016/S0007-8506(07)63458-1.

[DEUT13] DEUTSCHES INSTITUT FÜR NORMUNG. *DIN 8588:2013-08 : Manufacturing processes severing - Classification, subdivision, terms and definitions (German version)*. Berlin, 2013. DOI: https://dx.doi.org/10.31030/2047900.

[DEUT17] DEUTSCHES INSTITUT FÜR NORMUNG. *DIN EN 1330-9:2017-10: Non-destructive testing - Terminology - Part 9: Terms used in acoustic emission testing (German version)*. Berlin, 2017. DOI: https://doi.org/10.31030/2607064.

[DJAV10] DJAVANROODI, F., A. PIRGHOLI, and E. DERAKHSHANI. "FEM and ANN analysis in fine-blanking process". In: *Materials and Manufacturing Processes* 25(8) (2010), pp. 864–872. DOI: https://doi.org/10.1080/10426910903367444.

[DOŠ18] DOŠILOVIĆ, F. K., M. BRČIĆ, and N. HLUPIĆ. "Explainable artificial intelligence: A survey". In: *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. IEEE. 2018, pp. 0210–0215. DOI: https://doi.org/10.23919/MIPRO.2018.8400040.

[DRES04] DRESSLER, O., G. SCHNEIDER, G. STOCKMANNS, and E. KOCHS. "Awareness and the EEG power spectrum: analysis of frequencies". In: *British Journal of Anaesthesia* 93(6) (2004), pp. 806–809. DOI: https://doi.org/10.1093/bja/aeh270.

[EURO21A] EUROPEAN COMMISSION. *European Commission: Climate & energy framework*. [Accessed 25-Aug-2023]. 2021. URL: https://climate.ec.europa.eu/eu-action/climate-strategies-targets/2030-climate-energy-framework_en.

[EURO21B] EUROPEAN COMMISSION. *European Commission: Long-term strategy*. [Accessed 25-Aug-2023]. 2021. URL: https://climate.ec.europa.eu/eu-action/climate-strategies-targets/2050-long-term-strategy_en.

[EVER10] EVERITT, B. S. and A. SKRONDAL. *The Cambridge dictionary of statistics*. Vol. 2. Cambridge, 2010. ISBN: 978-0521766999.

[FREU99] FREUND, Y., R. SCHAPIRE, and N. ABE. "A short introduction to boosting". In: *Journal of Japanese Society For Artificial Intelligence* 14(771-780) (1999), p. 1612. URL: https://cseweb.ucsd.edu/~yfreund/papers/IntroToBoosting.pdf.

[FRIE01] FRIEDMAN, J. H. "Greedy function approximation: A gradient boosting machine". In: *The Annals of Statistics* (2001), pp. 1189–1232. DOI: https://doi.org/10.1214/aos/1013203451.

[GAAR09] GAARD, A., P. KRAKHMALEV, and J. BERGSTRÖM. "Wear mechanisms in galling: cold work tool materials sliding against high-strength carbon steel sheets". In: *Tribology Letters* 33 (2009), pp. 45–53. DOI: https://doi.org/10.1007/s11249-008-9390-6.

[GADE02] GADELMAWLA, E., M. M. KOURA, T. M. MAKSOUD, I. M. ELEWA, and H. SOLIMAN. "Roughness parameters". In: *Journal of Materials Processing Technology* 123(1) (2002), pp. 133–145. DOI: https://doi.org/10.1016/S0924-0136(02)00060-2.

[GAUG23] GAUGEL, S. and M. REICHERT. "PrecTime: A deep learning architecture for precise time series segmentation in industrial manufacturing operations". In: *Engineering Applications of Artificial Intelligence* 122 (2023), p. 106078. DOI: https://doi.org/10.1016/j.engappai.2023.106078.

[GESE02] GESELLSCHAFT FÜR TRIBOLOGIE. *Arbeisblatt 7: Tribologie–Verschleiß, Reibung, Definitionen, Begriffe, Prüfung*. [Accessed 28-Aug-2023]. 2002. URL: https://neu.gft-ev.de/wp-content/uploads/2002_AB_7_Tribologie.pdf.

[GHOJ23] GHOJOGH, B., M. CROWLEY, F. KARRAY, and A. GHODSI. *Elements of Dimensionality Reduction and Manifold Learning*. Springer, 2023. ISBN: 978-3031106019.

[GOOD16] GOODFELLOW, I., Y. BENGIO, and A. COURVILLE. *Deep learning*. MIT press, 2016. ISBN: 978-0262035613.

[GROO10] GROOVER, M. *Fundamentals of modern manufacturing: materials, processes, and systems*. 4th ed. John Wiley & Sons, 2010. ISBN: 978-1118393673.

[HALK11] HALKO, N., P.-G. MARTINSSON, and J. A. TROPP. "Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions". In: *SIAM review* 53(2) (2011), pp. 217–288. DOI: https://doi.org/10.1137/090771806.

[HALL99] HALL, M. A. "Correlation-based feature selection for machine learning". PhD thesis. The University of Waikato, 1999. URL: https://researchcommons.waikato.ac.nz/bitstream/handle/10289/15043/thesis.pdf.

[HARR20] HARRIS, C. R., K. J. MILLMAN, S. J. van der WALT, R. GOMMERS, P. VIRTANEN, D. COURNAPEAU, E. WIESER, J. TAYLOR, S. BERG, N. J. SMITH, R. KERN, M. PICUS, S. HOYER, M. H. van KERKWIJK, M. BRETT, A. HALDANE, J. F. del RÍO, M. WIEBE, P. PETERSON, P. GÉRARD-MARCHANT, K. SHEPPARD, T. REDDY, W. WECKESSER, H. ABBASI, C. GOHLKE, and T. E. OLIPHANT. "Array programming with NumPy". In: *Nature* 585(7825) (Sept. 2020), pp. 357–362. DOI: https://doi.org/10.1038/s41586-020-2649-2.

[HART75] HARTIGAN, J. A. *Clustering algorithms*. John Wiley & Sons, Inc., 1975. ISBN: 978-0471356455.

[HAST09] HASTIE, T., R. TIBSHIRANI, J. H. FRIEDMAN, and J. H. FRIEDMAN. *The elements of statistical learning: data mining, inference, and prediction*. Vol. 2. Springer, 2009. ISBN: 978-0387848570.

[HEIN02] HEINZEL, G., A. RÜDIGER, and R. SCHILLING. “Spectrum and spectral density estimation by the discrete Fourier transform (DFT), including a comprehensive list of window functions and some new at-top windows”. In: (2002). URL: https://hdl.handle.net/11858/00-001M-0000-0013-557A-5.

[HOTE33] HOTELLING, H. “Analysis of a complex of statistical variables into principal components”. In: *Journal of Educational Psychology* 24(6) (1933), p. 417. DOI: http://dx.doi.org/10.1037/h0071325.

[HU13] HU, B., Y. CHEN, and E. KEOGH. “Time series classification under more realistic assumptions”. In: *Proceedings of the 2013 SIAM International Conference on Data Mining*. SIAM. 2013, pp. 578–586. DOI: https://doi.org/10.1137/1.9781611972832.64.

[JAME13] JAMES, G., D. WITTEN, T. HASTIE, and R. TIBSHIRANI. *An introduction to statistical learning*. 112th ed. Springer, 2013. ISBN: 978-1071614174.

[JIA22] JIA, W., M. SUN, J. LIAN, and S. HOU. “Feature dimensionality reduction: a review”. In: *Complex & Intelligent Systems* 8(3) (2022), pp. 2663–2693. DOI: https://doi.org/10.1007/s40747-021-00637-x.

[JOHN67] JOHNSON, S. C. “Hierarchical clustering schemes”. In: *Psychometrika* 32(3) (1967), pp. 241–254. URL: http://cda.psych.uiuc.edu/psychometrika_highly_cited_articles/johnson_1967.pdf.

[JOOS12] JOOST, W. J. “Reducing vehicle weight and improving US energy efficiency using integrated computational materials engineering”. In: *The Journal of The Minerals, Metals Materials Society* 64 (2012), pp. 1032–1038. DOI: https://doi.org/10.1007/s11837-012-0424-z.

[JORD15] JORDAN, M. I. and T. M. MITCHELL. “Machine learning: Trends, perspectives, and prospects”. In: *Science* 349(6245) (2015), pp. 255–260. DOI: https://doi.org/10.1126/science.aaa841.

[KE17] KE, G., Q. MENG, T. FINLEY, T. WANG, W. CHEN, W. MA, Q. YE, and T.-Y. LIU. “LightGBM: A highly efficient gradient boosting decision tree”. In: *Advances in Neural Information Processing Systems* 30 (2017). URL: https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf.

[KEYE23] KEYENCE. *Arithmetical mean height (Ra, Pa, Wa) | Surface roughness parameters | Introduction to roughness*. [Accessed 14-Jun-2023]. 2023. URL: https://www.keyence.com/ss/products/microscope/roughness/line/parameters.jsp.

[KLOC01] KLOCKE, F. and H.-W. RAEDT. “Formulation and testing of optimised coating properties with regard to tribological performance in cold forging and fine blanking applications”. In: *International Journal of Refractory Metals and Hard Materials* 19(4-6) (2001), pp. 495–505. DOI: https://doi.org/10.1016/S0263-4368(01)00029-4.

[KLOC14] KLOCKE, F. *Manufacturing processes 4: Forming*. Springer, 2014. ISBN: 978-3642367717.

[KLOC16] KLOCKE, F., A. SHIROBOKOV, D. TRAUTH, and P. MATTFELD. “Deep rolling of fine blanking punch edges: Numerical and experimental investigation of a novel deep rolling tool for filleting of cylindrical punches”. In: *International Journal of Material Forming* 9 (2016), pp. 489–498. DOI: https://doi.org/10.1007/s12289-015-1235-x.

[KOHA97] KOHAVI, R. and G. H. JOHN. “Wrappers for feature subset selection”. In: *Artificial Intelligence* 97(1-2) (1997), pp. 273–324. DOI: https://doi.org/10.1016/S0004-3702(97)00043-X.

[KOLL18] KOLLMENT, W., P. O'LEARY, M. HARKER, T. KLÜNSNER, and S. ECK. "Force and acoustic emission measurements for condition monitoring of fine blanking tools". In: 2018 IEEE International Instrumentation and Measurement Technology Conference. IEEE. 2018, pp. 1–6. DOI: https://doi.org/10.1109/I2MTC.2018.8409569.

[KONG20] KONG, Q., T. SIAUW, and A. BAYEN. *Python programming and numerical methods: A guide for engineers and scientists*. Academic Press, 2020. ISBN: 978-0128195499.

[KUBI22] KUBIK, C., S. KNAUER, and P. GROCHE. "Smart sheet metal forming: importance of data acquisition, preprocessing and transformation on the performance of a multiclass support vector machine for predicting wear states during blanking". In: *Journal of Intelligent Manufacturing* 33(1) (2022), pp. 259–282. DOI: https://doi.org/10.1007/s10845-021-01789-w.

[KUHN13] KUHN, M. and K. JOHNSON. *Applied predictive modeling*. Springer, 2013. ISBN: 978-1461468486.

[LAW19] LAW, S. M. "Stumpy: A powerful and scalable python library for rime series data mining". In: *The Journal of Open Source Software* 4(39) (2019), p. 1504. DOI: https://doi.org/10.21105/joss.01504.

[LECU95] LECUN, Y., Y. BENGIO, et al. "Convolutional networks for images, speech, and time series". In: *The Handbook of Brain Theory and Neural Networks* 3361(10) (1995), p. 1995.

[LEE95] LEE, T., L. C. CHAN, and B. WU. "Straining behaviour in blanking process-fine blanking vs conventional blanking". In: *Journal of Materials Processing Technology* 48(1-4) (1995), pp. 105–111. DOI: https://doi.org/10.1016/0924-0136(94)01639-I.

[LIEW22] LIEWALD, M., T. BERGS, P. GROCHE, B.-A. BEHRENS, D. BRIESENICK, M. MÜLLER, P. NIEMIETZ, C. KUBIK, and F. MÜLLER. "Perspectives on data-driven models and its potentials in metal forming and blanking technologies". In: *Production Engineering* 16(5) (2022), pp. 607–625. DOI: https://doi.org/10.1007/s11740-022-01115-0.

[LOUP13] LOUPPE, G., L. WEHENKEL, A. SUTERA, and P. GEURTS. "Understanding variable importances in forests of randomized trees". In: *Advances in Neural Information Processing Systems* (2013), pp. 431–439. URL: https://proceedings.neurips.cc/paper/2013/file/53adaf494dc89ef7196d73636eb2451b-Paper.pdf.

[LUND17] LUNDBERG, S. M. and S.-I. LEE. "A unified approach to interpreting model predictions". In: *Advances in Neural Information Processing Systems* 30 (2017). URL: https://papers.nips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf.

[LUND20] LUNDBERG, S., G. ERION, H. CHEN, A. DEGRAVE, J. PRUTKIN, B. NAIR, R. KATZ, J. HIMMELFARB, N. BANSAL, and S.-I. LEE. "From local explanations to global understanding with explainable AI for trees". In: *Nature Machine Intelligence* 2(1) (2020), pp. 56–67. DOI: https://doi.org/10.1038/s42256-019-0138-9.

[MARW12] MARWALA, T. *Condition monitoring using computational intelligence methods: applications in mechanical and electrical systems*. 2012. ISBN: 978-1447123798.

[MATH23] MATHWORKS. *Mean frequency - MATLAB meanfreq*. [Accessed 31-May-2023]. 2023. URL: https://de.mathworks.com/help/signal/ref/meanfreq.html.

[MCIN17] MCINNES, L., J. HEALY, and S. ASTELS. "hdbscan: Hierarchical density based clustering". In: *The Journal of Open Source Software* 2(11) (2017), p. 205. DOI: https://doi.org/10.21105/joss.00205.

[McIn18] McInnes, L., J. Healy, and J. Melville. “Umap: Uniform manifold approximation and projection for dimension reduction”. In: *arXiv preprint arXiv:1802.03426* (2018). DOI: https://doi.org/10.48550/arXiv.1802.03426.

[McKi12] McKinsey Company. *Lightweight, heavy impact*. [Accessed 25-Aug-2023]. 2012. URL: https://www.mckinsey.com/~/media/mckinsey/dotcom/client_service/automotive%20and%20assembly/pdfs/lightweight_heavy_impact.ashx.

[McKi21] McKinsey Company. *Prediction at scale: How industry can get more value out of maintenance*. [Accessed 28-Aug-2023]. 2021. URL: https://www.mckinsey.com/capabilities/operations/our-insights/prediction-at-scale-how-industry-can-get-more-value-out-of-maintenance#/.

[Mitc97] Mitchell, T. M. *Machine Learning*. McGraw-Hill, 1997. ISBN: 978-0070428072.

[Mobl02] Mobley, R. K. *An introduction to predictive maintenance*. Elsevier, 2002. ISBN: 978-0123996374.

[Moln20] Molnar, C. *Interpretable machine learning*. Lulu.com, 2020. ISBN: 978-0244768522.

[Muee09] Mueen, A., E. Keogh, Q. Zhu, S. Cash, and B. Westover. “Exact discovery of time series motifs”. In: *Proceedings of the 2009 SIAM international conference on data mining*. SIAM. 2009, pp. 473–484. DOI: https://doi.org/10.1137/1.9781611972795.41.

[Muee22] Mueen, A., S. Zhing, Y. Zhu, M. Yeh, K. Kamgar, K. Viswanathan, C. Gupta, and E. Keogh. *The fastest similarity search algorithm for time series subsequences under Euclidean distance*. [Accessed 11-May-2023]. 2022. URL: http://www.cs.unm.edu/~mueen/FastestSimilaritySearch.html.

[Niem21] Niemietz, P., M. Unterberg, D. Trauth, and T. Bergs. “Autoencoder based wear assessment in sheet metal forming”. In: *IOP Conference Series: Materials Science and Engineering* 1157(1) (June 2021), p. 012082. DOI: https://doi.org/10.1088/1757-899X/1157/1/012082.

[Niem22] Niemietz, P., M. Kornely, D. Trauth, and T. Bergs. “Relating wear stages in sheet metal forming based on short-and long-term force signal variations”. In: *Journal of Intelligent Manufacturing* 33(7) (2022), pp. 2143–2155. DOI: https://doi.org/10.1007/s10845-022-01979-0.

[Niem23] Niemietz, P., M. Fencl, and T. Bergs. “Study on learning efficient stroke representations in clocked sheet metal processing: theoretical and practical evaluation”. In: *Production Engineering* 17(2) (2023), pp. 279–289. DOI: https://doi.org/10.1007/s11740-023-01182-x.

[Pedr11] Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830. URL: https://jmlr.csail.mit.edu/papers/volume12/pedregosa11a/pedregosa11a.pdf.

[Pere12] Pereira, M. P., W. Yan, and B. F. Rolfe. “Wear at the die radius in sheet metal stamping”. In: *Wear* 274 (2012), pp. 355–367. DOI: https://doi.org/10.1016/j.wear.2011.10.006.

[Pers19] Perslev, M., M. Jensen, S. Darkner, P. J. Jennum, and C. Igel. “U-time: A fully convolutional network for time series segmentation applied to sleep staging”. In: *Advances in Neural Information Processing Systems* 32 (2019). URL: https://proceedings.neurips.cc/paper_files/paper/2019/file/57bafb2c2dfeefba931bb03a835b1fa9-Paper.pdf.

[PHIN12] PHINYOMARK, A., S. THONGPANJA, H. HU, P. PHUKPATTARANONT, and C. LIMSAKUL. "The usefulness of mean and median frequencies in electromyography analysis". In: *Computational Intelligence in Electromyography Analysis - A Perspective on Current Applications and Future Challenges* 81 (2012), p. 67. DOI: https://doi.org/10.5772/50639.

[PROA07] PROAKIS, J. G. *Digital signal processing: principles, algorithms, and applications*. Pearson Education India, 2007. ISBN: 978-0133942897.

[REHO05] REHORN, A. G., J. JIANG, and P. E. ORBAN. "State-of-the-art methods and results in tool condition monitoring: a review". In: *The International Journal of Advanced Manufacturing Technology* 26 (2005), pp. 693–710. DOI: https://doi.org/10.1007/s00170-004-2038-2.

[REYN09] REYNOLDS, D. A. et al. "Gaussian mixture models". In: *Encyclopedia of Biometrics* 741(659-663) (2009). DOI: https://doi.org/10.1007/978-1-4899-7488-4_196.

[RUDI19] RUDIN, C. "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead". In: *Nature Machine Intelligence* 1(5) (2019), pp. 206–215. DOI: https://doi.org/10.1038/s42256-019-0048-x.

[SÁN20] SÁNCHEZ, C., A. HAGBERG, J. M. HURTADO, and A. PÉREZ-SUAY. *dcor: Distance correlation with confidence intervals*. Version 0.5.3. 2020. DOI: https://doi.org/10.5281/zenodo.3909791.

[SEEM16] SEEMUANG, N. "Non-destructive evaluation and condition monitoring of tool wear". PhD thesis. University of Sheffield, 2016. URL: https://core.ac.uk/download/pdf/42606234.pdf.

[SHAN17] SHANBHAG, V. V., P. M. PEREIRA, F. B. ROLFE, and N. ARUNACHALAM. "Time series analysis of tool wear in sheet metal stamping using acoustic emission". In: *Journal of Physics: Conference series*. Vol. 896. 1. IOP Publishing. 2017, p. 012030. DOI: https://doi.org/10.1088/1742-6596/896/1/012030.

[SHAN18] SHANBHAG, V. V., B. F. ROLFE, N. ARUNACHALAM, and M. P. PEREIRA. "Investigating galling wear behaviour in sheet metal stamping using acoustic emissions". In: *Wear* 414 (2018), pp. 31–42. DOI: https://doi.org/10.1016/j.wear.2018.07.003.

[SHAN19] SHANBHAG, V. V., B. F. ROLFE, J. M. GRIFFIN, N. ARUNACHALAM, and M. P. PEREIRA. "Understanding galling wear initiation and progression using force and acoustic emissions sensors". In: *Wear* 436 (2019), p. 202991. DOI: https://doi.org/10.1016/j.wear.2019.202991.

[SHAN20] SHANBHAG, V., B. ROLFE, and M. PEREIRA. "Investigation of galling wear using acoustic emission frequency characteristics". In: *Lubricants* 8(3) (2020), p. 25. DOI: https://doi.org/10.3390/lubricants8030025.

[SHAN49] SHANNON, C. E. "Communication in the presence of noise". In: *Proceedings of the IRE* 37(1) (1949), pp. 10–21. DOI: https://doi.org/10.1109/JRPROC.1949.232969.

[SHWA22] SHWARTZ-ZIV, R. and A. ARMON. "Tabular data: Deep learning is not all you need". In: *Information Fusion* 81 (2022), pp. 84–90. DOI: https://doi.org/10.1016/j.inffus.2021.11.011.

[SIEG15] SIEGERT, K. *Blechumformung: Verfahren, Werkzeuge und Maschinen*. Springer, 2015. ISBN: 978-3540024880.

[SOLO91] SOLOMON JR, O. M. "PSD computations using Welch's method". In: *NASA STI/Recon Technical Report N* 92 (1991), p. 23584. DOI: https://doi.org/10.2172/5688766.

[STAT23a] STATISTA. *Industriestrompreise inkl. Stromsteuer in Deutschland*. [Accessed 25-Aug-2023]. 2023. URL: https://de.statista.com/statistik/daten/studie/252029/umfrage/industriestrompreise-inkl-stromsteuer-in-deutschland/.

[STAT23b] STATISTA. *Stahl - Index zur Preisentwicklung in Deutschland bis 2022*. [Accessed 25-Aug-2023]. 2023. URL: https://de.statista.com/statistik/daten/studie/377846/umfrage/stahlpreisentwicklung-in-deutschland/.

[STOI05] STOICA, P., R. L. MOSES, et al. *Spectral analysis of signals*. Vol. 452. Pearson Prentice Hall Upper Saddle River, NJ, 2005. ISBN: 978-0131139565.

[STRA06] STRANG, G. *Linear algebra and its applications*. Belmont, CA: Thomson, Brooks/Cole, 2006. ISBN: 978-8131501726.

[SZEK07] SZ'EKELY, G. J., M. L. RIZZO, and N. K. BAKIROV. "Measuring and testing dependence by correlation of distances". In: *The Annals of Statistics* 35(6) (2007), pp. 2769–2794. DOI: https://doi.org/10.1214/009053607000000505.

[TIBS96] TIBSHIRANI, R. "Regression shrinkage and selection via the lasso". In: *Journal of the Royal Statistical Society, Series B* 58(1) (1996), pp. 267–288. DOI: https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.

[TROE17] TROEBER, P., P. DEMMEL, H. HOFFMANN, R. GOLLE, and W. VOLK. "On the influence of Seebeck coefficients on adhesive tool wear during sheet metal processing". In: *CIRP Annals* 66(1) (2017), pp. 293–296. DOI: https://doi.org/10.1016/j.cirp.2017.04.116.

[UMAP23] UMAP-LEARN. *How UMAP Works – umap 0.5 documentation*. [Accessed 11-May-2023]. 2023. URL: https://umap-learn.readthedocs.io/en/latest/how_umap_works.html#id9.

[UNTE19] UNTERBERG, M., P. NIEMIETZ, D. TRAUTH, K. WEHRLE, and T. BERGS. "In-situ material classification in sheet-metal blanking using deep convolutional neural networks". In: *Production Engineering* 13 (2019), pp. 743–749. DOI: https://doi.org/10.1007/s11740-019-00928-w.

[UNTE21] UNTERBERG, M., H. VOIGTS, I. F. WEISER, A. FEUERHACK, D. TRAUTH, and T. BERGS. "Wear monitoring in fine blanking processes using feature based analysis of acoustic emission signals". In: *Procedia CIRP* 104 (2021), pp. 164–169. DOI: https://doi.org/10.1016/j.procir.2021.11.028.

[UNTE23] UNTERBERG, M., M. BECKER, P. NIEMIETZ, and T. BERGS. "Data-driven indirect punch wear monitoring in sheet-metal stamping processes". In: *Journal of Intelligent Manufacturing* 33(1) (2023), pp. 259–282. DOI: https://doi.org/10.1007/s10845-023-02129-w.

[VAN 08] VAN DER MAATEN, L. and G. HINTON. "Visualizing data using t-SNE." In: *Journal of Machine Learning Research* 9(11) (2008). URL: http://jmlr.org/papers/v9/vandermaaten08a.html.

[VIRT20] VIRTANEN, P., R. GOMMERS, T. E. OLIPHANT, M. HABERLAND, T. REDDY, D. COURNAPEAU, E. BUROVSKI, P. PETERSON, W. WECKESSER, J. BRIGHT, S. J. VAN DER WALT, M. BRETT, J. WILSON, K. J. MILLMAN, N. MAYOROV, A. R. J. NELSON, E. JONES, R. KERN, E. LARSON, C. J. CAREY, İ. POLAT, Y. FENG, E. W. MOORE, J. VANDERPLAS, D. LAXALDE, J. PERKTOLD, R. CIMRMAN, I. HENRIKSEN, E. A. QUINTERO, C. R. HARRIS, A. M. ARCHIBALD, A. H. RIBEIRO, F. PEDREGOSA, P. VAN MULBREGT, and SCIPY 1.0 CONTRIBUTORS. "SciPy 1.0: Fundamental algorithms for scientific computing in Python". In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[VOIG21] VOIGTS, H. O. "Feinschneiden mit Hartmetallstempeln: Fine blanking with cemented carbide punches". PhD thesis. 2021. ISBN: 978-3985550081.

[WEIS05] WEISBERG, S. *Applied linear regression*. Vol. 528. John Wiley & Sons, 2005. ISBN: 978-1118386088.

[WELC67] WELCH, P. "The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms". In: *IEEE Transactions on Audio and Electroacoustics* 15(2) (1967), pp. 70–73. DOI: https://doi.org/10.1109/TAU.1967.1161901.

[XGBO23] XGBOOST. *XGBoost Parameters: xgboost 1.7.2 documentation*. [Accessed 04-Jan-2023]. 2023. URL: https://xgboost.readthedocs.io/en/stable/parameter.html.

[XIA20] XIA, K., J. HUANG, and H. WANG. "LSTM-CNN architecture for human activity recognition". In: *IEEE Access* 8 (2020), pp. 56855–56866. DOI: https://doi.org/10.1109/ACCESS.2020.2982225.

[YEH18] YEH, C.-C. M., Y. ZHU, L. ULANOVA, N. BEGUM, Y. DING, H. A. DAU, Z. ZIMMERMAN, D. F. SILVA, A. MUEEN, and E. KEOGH. "Time series joins, motifs, discords and shapelets: a unifying view that exploits the matrix profile". In: *Data Mining and Knowledge Discovery* 32 (2018), pp. 83–123. DOI: https://doi.org/10.1007/s10618-017-0519-9.

[ZHEN19] ZHENG, Q., X. ZHUANG, and Z. ZHAO. "State-of-the-art and future challenge in fine-blanking technology". In: *Production Engineering* 13(1) (2019), pp. 61–70. DOI: https://doi.org/10.1007/s11740-018-0839-7.

# A Appendix

*Anhang*

## A.1 Appendix for Chapter 4

**Table A.1:** Used blank holder and counter holder forces for the laboratory experiments as well as $R_{p0.2}$ and $R_m$ of the processed material (manufacturer's inspection certificate).
*Verwendete Nieder- und Gegenhalterkräfte der Laborexperimente sowie $R_{p0.2}$ und $R_m$ des prozessierten Werkstoffs (Prüfzeugnis des Herstellers).*

| Dataset | Blank holder force [kN] | Counter holder force [kN] | $R_{\mathrm{p0.2}}$ [N/mm$^2$] | $R_{\mathrm{m}}$ [N/mm$^2$] |
|---|---|---|---|---|
| L1 | 450 | 300 | 323 | 594 |
| L2 | 450 | 300 | 390 | 590 |
| L3 | 500 | 300 | 409 | 549 |

## A.2 Appendix for Chapter 6

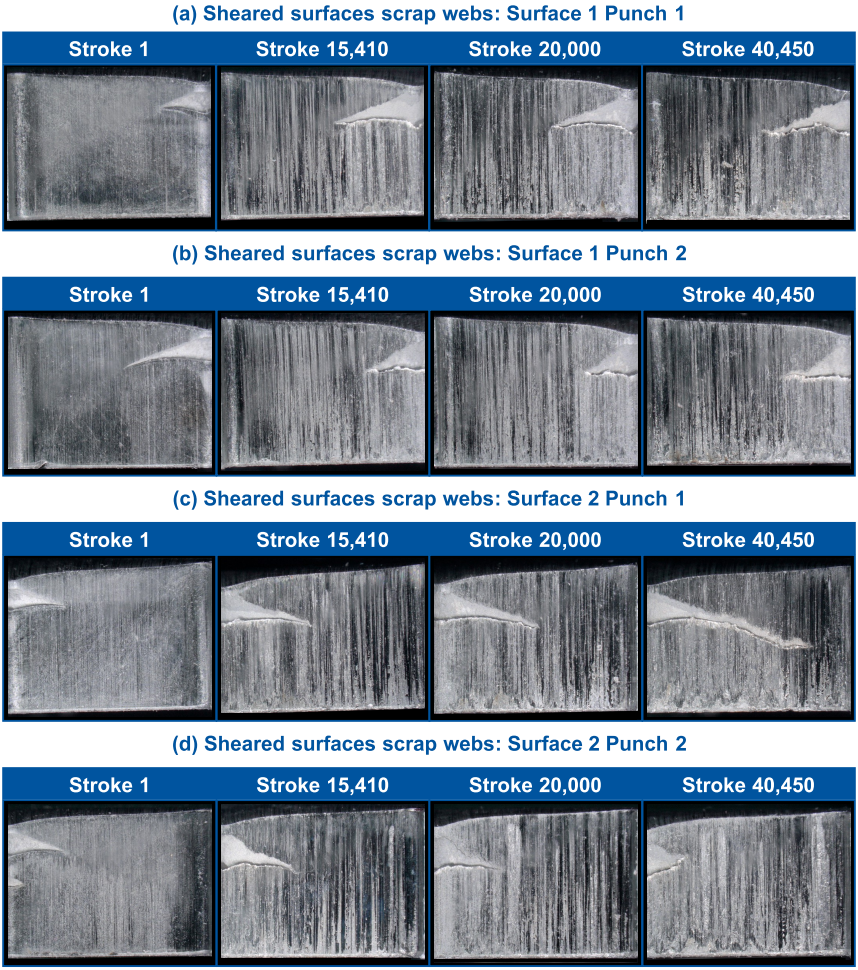



**Figure A.1:** Light microscopy images of the resulting scrap webs after 1, 15,410, 20,000, and 40,450 strokes for (a) Surface 1 Punch 1 (b) Surface 1 Punch 2 (c) Surface 2 Punch 1, and (d) Surface 2 Punch 2.
*Lichtmikroskopische Aufnahmen der entstandenen Stanzgitter nach 1, 15.410, 20.000 und 40.450 Hüben für (a) Fläche 1 Stempel 1 (b) Fläche 1 Stempel 2 (c) Fläche 2 Stempel 1 und (d) Fläche 2 Stempel 2.*

## A.3 Appendix for Chapter 8

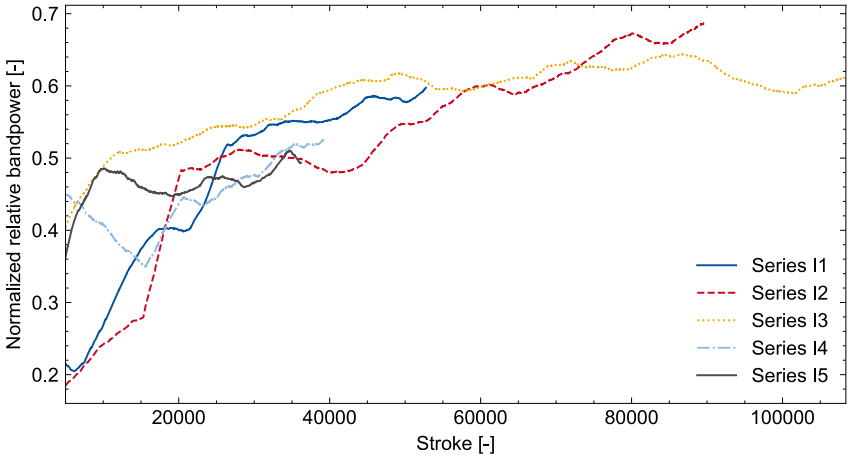



**Figure A.2:** Evolution of the series-wise normalized RBP feature (103–144 kHz) for the industrial stroke series (*rolling mean* window size 5,000).
*Entwicklung des Hubserien-weise normalisierten Features RBP (103–144 kHz) für die industrielle Hubserie (gleitender Mittelwert mit Fenstergröße 5.000).*

# B Abstract

*Kurzzusammenfassung*

The increasing scarcity of energy and the resulting rise in energy prices significantly impact operating costs in production facilities. This trend is particularly pronounced in sheet-metal stamping, where the use of higher-strength steels for lightweight construction leads to higher process friction, which increases the cost pressure on already tight margins due to unpredictably high tool wear. Due to the closed tool design, direct condition monitoring of sheet-metal stamping tools is not feasible.

This dissertation addresses these challenges by presenting a systematic approach for data-driven indirect condition monitoring of sheet-metal stamping tools using acoustic emission signals. Acoustic emission sensors are suitable signal sources for tool condition monitoring due to their cost efficiency and high information content of the signals. However, acoustic emission signals are difficult to process and interpret due to their complexity.

For this dissertation, 398,047 fineblanking process cycles from industrial and laboratory settings were acquired, preprocessed, and analyzed using a combination of machine learning and signal processing methods. Through domain-specific preprocessing, a process image representing the friction conditions between sheet metal and punch was isolated and exploratively analyzed. This data analysis revealed systematic patterns within and between the collected datasets that indicate a correlation with punch wear. To model the wear of the punch, a critical tool component, the quality of the sheared surface of the scrap web was observed and evaluated as a proxy variable for the punch wear in the laboratory setting. This proxy variable was used to generate labels for supervised machine learning models and to predict these labels from the acoustic emission data of the shearing and stripping phases of the fineblanking process. Evaluation of the models from a model performance and an explainable AI perspective, as well as analyses of the punches, revealed that adhesive wear of the shearing edges of the punches is a likely root cause of sheared surface quality degradation of the scrap web. These results were used to develop a scalar indicator for the tool condition, and the evolution of this indicator was evaluated for the datasets from the industrial setting. The analyses showed that the developed indicator allows the identification of problematic sequences of fineblanking process cycles based on the strength of its dispersion.

# Kurzzusammenfassung

Die zunehmende Energieknappheit und der daraus resultierende Anstieg der Energiepreise haben erhebliche Auswirkungen auf die Kosten für Stahlwerkstoffe und die Betriebskosten in Produktionsstätten. Besonders negativ wirkt sich dieser Trend im Bereich des Stanzens aus, wo der Einsatz höherfester Stähle für den Leichtbau durch erhöhte Prozessreibung den Kostendruck auf die ohnehin knappen Margen aufgrund des unvorhersehbar hohen Werkzeugverschleißes weiter erhöht. Aufgrund der geschlossenen Werkzeugkonstruktion beim Stanzen ist eine direkte Zustandsüberwachung der Werkzeuge nicht möglich.

Diese Dissertation adressiert die genannten Herausforderungen durch die Entwicklung eines systematischen Ansatzes zur datengetriebenen indirekten Zustandsüberwachung von Stanzwerkzeugen mittels Körperschallsignalen. Körperschallsensoren stellen aufgrund ihrer geringen Anschaffungskosten und des hohen Informationsgehalts der Signale geeignete Signalquellen für die Zustandsüberwachung von Werkzeugen dar. Allerdings sind Körperschallsignale aufgrund ihrer Komplexität schwierig zu verarbeiten und zu interpretieren.

Insgesamt wurden für diese Dissertation 398.047 Feinschneid-Prozesszyklen aus Industrie- und Laborumgebungen erfasst, vorverarbeitet und mit einer Kombination aus maschinellen Lernverfahren und Signalverarbeitungsmethoden analysiert. Durch eine domänenspezifische Vorverarbeitung wurde ein Prozessabbild als Repräsentation der Reibbedingungen zwischen Blech und Stempel isoliert und explorativ analysiert. Diese Datenanalyse ergab systematische Muster innerhalb und zwischen den erhobenen Datensätzen, die auf eine Korrelation mit dem Stempelverschleiß hindeuten. Um den Verschleiß des Stempels, der eine kritische Werkzeugkomponente darstellt, zu modellieren, wurde die Qualität der Schnittfläche des Stanzgitters in der Laborumgebung als Proxy-Variable für den Stempelverschleiß beobachtet und evaluiert. Diese Proxy-Variable wurde verwendet, um Labels für überwachte maschinelle Lernmodelle zu generieren und diese aus den Körperschallsignalen der Schneid- und Abstreifphasen des zugehörigen Feinschneidprozesses vorherzusagen. Evaluierung der Modelle aus der Perspektive der Modellgüte und der erklärbaren KI sowie Analysen der Stempel ergaben, dass adhäsiver Verschleiß der Schneidkanten an den Stempeln eine wahrscheinliche Hauptursache für die Beeinträchtigung der Schnittflächenqualität am Stanzgitter ist. Diese Ergebnisse wurden verwendet, um einen skalaren Indikator für den Werkzeugzustand zu entwickeln, dessen Evolution für die Datensätze aus dem industriellen Umfeld evaluiert wurde. Die Analyse ergab, dass der entwickelte Indikator eine Identifizierung problematischer Sequenzen von Feinschneidprozesszyklen anhand der Stärke seiner Streuung ermöglicht.