

# AcousticIntelligence - KI-basiertes Geräuschmonitoring zur objektiven Qualitätssicherung in Produktionsprozessen

Jan Troge<sup>1</sup>, Julie Damay<sup>2</sup>, Dennis M. Heim<sup>2</sup>, Michael Bortz<sup>2</sup>, Eric Hensel<sup>1</sup>, Peter Scholz<sup>1</sup>

<sup>1</sup> Fraunhofer-Institut für Werkzeugmaschinen und Umformtechnik IWU, 09126 Chemnitz, Deutschland

Email: jan.troge@iwu.fraunhofer.de

<sup>2</sup> Fraunhofer-Institut für Techno- und Wirtschaftsmathematik ITWM, 67663 Kaiserslautern, Deutschland

## Einleitung

Um die Effizienz von Produktionsanlagen und -maschinen nachhaltig zu steigern, sind zwei wichtige Aspekte zu berücksichtigen: die Überwachung der Produktionsprozesse und des Maschinenpark-Zustandes. Durch eine kontinuierliche und zuverlässige Prozessüberwachung werden Fehler in der Produktion frühzeitig erkannt und Korrekturen können eingeleitet werden. Die Kosten für Ausschuss und Nacharbeit stellen einen wesentlichen Stellhebel zur Verbesserung des Betriebsergebnisses von Unternehmen dar. Eine Reduktion von Ausschuss steigert die Produktivität und Kosteneffizienz und schont zudem Ressourcen. Durch eine möglichst frühzeitige Prognose bevorstehender Maschinenausfälle können Wartungsarbeiten geplant und eine hohe Verfügbarkeit sichergestellt werden.

Das menschliche Gehör eines erfahrenen Maschinenbedieners ist in der Lage, aus Prozessgeräuschen von Produktionsanlagen relevante akustische Informationen herauszufiltern und zu interpretieren. So werden Fehler und Abweichungen oft schnell erkannt und Gegenmaßnahmen können umgehend eingeleitet werden. Der vorliegende Beitrag zeigt Ansätze zur Objektivierung der Geräuschwahrnehmung eines erfahrenen Maschinenbedieners, um diese als Klassifizierungsinstrument in einem technischen Prozessmonitoringsystem zur Qualitätssicherung von Produktionsprozessen nutzen zu können. Um die sehr komplexe menschliche Wahrnehmung von Geräuschen nachzubilden, werden verschiedene Klassifizierungsmethoden auf Basis psychoakustischer Daten am Beispiel eines einfachen Tiefziehprozesses im Labormaßstab (Einzelhubversuche) untersucht und mit Lernverfahren des maschinellen Lernens (ML) kombiniert.

## Stand der Technik

Intelligente und selbstlernende Systeme für die Qualitäts- und Maschinenüberwachung in der Produktion leisten einen wesentlichen Beitrag zur Sicherung der langfristigen Wettbewerbsfähigkeit von Unternehmen. Mit der fortschreitenden Digitalisierung in der Produktion stehen mittlerweile umfangreiche Daten von Produktionsprozessen zur Verfügung, die zur Qualitätsüberwachung genutzt werden können. Prozessüberwachungssysteme werden bereits in zahlreichen Produktionsprozessen eingesetzt. Dass bei der Blechumformung eine Rissdetektion über Schallemissionsanalysen mittels Körperschall möglich ist, hat STRAUBE bereits vor knapp 20 Jahren gezeigt [1]. Das Potential von Acoustic Emission Sensoren zur Prozessüberwachung in der Blechumformung wurde

von HAO et al. [2] untersucht. In [3] konnte zudem nachgewiesen werden, dass die Fähigkeit für eine onlinefähige Prozessüberwachung gegeben ist. Die Auswertung von Körperschallsignalen mit ML-Methoden zur Erkennung des Umformzustandes untersuchten TSURUYA et al. [4]. Die akustische Analyse mittels Luftschall am Beispiel eines Scherschneidprozesses haben MÜHLBAUER et al. untersucht und mittels Methoden zur Anomalieerkennung gezeigt, dass durch einen modellbasierten Ansatz schleichende und plötzliche Prozessänderungen erkannt werden können [5].

Viele Klassifizierungsmethoden basieren auf der seit Jahrzehnten etablierten Fast Fourier Transformation (FFT) [6] und vergleichen einzelne Frequenzanteile oder -bänder mit definierten Referenzkurven. Im Bereich der Psychoakustik werden objektive und messbare Bewertungskriterien eines Hörereignisses entwickelt, die mit der menschlichen Hörwahrnehmung korrelieren, um Aussagen zur Schallwahrnehmung zu treffen. Eigene Untersuchungen zur subjektiven Qualitätsüberwachung bei einem Fräsprozess zeigen, dass psychoakustische Auswertungen eine Korrelation zum Bearbeitungsprozess und -qualität aufweisen [7]. Aufgrund der Komplexität der Schallwahrnehmung werden zunehmend verschiedene ML-Verfahren eingesetzt, wie z. B. Klassifikationsverfahren [8, 9], die Hauptkomponentenanalyse [10, 11] oder künstliche neuronale Netze als biologisch inspirierte Algorithmen [12, 13].

## Versuchsaufbau und -durchführung

Zur Realisierung eines Umformprozesses im Versuchsmaßstab wurde eine am Fraunhofer IWU vorhandene Versuchsmaschine eingesetzt (Abbildung 1). In diese Versuchsmaschine ist ein dreiteiliges Tiefziehwerkzeug integriert, das aus einem feststehenden Stempel und einem beweglichen Niederhalter und Matrize besteht. Zur Datenaufnahme wurden quadratische Blechplatinen (100 mm x 100 mm) eingelegt und durch einen halbkugelförmigen Stempel ( $R = 27,5$  mm) auf eine vorgegebene Tiefe umgeformt. Zur Variation der Prozessparameter wurden unterschiedliche Blechwerkstoffe (Edelstahlblech: 1.4301 in 0,6 mm und 0,8 mm; Aluminiumblech: AC300 in 1,0 mm und Tiefziehstahl: DX54 in 0,7 mm) eingesetzt und die Beölung variiert (vgl. Tabelle 1). Während des Versuchs werden für jeden Werkzeughub Messdaten über die verwendeten Sensoren aufgenommen.

Das gewählte Messsetup des Vorversuchs ist im linken Teil von Abbildung 1 zu sehen. Im dargestellten Versuchsaufbau ist zu erkennen, dass im Rahmen der Versu-

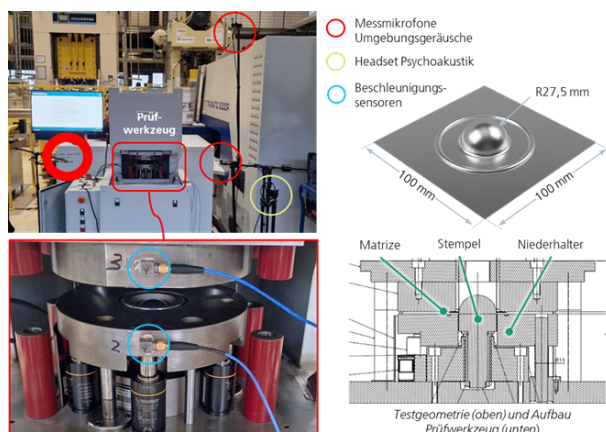


Abbildung 1: Versuchsaufbau Einzelhubversuch

che Messgrößen mit Beschleunigungsaufnehmern, Mikrofonen und einem binauralen Recording-Headset erfasst wurden. Da im vorliegenden Beitrag der Fokus auf der Verwendung psychoakustischer Parameter liegt, wird im Folgenden nicht näher auf die Körperschallsensoren eingegangen. Eine Korrelationsanalyse der extrahierten Merkmale, welche im nachfolgenden Abschnitt näher erläutert werden, ergab, dass die erfassten Luftschallsignale eine hohe Ähnlichkeit aufweisen. Daher wurden im Rahmen der vorgestellten Untersuchung nur die beiden Kanäle des binauralen Recording-Headsets (Typ: Brüel & Kjær 4965) näher betrachtet. Das Headset wurde während der Versuche vom Maschinenbediener getragen. Da beide Mikrofonkanäle des Headsets zu vergleichbaren Ergebnissen führten, werden im Nachfolgenden nur die Ergebnisse basierend auf den erfassten Messdaten des linken Mikrofons dargestellt. Die Versuchsmatrix der Hauptmessung ist in Tabelle 1 dargestellt. Wie zu erkennen ist, wurden insgesamt 341 Messungen für unterschiedliche Materialien und Blechdicken durchgeführt. Zudem wurden Messungen mit und ohne Beölung sowie das Schadensbild Riss messtechnisch erfasst. Die Messungen erfolgten mit einer Abtastfrequenz von 48 kHz und einem Zeitfenster von 3 s. Ein Triggersignal aus der Maschine stellte eine zeitsynchrone Messung sicher.

Tabelle 1: Versuchsmatrix der Hauptmessung (Angaben der Blechdicke in  $\mu\text{m}$ )

Riss	Beölung	Material	Dicke	Anzahl
Ja	Ja	1.4301	600	13
Nein	Nein	1.4301	600	6
Ja	Nein	1.4301	600	39
Nein	Ja	1.4301	800	41
Ja	Ja	1.4301	800	2
Ja	Nein	1.4301	800	40
Ja	Ja	AC300	1000	50
Ja	Nein	AC300	1000	50
Nein	Ja	DX54	700	50
Nein	Nein	DX54	700	50
			<b>Summe</b>	<b>341</b>

### Extraktion von Signalmerkmalen

Für alle erfassten Zeitdaten wurden unterschiedliche Signalparameter ermittelt. Die extrahierten Signalparameter

ter wurden basierend auf der zugrundeliegenden Berechnung in die nachfolgend beschriebenen Gruppen unterteilt. Zunächst erfolgte die Berechnung von Signalparametern im Zeitbereich (Bezeichnung  $\mathbf{t}$ , 11 Merkmale). Neben klassischen Parametern wie Mittelwert, Maximum, Minimum und Spannweite wurden zusätzliche Parameter wie Crest-, Form-, Impuls- und *Margin*-Faktor (vgl. [14]) berechnet. Die zweite Gruppe von Parametern beinhaltet statistische Größen (Bezeichnung  $\mathbf{p}$ , 5 Merkmale) wie Wölbung, Schiefe, Standardabweichung und dem *random error of average* (vgl. [15]). Die Berechnung der Signalparameter der dritten Gruppe basiert auf dem Amplitudenspektrum eines jeweiligen Zeitsignals (Bezeichnung  $\mathbf{f}$ , 7 Merkmale). Anhand des Amplitudenspektrums wurden alle oben genannten statistischen Kenngrößen sowie Mittelwert, Maximum und Summe der frequenzabhängigen Größe bestimmt. Die Parameter der Gruppen  $\mathbf{t}$ ,  $\mathbf{p}$  und  $\mathbf{f}$  werden im Folgenden als Standard-Signalparameter bezeichnet. Eine weitere Gruppe von Parametern stellen die psychoakustischen Kenngrößen dar. Die Psychoakustik-Parameter wurden mit der Python-Toolbox MOSQUITO [16] berechnet. Die psychoakustischen Kenngrößen werden nachfolgend mit  $\mathbf{SQ}$  gekennzeichnet. Es wurden insgesamt 50 Merkmale aus den berechneten Größen Lautheit, Schärfe, Rauigkeit, prominence ratio und tone-2-noise ratio extrahiert, wobei sowohl Einzahlwerte als auch zeitabhängige Größen berücksichtigt wurden. Die Reduktion der zeitabhängigen Größen auf Einzahlwerte erfolgte durch die Berechnung von Spannweite, Mittelwert, Median und Maximum. Weiterhin wurde das Terzspektrum für Terzmittenfrequenzen von 100 Hz bis 10 kHz berechnet und die Werte als zusätzliche Signalparameter hinterlegt (Bezeichnung **Terz**, 21 Merkmale). In Summe ergeben sich somit 94 Merkmale, welche nachfolgend zur Klassifikation der in Tabelle 1 dargestellten Zielgrößen verwendet wurden.

### Klassifikation

Mit den extrahierten Signalmerkmalen wurde ein Klassifikator trainiert, um die vier Zielgrößen aus Tabelle 1 vorherzusagen. Verschiedene Klassifikationsverfahren wurden eingesetzt und miteinander verglichen:  $k$ -nearest neighbors (kNN) [17], die Logistische Regression [17], die probabilistische Methode eSPA [18] und zwei neuronale Netzwerke mit unterschiedlicher Architektur. Im Folgenden werden die Ergebnisse des kNN Klassifikators vorgestellt und diskutiert. Dieses Klassifikationsverfahren liefert trotz seiner Einfachheit für alle Zielgrößen vergleichbar gute bzw. genauere Vorhersagen als die anderen Verfahren.

Mit kNN werden Vorhersagen durch Abstimmen der  $k$  nächsten Nachbarn im Raum der Merkmale gemacht. Die Anzahl  $k$  der Nachbarn ist ein Hyperparameter, der gesetzt werden muss. Für verschiedene  $k$ -Werte wird ein Klassifikator trainiert und die Vorhersagegenauigkeit von jedem Klassifikator auf einem Validierungsdatensatz berechnet. Der Hyperparameter wird dann auf den Wert gesetzt, der zur besten Vorhersagegenauigkeit für den Validierungsdatensatz führt. Die Vorhersagegenauigkeit ist in diesem Beitrag als der Anteil der richtig klassifizierten Punkte definiert. Da der Abstand der Punkte zueinander

von der Größenordnung der Merkmale stark beeinflusst wird, wurden die Merkmale zwischen 0 und 1 skaliert. Außerdem kann eine Ungleichverteilung der Trainingspunkte in den unterschiedlichen Klassen zu einer schlechteren Vorhersagegenauigkeit für die unterrepräsentierte(n) Klasse(n) führen. Deswegen wurde der Trainingsdatensatz für jede Zielgröße durch zufällige Auswahl (ohne Duplikate) der gleichen Anzahl an Punkten in jeder Klasse balanciert. Um die Vorhersagegenauigkeit des Klassifikators zu beurteilen, wurde der gesamte balancierte Datensatz in Trainings- und Testdatensatz aufgeteilt. Dabei wurde eine 10-fache Kreuzvalidierung (vgl. [17]) angewandt, um den Einfluss des Testdatensatzes zu verringern. Die Gleichverteilung der Punkte in den verschiedenen Klassen wurde sowohl im Trainings- als auch im Testdatensatz erhalten. Für jeden Trainingsdatensatz der Kreuzvalidierung wurde der Hyperparameter  $k$  neu bestimmt. Dabei wurden für  $k$  die Werte 2, 3, 5, 10 und 20 eingesetzt. Der beste  $k$ -Wert wurde anschließend benutzt, um Vorhersagen für die Punkte im Testdatensatz zu machen. Tabelle 2 zeigt die Aufteilung der Punkte zwischen Trainings- und Testdatensatz für jede Zielgröße. Die Anzahl der Punkte kann in der Kreuzvalidierung um einen Punkt abweichen, wenn die Gesamtzahl der Punkte nicht durch 10 teilbar ist. Für die Klassifikation wurde auf Methoden der Python-Bibliothek sklearn [19] zurückgegriffen.

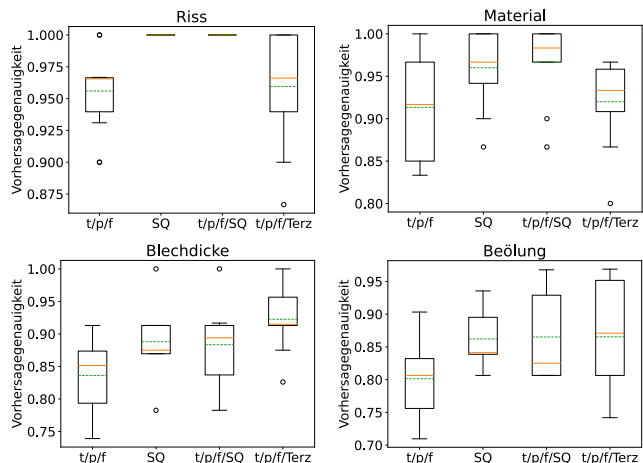
**Tabelle 2:** Anzahl der Datenpunkte im Trainings- und Testdatensatz für jede Zielgröße

Zielgröße	Punktezahl	
	Training	Test
Riss	264	30
Blechdicke	208	24
Beölung	280	32
Material	270	30

## Ergebnisse und Diskussion

Durch die Anwendung der 10-fachen Kreuzvalidierung wurden zehn verschiedene Vorhersagegenauigkeitswerte, d.h. einer für jeden Testdatensatz, berechnet. In Abbildung 2 sind diese Ergebnisse für verschiedene Merkmalsätze und Zielgrößen als Boxplot dargestellt. Der Boxplot zeigt die Streuung der Vorhersagegenauigkeit für die verschiedenen Trainingsdatensätze. Je kleiner die Box ist, desto unempfindlicher ist die Vorhersagegenauigkeit gegenüber Variationen im Trainingsdatensatz. In jedem Boxplot sind der Median und der Mittelwert, der im Vergleich zum Median gegenüber Ausreißer empfindlich ist, angegeben.

Zunächst ist zu erkennen, dass Risse, die während des Prozesses am Bauteil entstehen, von allen Modellen mit hoher Genauigkeit identifiziert werden können. Besonders hervorzuheben sind die beiden Modelle, die mit den Psychoakustik-Merkmalen (SQ) trainiert wurden. Durch die Verwendung der SQ-Merkmale wurden alle Risse im Fertigungsprozess richtig erkannt, weshalb im Diagramm nur eine Linie zu erkennen ist. Die beiden anderen Klassifikatoren, die mit den Standard-Merkmalen t/p/f und den Amplituden des Terzspektrums trainiert wurden, zeigen



**Abbildung 2:** Boxplot der Vorhersagegenauigkeit für verschiedene Merkmalsätze für alle Zielgrößen; Rote durchgezogene Linie: Median, grüne gestrichelte Linie: Mittelwert

hingegen eine deutlich geringere Vorhersagegenauigkeit im Vergleich zum Klassifikator mit SQ-Merkmalen. Aufgrund der ausgeprägten Signalcharakteristik beim Auftreten eines Risses (impulsförmiger Signalverlauf) sind die vergleichsweise hohen Genauigkeiten im Vergleich zu den anderen Zielgrößen plausibel. Für die weiteren betrachteten Zielgrößen *Material*, *Blechdicke* und *Beölung* sind die Unterschiede in den Vorhersagegenauigkeiten der zugrundeliegenden Klassifikatoren mit verschiedenen Merkmalen nicht so eindeutig. Auf Basis der dargestellten Boxplots kann zunächst festgehalten werden, dass die Verwendung von SQ-Merkmalen eine deutliche Steigerung der Genauigkeiten im Vergleich zur Verwendung von Standard-Signalparametern (t/p/f) erreichen lässt. Dies ist für alle Zielgrößen zu beobachten. Neben der höheren Genauigkeit kann eine Verringerung der Streuung festgestellt werden (vgl. bspw. Boxplots zur Blechdicke aus Abbildung 2), was als weiterer Vorteil der Berücksichtigung von SQ-Merkmalen angesehen werden kann. Weiterhin ist zu erkennen, dass die Verwendung von SQ-Parametern nicht für alle verwendeten Zielgrößen zu einer verbesserten Vorhersagegenauigkeit führt. Dies kann am Beispiel der Blechdicke verdeutlicht werden. Unter Verwendung der Amplituden des Terzspektrums kann eine größere Genauigkeit im Vergleich zur Nutzung der SQ-Merkmale festgestellt werden. Die verbesserte Genauigkeit unter Verwendung der Terzen deutet auf ausgeprägte Muster im Frequenzbereich am Beispiel der Zielgröße *Blechdicke* hin. Der Umformprozess größerer Blechdicken erfordert größere Prozesskräfte, die von der entsprechenden Antriebs- bzw. Lastmaschine aufgebracht werden müssen. Die Veränderung spektraler Anteile im betrachteten Luftschall erscheint damit plausibel und deutet auf eine bessere Differenzierbarkeit anhand spektraler Anteile hin. Allerdings beinhalten die SQ-Parameter die Lautheiten der einzelnen Frequenzgruppen und damit müsste auch unter Verwendung der SQ-Merkmale eine Bewertung spektraler Anteile erfolgen. Eine mögliche Erklärung liegt in der Lage pegeldominanter Anteile im Frequenzbereich. Es kann auf Grundlage der dargestellten Ergebnisse nicht ausgeschlossen werden, dass die zu erwartenden schmalbandigen Anteile der Lastmaschine an der Grenze zweier

Frequenzgruppen liegt und damit für verschiedene Messungen zwischen diesen beiden Gruppen hin und her wandert. Die Analyse der genannten Unterschiede und deren Ursachen sind Bestandteil weiterer Arbeiten anhand des gezeigten Datensatzes.

Abschließend kann festgehalten werden, dass die Verwendung psychoakustischer Kenngrößen im Rahmen der Maschinen- und Prozessüberwachung einen vielversprechenden Ansatz darstellt, um die Vorhersagegenauigkeit von Überwachungssystemen zu steigern. Insbesondere die sichere Erkennung von Rissen als Bauteilfehler aus dem Prozess kann an dieser Stelle hervorgehoben werden. Es lässt sich sicherlich argumentieren, dass speziell auf den Prozess angepasste Signalparameter gefunden werden können, die eine vergleichbar gute Vorhersagegenauigkeit ermöglichen. Allerdings sind auf einen Prozess angepasste Parameter im Sinne der Übertragbarkeit auf andere Prozesse als kritisch anzusehen.

### Zusammenfassung

Der vorliegende Beitrag beschäftigt sich mit der Anwendung psychoakustischer Kenngrößen (SQ-Merkmale) zur Klassifikation von Prozess- und Bauteilfehlern am Beispiel eines Tiefziehprozesses. Es konnte gezeigt werden, dass die Verwendung von SQ-Merkmalen einen vielversprechenden Ansatz darstellt, um die Vorhersagegenauigkeit eines Klassifikators (hier kNN) zu steigern. Die Beurteilung erfolgte anhand unterschiedlicher Modelle, die mit verschiedenen Signalmerkmalen trainiert und verglichen wurden. Die Detektion von Rissen als prozessbedingte Bauteilfehler ist hervorzuheben, da diese unter Verwendung von SQ-Merkmalen auf 100 % gesteigert werden konnte. Alle weiteren Zielgrößen *Material*, *Blechdicke* und *Beölung* konnten mit einer Genauigkeit größer 85 % klassifiziert werden. Die Verwendung von SQ-Merkmalen ist nicht an einen spezifischen Prozess gebunden und lässt sich prinzipiell auf andere Prozesse und Anlagen übertragen, wobei ein separates Training erforderlich ist. Als weiterer Vorteil ist die notwendige (Standard-)Messtechnik sowie die nicht erforderliche Applikation von Sensoren im Prozessraum anzusehen. Die Übertragbarkeit auf weitere Prozesse ist Bestandteil weiterer Arbeiten.

### Literatur

- [1] Straube, O.: Körperschallanalyse an Tiefziehwerkzeugen. Zeitschrift für wirtschaftlichen Fabrikbetrieb **88**(10) (1993), 457-459
- [2] Hao, S. et al.: Acoustic emission monitoring of sheet metal forming: characterization of the transducer, the work material and the process. Journal of Materials Processing Technology 101 (2000), Issues 1–3, 124-136
- [3] DFG-Projekt *Akustische Emissionsanalyse zur Online-Prozessüberwachung in der Blechumformung*, URL: <https://gepris.dfg.de/gepris/projekt/263066528>, abgerufen am 27.03.2024
- [4] Tsuruya, T. et al.: Forming state recognition in deep drawing process with machine learning. Journal of Advanced Mechanical Design, Systems, and Manufacturing **13**(3) (2019), JAMDSM0066
- [5] Mühlbauer, M. et al.: Automated data labeling and anomaly detection using airborne sound analysis. Procedia CIRP 93 (2020), 1247–1252
- [6] Schäfer, C. et al.: Körperschallanalyse zur Überwachung von Präzisions- und Ultra-Präzisionsfertungsverfahren. 16. Kolloquium Schallemission, DGZfP-Berichtsband BB 105 (2007)
- [7] Trautmann, F. et al.: Subjective hearing sensation of process variations at a milling machine. How reliable will chatter marks be detected?. INTER-NOISE and NOISE-CON Congress and Conference Proceedings **263**(2) (2021), 4088-4099
- [8] Kane, P.V. und Andhare, A.B.: Application of psychoacoustics for gear fault diagnosis using artificial neural network. Journal of Low Frequency Noise, Vibration and Active Control **35**(3) (2016), 207-220
- [9] Rafaelof, M. und Schroeder, A.: Investigation of machine learning algorithms to model perception of sound. Proceedings of Meetings on Acoustics **33**(1) (2018)
- [10] Bowen, D.L.: Correlating Sound Quality Metrics and Jury Ratings. Journal of Sound & Vibration **42**(9) (2008), 12
- [11] Lopes, B. et al.: Neural Network Models For the Subjective And Objective Assessment Of A Propeller Aircraft Interior Sound Quality. INTER-NOISE and NOISE-CON congress and conference proceedings **259**(5) (2019), 4124-4135
- [12] Lopez-Ballester, J. et al.: Computation of psychoacoustic annoyance using deep neural networks. Journal of Applied Sciences, **9**(15) (2019), 3136
- [13] Sottek, R. und Henrique, T.: AI-SQ Metrics: Artificial Intelligence in Sound Quality Metrics. INTER-NOISE and NOISE-CON Congress and Conference Proceedings **261**(3) (2020), 3083-3091
- [14] Hui, K. et al.: An improved wrapper-based feature selection method for machinery fault diagnosis. PloS one **12**(12) (2017), e0189143
- [15] Brandt, A.: Noise and vibration analysis: signal analysis and experimental procedures. John Wiley & Sons, 2011.
- [16] Green Forge Coop. MOSQUITO (v1.1.0), URL: <https://doi.org/10.5281/zenodo.5284054>, abgerufen am 27.03.2024
- [17] Hastie, T., et al.: The elements of statistical learning: data mining, inference, and prediction. Vol. 2. New York: Springer, 2009
- [18] Horenko, I.: On a scalable entropic breaching of the overfitting barrier for small data problems in machine learning. Neural Computation **32**(8) (2020): 1563-1579.
- [19] Pedregosa, F., et al.: Scikit-learn: Machine learning in Python. The Journal of machine Learning research **12** (2011), 2825-2830