Bei diesem Beitrag handelt es sich um einen wissenschaftlich begutachteten und freigegebenen Fachaufsatz („reviewed paper“).

doi.org/10.37544/1436-4980-2026-06-44
Datum der Einreichung: 03.03.2026
Datum der Annahme: 28.04.2026
Datum der Veröffentlichung: 27.07.2026

Datengetriebene Wertschöpfung durch Clusterverfahren in Scherschneidprozessen

Datengetriebene Wertschöpfung durch Clustering

H. von Linde, M. Lück, A. Schenek, O. Riedel

ZUSAMMENFASSUNG Zuverlässige Zustandsüberwachung im Scherschneiden scheitert oft an fehlenden Labels und überlagernden Werkstoffeffekten in Kraft-Zeit-Kurven. Der Beitrag untersucht ein rein unüberwachtes Vorgehen, das Kurven gezielt aufbereitet und anschließend mit mehreren Clustering-Familien strukturiert. Es wird gezeigt, welche Schritte der Datenrepräsentation die Trennbarkeit von Werkzeugzuständen maßgeblich prägen und wo materialbedingte Einflüsse Verschleißsignaturen verdecken. Daraus leitet sich ab, wie Clustering als Vorverdichtung und zur Driftbeobachtung in robuste Monitoring-Pipelines eingebettet werden kann.

STICHWÖRTER
Scherschneiden, Zustandsüberwachung, Clustering

Data-driven value creation using clustering

ABSTRACT Reliable condition monitoring in shear cutting is often limited by missing labels and material-related variability in force-time curves. This paper studies a fully unsupervised workflow that transforms the curves into a consistent representation and then structures the data using multiple clustering families. As a contribution, it highlights which representation choices primarily drive the separability of tool states and where material effects can obscure wear signatures. The results are positioned in terms of practical benefit: clustering can serve as data pre-aggregation and drift monitoring to support robust, scalable monitoring pipelines.

1 Einleitung

Scherschneid- und Schneidprozesse sind zentrale Operationen der Blechumformtechnik [1]. Sie erlauben hohe Taktzahlen bei geringen Stückkosten, sind aber stark vom Werkzeugzustand und von Werkstoffkennwerten abhängig [2]. In der Praxis werden Stempel und Matrizen meist in festen Intervallen nachgeschliffen oder getauscht, zum Beispiel nach einer vorgegebenen Anzahl von Pressenhüben [2–4]. Das reduziert ungeplante Stillstände, führt aber dazu, dass Werkzeuge entweder zu früh gewartet werden (ungenutzte Reststandzeit) oder zu spät, was Ausschuss, Gratbildung, Werkzeugbruchrisiken und Folgekosten erhöht.

Die Intervallwartung ignoriert Belastungs- und Verschleißstreuung und erzwingt Sicherheitsmargen, die den Prozess zwar robust machen, aber Potenziale bei Prozesszeit, Werkzeugnutzung und Qualitätsoptimierung verschenken.

Eine zustandsorientierte Wartung erfordert messgrößenbasierte Indikatoren. In Scherschneidprozessen ist die Schneidkraft geeignet, weil sie viele Einflussgrößen bündelt und oft ohne invasive Sensorik verfügbar ist [5, 6]. Die zentrale Herausforderung ist ihre Ambiguität: Mehrere Ursachen können ähnliche Änderungen der Kraftkurve bewirken [7, 8]. Algorithmen müssen daher die für die Ursache verantwortlichen Effekte entkoppeln oder Invarianten bilden, die Störgrößen unterdrücken.

Eine datengetriebene Zustandsüberwachung koppelt Wartungsentscheidungen direkt an Prozesssignaturen. Als Sensorsignal bietet sich die Schneidkraft beziehungsweise die Kraft-Zeit- oder Kraft-Weg-Kurve pro Hub an [5, 6, 9–11]. In idealisierten Versuchsreihen unterscheiden sich diese Kurven zwischen Werkzeugzuständen oft deutlich. In der realen Fertigung überlagern aber Werkstoffschwankungen (zum Beispiel Zugfestigkeit innerhalb eines Coils), thermische Effekte, Schmierungseinflüsse sowie Drift- und Alterungseffekte die Signale [12]. Entscheidend ist, unter welchen Randbedingungen robuste, generalisierbare Aussagen möglich sind.

Verschleiß wirkt typischerweise graduell [2]. Die Signaturänderung ist daher kein sprunghafter Zustandswechsel, sondern ein Drift im Merkmalsraum. Ein Clustering, das diskrete Gruppen erwartet, kann Drift nur dann sinnvoll abbilden, wenn die Repräsentation entweder klare Regimewechsel sichtbar macht oder Driftabschnitte als kompakte Regionen repräsentiert.

Dieser Beitrag untersucht eine technologisch bewusst einfach umzusetzende Methodik. **Bild 1** zeigt den Versuchsaufbau schematisch.

Es wird kein überwachtes Modell trainiert, vielmehr werden Ähnlichkeiten zwischen Hubsignalen in einer Feature-Darstellung gesucht und durch Clusterverfahren gruppiert. Viele Fertigungsdaten liegen ungelabelt vor und aufwendiges Feature Engineering erfordert Domänenwissen, das nicht immer verfügbar ist. Der Schwerpunkt liegt daher auf (i) minimaler, reproduzierbarer Vorverarbeitung, (ii) dem Vergleich mehrerer Clustering-Familien und (iii) einer quantitativen Validierung an einer Fallstudie.

Die Fallstudie adressiert einen konkreten Anwenderwunsch: Die Verschleißerkennung soll möglichst unabhängig vom Werkstoff funktionieren. Eine explizite Werkstoffklassifikation wird nicht als Ziel betrachtet, weil sie für Wartungsentscheidungen

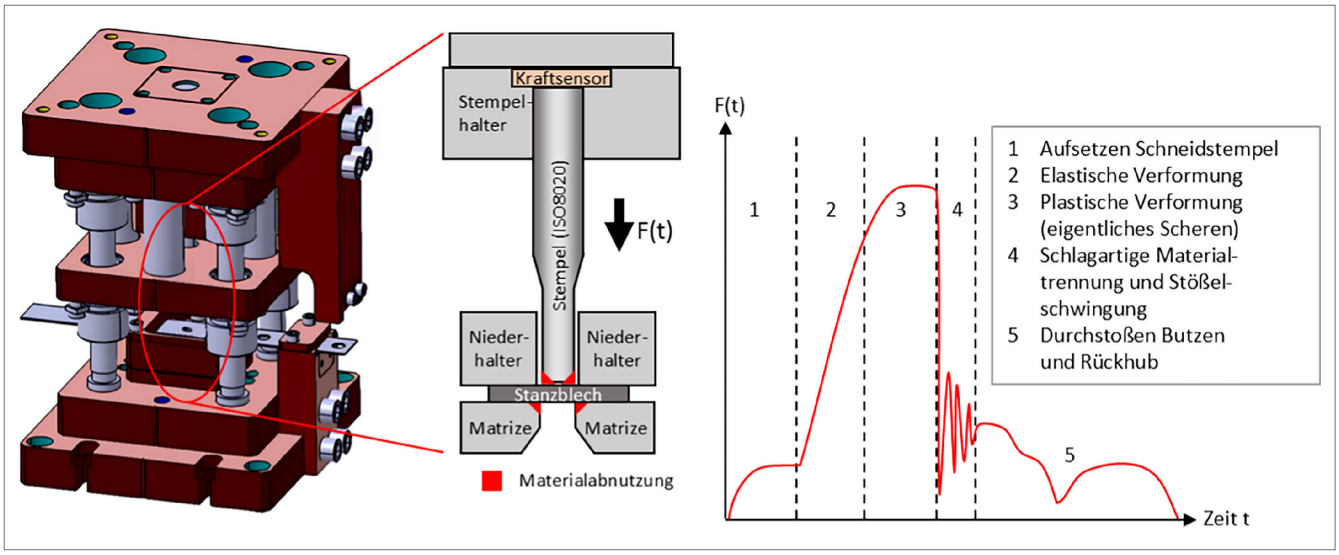

Bild 1 Rendering des Versuchsaufbaus und Querschnitt. Darstellung der Abschnitte der Kraftkurve. *Grafik: Rendering nach [6], Kraftkurve nach [13].*

sekundär ist und in der Praxis durch Materialstreuung leicht verfälscht werden kann [2, 13].

Der Nutzen des Ansatzes ist nicht zwingend eine vollautomatische Diagnose im Sinne einer „harten" Zustandsklassifikation. Bereits eine robuste Strukturierung der Daten, etwa die Trennung von stabilen Regimen, transienten Phasen und Ausreißerhüben, ist wertschöpfend, weil sie Ursachenanalysen fokussiert und gezielte Validierungsmessungen ermöglicht. Clustering wird daher als Baustein verstanden, der zwischen Datenerfassung und Wartungsentscheidung eine datenbasierte Vorverdichtung liefert [2, 5].

Der Beitrag dieser Arbeit ist (1) eine reproduzierbare Pipeline für kurvenbasierte Scherschneidkraft-Clusterung ohne domänenspezifische Kennwerte, (2) eine familienübergreifende Gegenüberstellung von partitionierungsbasierten, hierarchischen, dichte- und graphbasierten Verfahren bei konstanter Parametrisierung, (3) die empirische Evidenz, dass die Datenrepräsentation (Normierung/Alignment) die Trennbarkeit stärker bestimmt als die Algorithmuswahl, (4) eine Diskussion der Grenzen der Werkstoffunabhängigkeit und der Trade-offs zwischen Invarianz und Sensitivität sowie (5) ein konkreter Experimentplan für robuste, werkstofftolerante Inline-Überwachung.

2 Stand der Technik

Die Auswertung von Prozesssignalen zur Zustands- und Qualitätsüberwachung ist in der Scherschneidtechnik etabliert, jedoch meist als Grenzwert- oder Hüllkurvenüberwachung implementiert. Solche Systeme detektieren Störungen zuverlässig erst, wenn Abweichungen definierte Schwellen überschreiten. Für Assistenzsysteme wird daher eine feinere Diagnostik gefordert, die auch kleine Parameteränderungen identifiziert, Ursachen unterscheidet und Handlungsempfehlungen ableitet. Kraft-Weg- oder Kraft-Zeit-Kurven werden hierfür als informationsreiche Signaturen genutzt.

Kubik et al. untersuchen Kraft-Weg-Kurven beim Scherschneiden mit dem Ziel, Eigenschaftsabweichungen zu identifizieren und Korrelationen zwischen Qualitätsmerkmalen und Kurven-Features aufzubauen [2]. Kurven werden in charakteristische Kennwerte überführt, die anschließend in Korrelationen und Regressionsmodellen für ein Entscheidungsunterstützungssystem genutzt werden. Gleichzeitig wird betont, dass (a) viele Einflussgrößen den Prozess prägen (Werkzeug, Werkstoff, Maschine, Prozessführung), (b) industrielle Abweichungen oft nur wenige Prozent betragen und (c) bei hohen Hubzahlen dynamische Effekte die Signale überlagern können. Ein rein datengetriebener Ansatz ohne Prozesswissen kann daher an korrelierten Störgrößen scheitern, während prozessgetriebene Ansätze zusätzlichen Modellierungs- und Engineering-Aufwand erfordern.

Weitere Untersuchungen analysieren den Informationsgehalt in-situ gemessener Schneidkraft-Weg-Verläufe für Zustands- und Qualitätsaussagen [2, 8, 12, 14, 15]. *Hoppe et al.* schlagen eine phasenbasierte Zerlegung der Stempelkraft-Weg-Kurve in Schneid-, Durchdrück- und Abstreif-/Rückhubphase vor [14]. Darauf aufbauend werden Feature-Engineering-Ansätze und Feature-Extraction-Methoden wie die Principal Component Analysis (PCA) in Kombination mit Klassifikatoren (wie etwa Support-Vektor-Machine, SVM) eingesetzt [14, 15]. Es wird gezeigt, dass eine differenzierte Betrachtung der Kraft-Weg-Signatur sowohl die Klassifikation von Verschleißzuständen an Schneidaktivelementen als auch die Kopplung zu nachgelagerten Operationen (zum Beispiel Vorhersage von Biegewinkeln) ermöglicht [14, 15]. Ergänzend untersuchen *Kubik et al.* den Einfluss von Datenerfassung, Vorverarbeitung und Transformation auf die Echtzeitfähigkeit einer Werkzeugzustandserkennung [8, 15].

Für die vorliegende Arbeit sind zwei Schlussfolgerungen zentral. Erstens ist eine stabile zeitliche Ausrichtung des Signals entscheidend, weil die semantische Bedeutung eines Kurvenabschnitts (wie etwa Schneidphase versus Nachlauf) nur bei vergleichbarer Phasenlage über verschiedene Hübe hinweg erhalten bleibt. Zweitens ist die Merkmalsdarstellung der dominierende Freiheitsgrad, weil sie festlegt, ob Variationen primär durch Werkstoff, Prozesszustand oder Werkzeugzustand getrennt werden.

Mit den technologischen Fortschritten lassen sich Prozess- und Sensordaten heute in großer Menge und zu geringen Kosten erfassen [16]. Diese Datenfülle führt jedoch zum von *Bellman* beschriebenen „Curse of Dimensionality". Mit wachsender Merkmalsanzahl steigen Rechenaufwand und Speicherbedarf stark an, während Visualisierung und robuste Modellbildung erschwert werden [17, 18]. Dimensionsreduktion ist daher ein wesentlicher

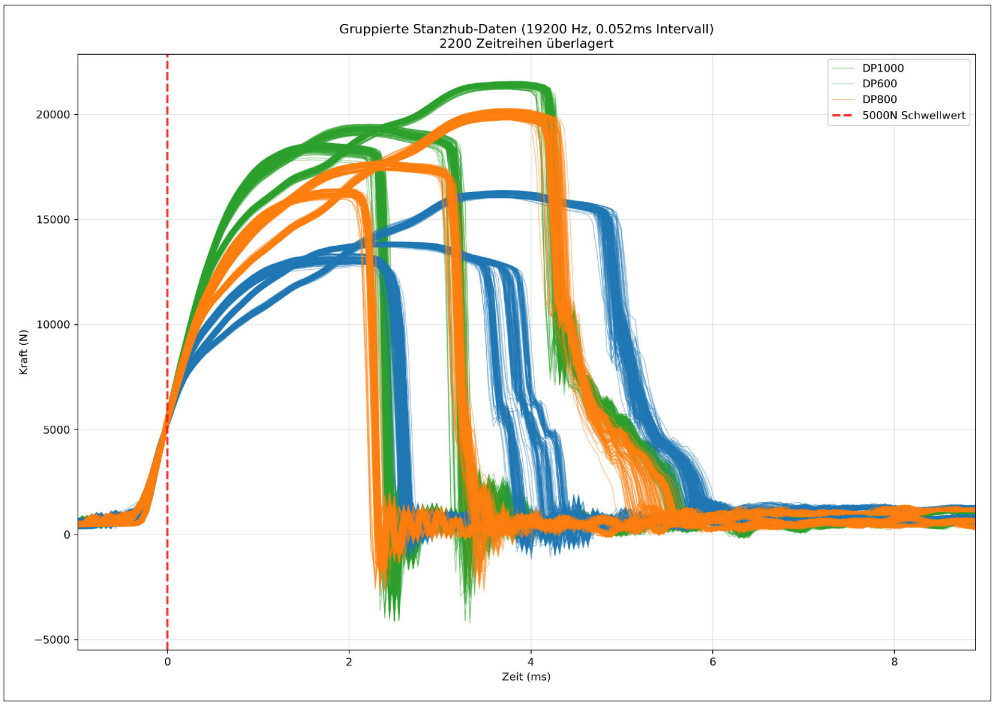

Bild 2 Kraftverläufe der Scherschneidversuche nach Werkstoff und Werkzeugzustand. Eingefärbt nach Materialart. *Grafik: Eigene Darstellung*

Baustein des maschinellen Lernens [19]. Maßnahmen reichen von Vorverarbeitungen ohne Änderung der Dimensionalität (zum Beispiel Normalisierung, Standardisierung) bis zu Verfahren, die den Merkmalsraum gezielt verkleinern oder – etwa durch nichtlineare Expansionen – vergrößern können [19, 20].

Trotz uneinheitlicher Taxonomie lassen sich im Kern vier Strategiefamilien unterscheiden [19–23]. Feature Extraction erzeugt automatisiert neue Merkmale aus Rohdaten (zum Beispiel PCA, Wavelets, Einbettungen). Feature Selection wählt relevante Teilmengen vorhandener Merkmale anhand statistischer Kriterien beziehungsweise Suchstrategien (Filter-, Wrapper-, Embedded-Ansätze) [20–23]. Feature Engineering beschreibt die manuelle, wissensgetriebene Konstruktion prozessspezifischer Kennwerte durch Domänenexperten [24–26]. Feature Importance bewertet Merkmale erklärungsorientiert (global/lokal) und dient häufig als Grundlage für Auswahlentscheidungen [27–30].

Feature Engineering ist typischerweise gut interpretierbar und nah an Physik und Prozesslogik, bindet jedoch Expertenwissen, verursacht hohen Zeitaufwand und skaliert nur begrenzt über Werkstoffe, Werkzeuge und Anlagen [24–26]. Feature Extraction reduziert den manuellen Aufwand und kann eine komprimierte Datenrepräsentation liefern, verlangt in der Praxis aber große, qualitativ hochwertige Datenmengen, sorgfältige Synchronisation sowie robuste Strategien gegen zeitliche Drift und Nichtlinearitäten [20, 31, 32].

Aus industrieller Sicht ist zudem die Labelknappheit relevant. Eine Ground-Truth-Erhebung zum Werkzeugzustand ist teuer, weil sie Prozessunterbrechungen oder Laboranalysen erfordert. Unüberwachte Verfahren sind daher attraktiv als Vorstufe. Sie strukturieren Datenräume, identifizieren Ausreißer und liefern Hypothesen über Prozessregime, bevor gezielte Validierungen erfolgen.

Clustering ist in diesem Kontext attraktiv, weil es ohne Labeling auskommt und als Erstdiagnose oder Strukturierungsschritt dienen kann. Zu beachten ist, dass Cluster nicht automatisch Zustände sind, sondern Gruppen im gewählten Merkmalsraum. Wenn dominierende Effekte (wie etwa Kraftniveau durch Zugfestigkeit) die Struktur bestimmen, spiegelt das Clustering primär diese Effekte wider. Eine fachlich sinnvolle Zustandsüberwachung erfordert daher eine methodische Klammer aus Vorverarbeitung, Merkmalswahl, algorithmischer Robustheit und validierter Zuordnung der Cluster zu Zustandsklassen.

Diese Arbeit verzichtet bewusst auf umfangreiches Feature Engineering. Es werden keine domänenspezifischen Kennwerte wie Maximalwerte, Integrale einzelner Phasen oder geometrische Kantenparameter konstruiert. Stattdessen wird eine feingranulare, aber einfache Kurvenrepräsentation genutzt, die als Feature-Matrix direkt aus den Zeitreihen entsteht. Der Vorteil ist die Reduktion des Engineering-Aufwands, der Nachteil liegt darin, dass physikalisch motivierte Invarianten nicht automatisch entstehen, sondern durch Normierung und Standardisierung nur approximiert werden.

3 Methode

3.1 Datenbasis und Zielsetzung

Die Fallstudie basiert auf 2200 auswertbaren Scherschneidhüben aus Versuchen an drei Dualphasenstählen (DP600, DP800, DP1000). Die Versuchsplanung ist voll faktoriell hinsichtlich Werkstoff- und Werkzeugzustand angelegt. Pro Werkstoff wurden 750 Hübe aufgenommen, verteilt auf die drei Werkzeugzustände scharf, mittel und kaputt mit jeweils 250 Hüben. Vorgesehen waren 2250 Hübe; nach Ausschluss fehlerhafter beziehungsweise unvollständiger Messungen verblieben 2200 Datensätze.

Die Schneidkraft wird mit einer im Sockel integrierten Kraftmessdose mit einer Abtastrate von 19,2 kHz erfasst; Bild 1 zeigt den Aufbau. Pro Hub liegt ein Kraft-Zeit-Verlauf als Zeitreihe vor, ein Plot aller Rohdaten ist in **Bild 2** dargestellt. In der für das Clustering verwendeten, diskretisierten Darstellung umfasst der Datensatz rund 420 000 Stützstellenwerte. Zu jedem Hub stehen Metadaten (Werkstoff, Werkzeugzustand, Dateiname, Peak-Index) zur Verfügung.

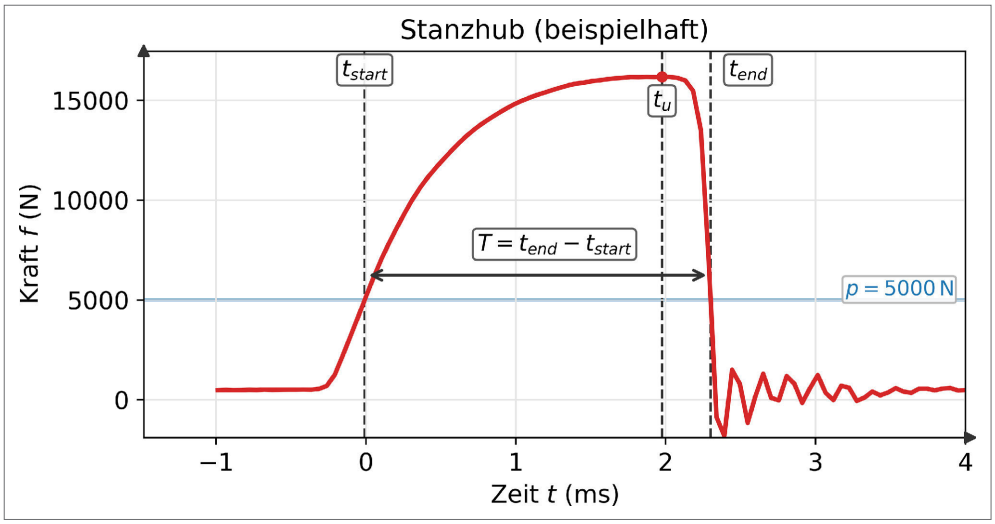

Bild 3 Beispielhafte Aufteilung eines Scherschneidhubs mit Zeitpunkten t_{start}, t_u, t_{end}. *Grafik: Eigene Darstellung*

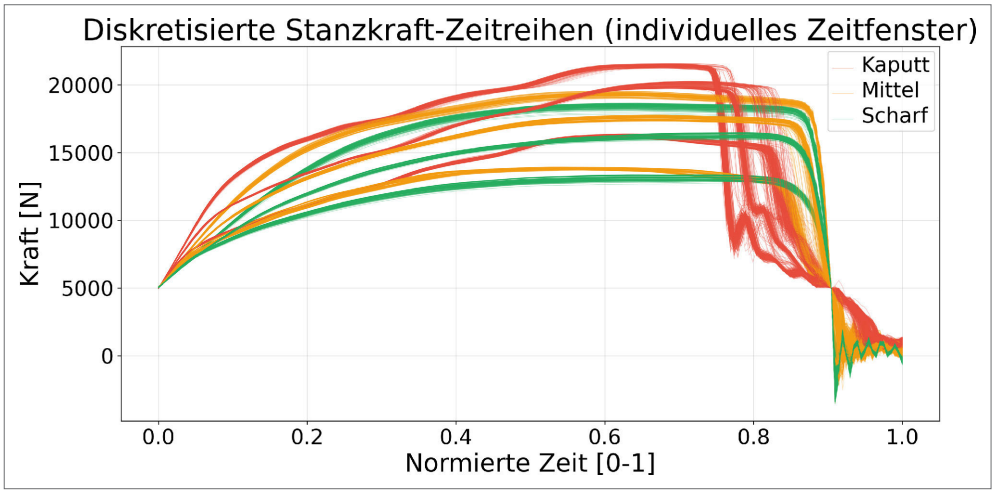

Bild 4 Darstellung der diskretisierten Kraftverläufe der Scherschneidversuche nach Werkstoff und Werkzeugzustand. Eingefärbt nach Werkzeugzustand. *Grafik: Eigene Darstellung*

Die Werkzeuge werden in drei definierten Zuständen eingesetzt: scharf, mittel und kaputt. Die Zustände mittel und kaputt werden gezielt im Labor hergestellt und entsprechen einer Kantenrundung von 50 µm beziehungsweise 100 µm. Damit sind die Zustände reproduzierbar beziehungsweise deterministisch. Als Messgröße dient der Kraftverlauf pro Hub.

Die übergeordnete Zielsetzung ist anwendungsnah. Es wird geprüft, ob Werkzeugzustände aus Kraftsignalen automatisch geclustert werden können – möglichst unabhängig vom Werkstoff und ohne nutzerorientiertes Feature Engineering im Vorfeld der Datenverarbeitung. Eine explizite Werkstoffklassifikation wird nicht verfolgt, weil Schwankungen innerhalb eines Coils oder Materialwechsel in der Praxis die Übertragbarkeit einer reinen Materialerkennung einschränken.

3.2 Vorverarbeitung: Fensterung und Diskretisierung

Jede Zeitreihe wird zuerst auf ein reproduzierbares Zeitfenster um die Schneidphase reduziert. Dazu wird der Zeitpunkt t_{start} bestimmt, an dem die Kraft erstmals eine definierte Schwelle F_{th} überschreitet (hier 5000 N). Anschließend wird t_u ermittelt, an dem die Kraft nach der Überschreitung wieder unter F_{th} fällt. Das Fensterende wird als $t_{end} = t_u + \Delta t$ gesetzt, wobei Δt eine kurze Nachlaufzeit abbildet.

Die Dauer der Schneidphase variiert zwischen den Hüben infolge unterschiedlicher Schneidkantenverrundungen und Duktilität der Blechwerkstoffe. Um Kurven vergleichbar zu machen, wird die Zeitachse normiert und das Fenster in B diskrete Intervalle zerlegt (hier $B = 200$). Zusätzlich wird der Unterschreitungszeitpunkt t_u auf eine feste normierte Position $p \in (0,1)$ abgebildet (hier $p = 0{,}9$). Bereiche vor und nach der Unterschreitung werden getrennt skaliert, sodass die charakteristische Phasenlage in allen Samples konsistent repräsentiert ist (**Bild 3**).

Formal wird die relative Zeit $\tau = t_{end} - t_{start}$ betrachtet und in die Intervalle $[0, \tau_u)$ und $[\tau_u, t_{end}]$ mit $\tau_u = t_u - t_{start}$ und $\tau = t_{end} - t_{start}$ geteilt. Die Abbildung auf die normierte Zeit $x \in [0,1]$ erfolgt stückweise linear. Für $\tau \in [0,t_u)$ gilt $x = p\,\tau/t_u$; für $\tau \in [t_u, t_{end}]$ gilt $x = p + (1-p)(T-\tau_u)/(t_{end}-\tau_u)$. Anschließend werden B Stützstellen x_i gleichmäßig auf $[0,1]$ verteilt und die Kraftwerte durch lineare Interpolation abgetastet. Damit ergibt sich pro Hub ein Vektor $\mathbf{f} \in \mathbb{R}^B$.

Das Ergebnis ist eine Feature-Matrix, in der jede Zeile einen Hub und jede Spalte einen normierten Zeitstützpunkt beschreibt (**Bild 4**). Domänenspezifische Kennwerte wie maximale Kraft, Scherflächenanteile oder Gratparameter werden bewusst nicht gebildet, stattdessen wird der Kurvenverlauf direkt abgebildet.

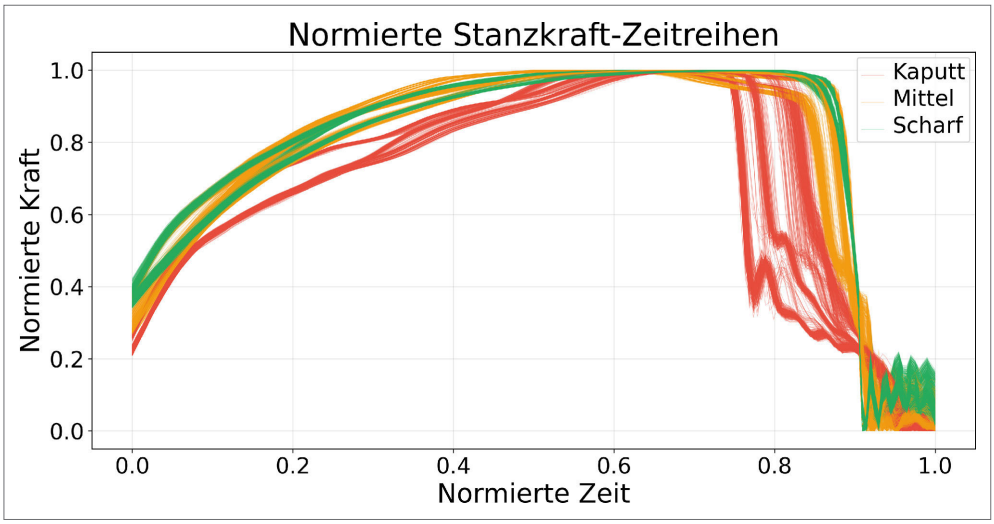

Bild 5 Darstellung der normierten Kraftverläufe der Scherschneidversuche nach Werkstoff und Werkzeugzustand. Eingefärbt nach Werkzeugzustand. *Grafik: Eigene Darstellung*

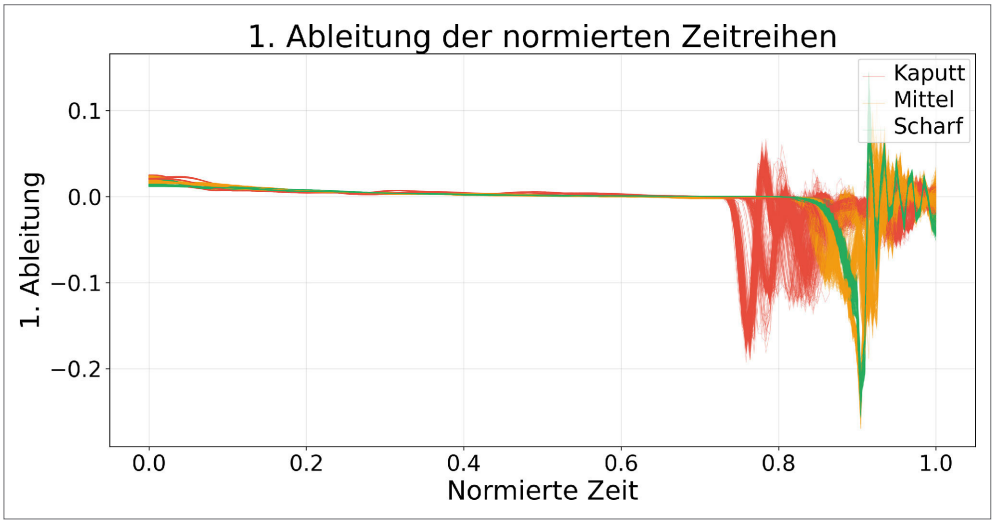

Bild 6 Darstellung der 1. Ableitung der normierten Kraftverläufe der Scherschneidversuche nach Werkstoff und Werkzeugzustand. Eingefärbt nach Werkzeugzustand. *Grafik: Eigene Darstellung*

3.3 Normierung und Ableitungen

Auf Basis der diskretisierten Kurven werden drei Varianten betrachtet:

- Normierte Kurve: Min-Max-Normierung der Kraftwerte auf $[0,1]$ pro Hub (**Bild 5**)
- Erste Ableitung: numerische Ableitung der normierten Kurve mittels <Dreieck steht auf Spitze> (Gradient) als Proxy für Steigungs- beziehungsweise Änderungsratenmuster (**Bild 6**)
- Zweite Ableitung: numerische zweite Ableitung als Proxy für Krümmungs- beziehungsweise Beschleunigungsmuster (**Bild 7**)

Die Ableitungen sollen Form- und Gradienteninformationen hervorheben, während absolute Niveaus (zum Beispiel werkstoffbedingte Kraftamplituden) abgeschwächt werden. Jede Transformation erhöht jedoch die Sensitivität gegenüber Rauschen und Synchronisationsfehlern, dieser Zielkonflikt wird im Validierungsteil quantitativ bewertet.

Die Min-Max-Normierung pro Hub ist definiert als $\tilde{f}_i = \frac{f_i - \min_j f_j}{\max_j f_j - \min_j f_j}$. Die numerische Ableitung erfolgt diskret über ∇ auf dem äquidistanten Gitter x_i. Die erste Ableitung approximiert lokale Steigungen, die zweite lokale Krümmungen. Diese Operationen erhöhen die Formempfindlichkeit, sind aber besonders sensitiv gegenüber Rauschen und kleinen Phasenverschiebungen.

3.4 Clustering-Verfahren

Es werden acht Verfahren aus unterschiedlichen Clustering-Methoden eingesetzt, um die Robustheit gegenüber Modellannahmen zu prüfen:

- Partitionierend: k-Means
- Hierarchisch agglomerativ: Single Linkage, Complete Linkage, Average Linkage, Ward
- Hierarchisch (BIRCH): CF-Tree-basierte Vorstrukturierung mit anschließender Clusterbildung
- Dichtebasiert: DBSCAN
- Graph-basiert: Spectral Clustering

Für alle Verfahren außer DBSCAN wird die Zielanzahl der Cluster auf $k = 3$ gesetzt, um die drei erwarteten Werkzeugzustände abzubilden. DBSCAN bestimmt die Clusterzahl indirekt über ε und min_samples.

Vor dem Clustering werden die Feature-Dimensionen standardisiert (z-Transformation), um unterschiedliche Skalen und Varianzen der Feature-Spalten auszugleichen. Dies ist insbesondere

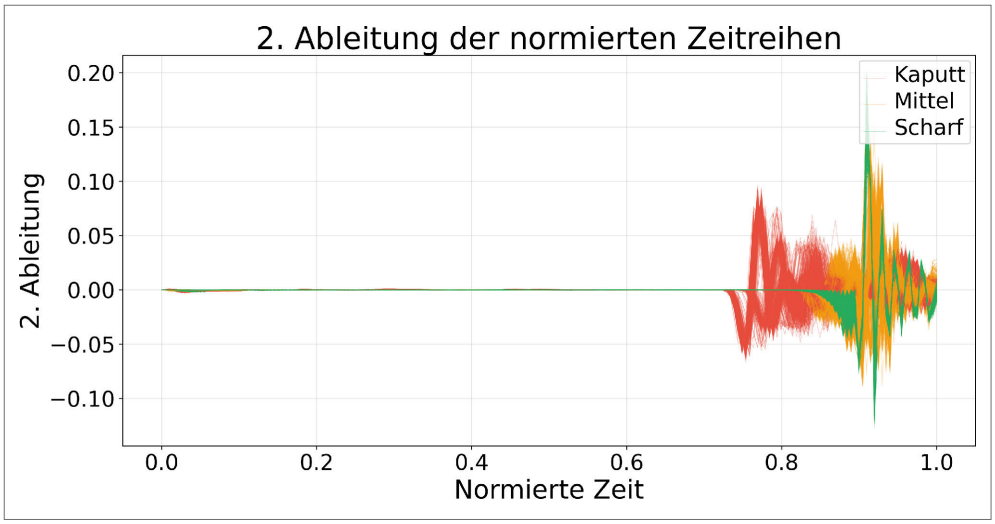

Bild 7 Darstellung der 2. Ableitung der Kraftverläufe der Scherschneidversuche nach Werkstoff und Werkzeugzustand. Eingefärbt nach Werkzeugzustand.
Grafik: Eigene Darstellung

Tabelle 1 Zentrale Pipeline-Parameter der Fallstudie.

Kategorie	Parameter	Wert
Fensterung	Kraftschwelle Fth	5.000 N
Fensterung	Nachlauf Δt	1 ms
Diskretisierung	Buckets B	200
Diskretisierung	Unterschreitungsposition p	0,9
Clustering	Ziel-Clusterzahl k	3 (außer DBSCAN)
Skalierung	Standardisierung	z-Transformation pro Feature

für distanzbasierte Verfahren relevant und wird als $\tilde{f}_i = \frac{\tilde{f}_i - \mu_i}{\sigma_i}$ implementiert, wobei μ_i und σ_i Mittelwert und Standardabweichung der i-ten Feature-Spalte über alle Stanzhübe sind. f_i steht für den Kraftwert am i-ten normierten Zeitstützpunkt eines Hubs, $\tilde{f}_i$ für den gleichen Wert nach Min-Max-Normierung pro Hub und $\tilde{f}_i$ für die z-standardisierten Featurewerte. Die Standardisierung reduziert die Dominanz einzelner Zeitpunkte mit hoher Varianz.

Die Wahl mehrerer Algorithmus-Familien dient als Robustheitsprüfung. k-Means ist eine Partitionierungsbaseline mit konvexen Clusterannahmen. Agglomerative Linkage-Varianten unterscheiden sich in der impliziten Clusterform (Kettenbildung versus kompakte Gruppen). Ward minimiert Varianzanstiege beim Mergen. BIRCH approximiert Cluster über eine hierarchische Kompressionsstruktur. DBSCAN modelliert dichtebasierte Cluster und kann Drift eher als Kontinuum abbilden, erfordert aber eine robuste Parametrisierung. Spectral Clustering operiert auf einer Ähnlichkeitsmatrix und kann nicht-konvexe Strukturen repräsentieren, ist jedoch sensitiv gegenüber der Graphkonstruktion.

3.5 Implementierte Pipeline und Parameter

Die Methodik ist als wiederholbare Pipeline umgesetzt. Die wesentlichen Parameter werden nicht pro Datensatz angepasst, sondern einmalig festgelegt, um eine konservative Bewertung zu erhalten. **Tabelle 1** fasst die zentralen Einstellungen zusammen.

Neben diesen Parametern existieren verfahrensspezifische Einstellungen (zum Beispiel Initialisierungen bei k-Means oder ε bei DBSCAN). Da die betrachteten Datenvektoren hochdimensional sind (B = 200), ist die Parametrisierung insbesondere für dichtebasierte Verfahren kritisch. Der Vergleich mehrerer Familien fungiert damit zugleich als Sensitivitätsanalyse gegenüber impliziten Geometrieannahmen.

Die hohe Dimensionalität ist zugleich methodischer Kern und Risiko. Sie erlaubt, den Kurvenverlauf ohne starke Informationsverdichtung abzubilden, erschwert aber Trennschärfe und Robustheit. In hohen Dimensionen verlieren euklidische Distanzen an Aussagekraft und Rauschen beeinflusst viele Feature-Dimensionen gleichzeitig. Klassische Signalpipelines adressieren dies durch explizite Dimensionsreduktion (zum Beispiel PCA) oder domänenspezifische Kennwerte. Darauf wird bewusst verzichtet, um die Aussage „ohne aufwendiges Feature Engineering“ nicht zu verwässern. Für Folgearbeiten ist jedoch zu prüfen, ob eine schwachparametrische Reduktion (zum Beispiel wenige Hauptkomponenten) die Robustheit gegenüber Werkstoff- und Prozessstreuung erhöht, ohne die Interpretierbarkeit vollständig zu verlieren.

3.6 Evaluationskonzept

Clustering ist unüberwacht, dennoch wird für die Fallstudie eine ex-post Validierung durchgeführt, weil die Werkzeugzustände in den Metadaten bekannt sind. Die Clusterlabels werden dazu mit den Zustandslabels verglichen. Da Clusterlabels beliebig per-

Tabelle 2 Ausgewählte Evaluationsergebnisse (Werkzeugzustand als Referenz). Werte aus der Pipeline-Auswertung.

Verfahren	Datensatz	Jaccard	ARI	NMI	Silhouette
k-Means	Normiert	0,578	0,578	0,71	0,5
Ward	Normiert	0,578	0,578	0,71	0,5
BIRCH	Normiert	0,578	0,578	0,71	0,5
Average Linkage	1. Ableitung	0,51	0,46	0,65	0,28
BIRCH	2. Ableitung	0,51	0,46	0,65	0,11
DBSCAN	1. Ableitung	0,33	0	0	0
Spectral	Normiert	0,31	0,05	0,13	0,07

mutiert sind, kommen paar- und informationsbasierte Metriken zum Einsatz:

- Jaccard-Index (paarbasiert)
- Adjusted Rand Index (ARI)
- Normalized Mutual Information (NMI)
- Silhouette-Score (intra-/interclusterbezogen; ohne Ground Truth)

Dieses Metrikbündel adressiert zwei Perspektiven, die Übereinstimmung mit der Zustandsstruktur (Jaccard, ARI, NMI) sowie die geometrische Trennbarkeit im Merkmalsraum (Silhouette). In industriellen Anwendungen ist letztere relevant, da Ground Truth häufig nicht verfügbar ist.

Zusätzlich zur Zustandsvalidierung wird die Werkstoffrobustheit der Clusterstruktur ausgewertet. Dazu werden die Clusterlabels sowohl mit den Zustandslabels als auch mit den Materiallabels (DP600/DP800/DP1000) verglichen. Dazu werden analog zu ARI/NMI auch Kontingenztafeln „Cluster × Werkstoff“ sowie „Cluster × Zustand“ berichtet. Ziel ist, empirisch zu prüfen, ob Cluster primär Werkzeugzustände oder lediglich Werkstoff- beziehungsweise Prozesseffekte abbilden.

Ein hoher Silhouette-Wert zeigt eine geometrisch plausible Partition, garantiert aber keine korrekte Zustandsdiagnose. Umgekehrt kann eine moderate Silhouette bei sinnvoller Zustandszuordnung auftreten, wenn Übergänge weich sind oder Werkstoff- und Verschleißeffekte überlagert wirken. Daher werden die Metriken bewusst kombiniert und nicht auf eine Kennzahl reduziert.

4 Validierung

4.1 Versuchsaufbau und Implementierung

Die Analyse wird als deterministische Pipeline umgesetzt. Die Verarbeitungsschritte sind:

1. Fensterung und Diskretisierung,
2. Normierung und Ableitungen,
3. Clustering,
4. Evaluation und Visualisierung.

Es existiert keine Trainingsphase im Sinne einer Modellanpassung über Epochen; stattdessen werden Parameter einmalig gesetzt und alle Daten anschließend konsistent mit Python-Skripten berechnet. Dies schafft eine reproduzierbare Grundlage für den Vergleich von Vorverarbeitungs- und Clustering-Varianten.

Die Parametrisierung folgt einem Minimalprinzip, um Interpretierbarkeit zu erhalten. Die Kraftschwelle F_{th} definiert die betrachtete Prozessphase. Die Bucketanzahl B legt die zeitliche Auflösung fest. Die Clusterzahl k wird auf drei gesetzt, um eine Zuordnung zu den drei Werkzeugzuständen zu ermöglichen. Es wird kein Best-Case-Tuning vorgenommen, sondern eine konservative Bewertung angestrebt.

4.2 Ergebnisse: Vergleich der Vorverarbeitungsvarianten

Tabelle 2 zeigt die Kennzahlen für ausgewählte Kombinationen aus Verfahren und Vorverarbeitung. In der Fallstudie erreichen mehrere Verfahren in der Variante „normierte Kurve“ die höchsten Übereinstimmungswerte mit der bekannten Zustandsstruktur. Für die Ableitungsvarianten sinken die Übereinstimmungswerte tendenziell; die zweite Ableitung ist in mehreren Fällen besonders instabil.

Die Resultate sind in zweierlei Hinsicht aufschlussreich. Erstens zeigen ähnliche Kennzahlen über verschiedene Verfahren, dass unterschiedliche Algorithmen im gleichen Merkmalsraum zu sehr ähnlichen Partitionen konvergieren. Das ist plausibel, wenn die Daten eine dominante, leicht trennbare Struktur enthalten, und deutet darauf hin, dass der Engpass weniger in der Wahl des konkreten Clustering-Algorithmus als in der Datenrepräsentation liegt.

Zweitens bleibt das Niveau der Übereinstimmung begrenzt. Selbst die besten Konfigurationen liefern keine nahezu perfekte Trennung der drei Zustände. Daraus folgt, dass mindestens einer der folgenden Faktoren wirkt:

1. die Zustandsklassen sind im abgeleiteten Kraftsignal nicht vollständig separierbar,
2. die Vorverarbeitung entfernt relevante Informationen,
3. der Merkmalsraum wird von dominierenden Störgrößen geprägt, die nicht verschleißbedingt sind, oder
4. die Zustandsdefinition scharf/mittel/kaputt bildet einen kontinuierlichen Übergang ab, der in Messsignalen nur graduell sichtbar ist.

Für DBSCAN zeigt sich in der vorliegenden Parametrisierung eine geringe Stabilität, vor allem auf Ableitungsdaten. Dies deutet darauf hin, dass die Dichtegeometrie im Merkmalsraum stark von Skalierung und Rauschen beeinflusst wird. Dichtebasierte Verfahren sind damit nicht ausgeschlossen, erfordern aber eine systematische Parameterauswahl, die Drift- und Rauschrobustheit explizit berücksichtigt.

4.3 Robustheits- und Sensitivitätsanalyse

Um die Aussagekraft der Ergebnisse unabhängig von Einzelentscheidungen der Vorverarbeitung zu beurteilen, wird eine

Tabelle 3 Sensitivitätsanalyse der Pipeline-Parameter (k-Means/norm, One-at-a-Time-Variation, jeweils Defaults für alle anderen Parameter).

Variierter Parameter	Wert	T [kN]	Δt [ms]	B	p	ARI	NMI	Jaccard	Silhouette
Kraftschwelle T	**3.000**	3	1	200	0,9	0,519	0,593	0,519	0,395
Kraftschwelle T	**5.000**	5	1	200	0,9	0,587	0,712	0,589	0,433
Kraftschwelle T	**7.000**	7	1	200	0,9	0,462	0,656	0,512	0,537
Nachlauf Δt	**0,5**	5	0,5	200	0,9	0,587	0,712	0,589	0,45
Nachlauf Δt	**1**	5	1	200	0,9	0,587	0,712	0,589	0,433
Nachlauf Δt	**2**	5	2	200	0,9	0,519	0,593	0,52	0,415
Bucketzahl B	**100**	5	1	100	0,9	0,519	0,593	0,52	0,423
Bucketzahl B	**200**	5	1	200	0,9	0,587	0,712	0,589	0,433
Bucketzahl B	**400**	5	1	400	0,9	0,587	0,712	0,589	0,436
Alignment p	**0,8**	5	1	200	0,8	0,587	0,712	0,589	0,417
Alignment p	**0,85**	5	1	200	0,85	0,519	0,593	0,52	0,375
Alignment p	**0,9**	5	1	200	0,9	0,587	0,712	0,589	0,433
Alignment p	**0,95**	5	1	200	0,95	0,519	0,593	0,52	0,437
Zusammenfassung						**Median 0,587**	**Median 0,712**	**Median 0,589**	**Median 0,433**
						Range [0,46; 0,59]	Range [0,59; 0,71]	Range [0,51; 0,59]	Range [0,38; 0,54]

Tabelle 4 Rausch-Robustheit (k-Means/norm, additives Gaußsches Rauschen).

Rauschpegel (% Signalspanne)	ARI	NMI	Jaccard	Silhouette
0 % (Referenz)	0,587	0,712	0,589	0,433
1%	0,519	0,593	0,52	0,296
2%	0,519	0,593	0,52	0,241
5%	0,462	0,656	0,512	0,231
10%	0,462	0,656	0,512	0,13

Tabelle 5 Synchronisationsfehler / Jitter (k-Means/norm, zufälliger Zeitversatz pro Hub).

Jitter (±Samples)	ARI	NMI	Jaccard	Silhouette
0 (Referenz)	0,587	0,712	0,589	0,433
±1	0,586	0,709	0,589	0,391
±2	0,584	0,708	0,588	0,362
±5	0,512	0,585	0,509	0,263

systematische Sensitivitätsanalyse der zentralen Pipeline-Parameter durchgeführt. Variiert werden (bei ansonsten identischer Pipeline): Kraftschwelle T (3000/5000/7000 N), Nachlauf Δt (0,5/ 1,0/2,0 ms), Bucketzahl B (100/200/400) sowie Alignment-Position p (0,80/0,85/0,90/0,95). Für jede Parameterkombination werden die Kennzahlen ARI/NMI/Jaccard/Silhouette berichtet (**Tabelle 3**), zusätzlich werden Robustheitszusammenfassungen (Median und Spannweite über den Grid) angegeben.

Zudem wird die Robustheit gegenüber Messartefakten explizit geprüft, indem kontrolliert (i) additives Rauschen (**Tabelle 4**) und (ii) Synchronisationsfehler (**Tabelle 5**) injiziert werden.

Für Rauschen wird auf den diskretisierten Kurven ein additives, normalverteiltes Rauschen mit relativer Amplitude 1–10 % der Signalspanne aufgebracht; für Synchronisationsfehler wird ein Zeitversatz von ±1 bis ±5 Samples beziehungsweise eine Verschiebung der Alignment-Position um ±2,5 (entspricht ±5 Samples bei B = 200) des normierten Fensters simuliert. Dadurch

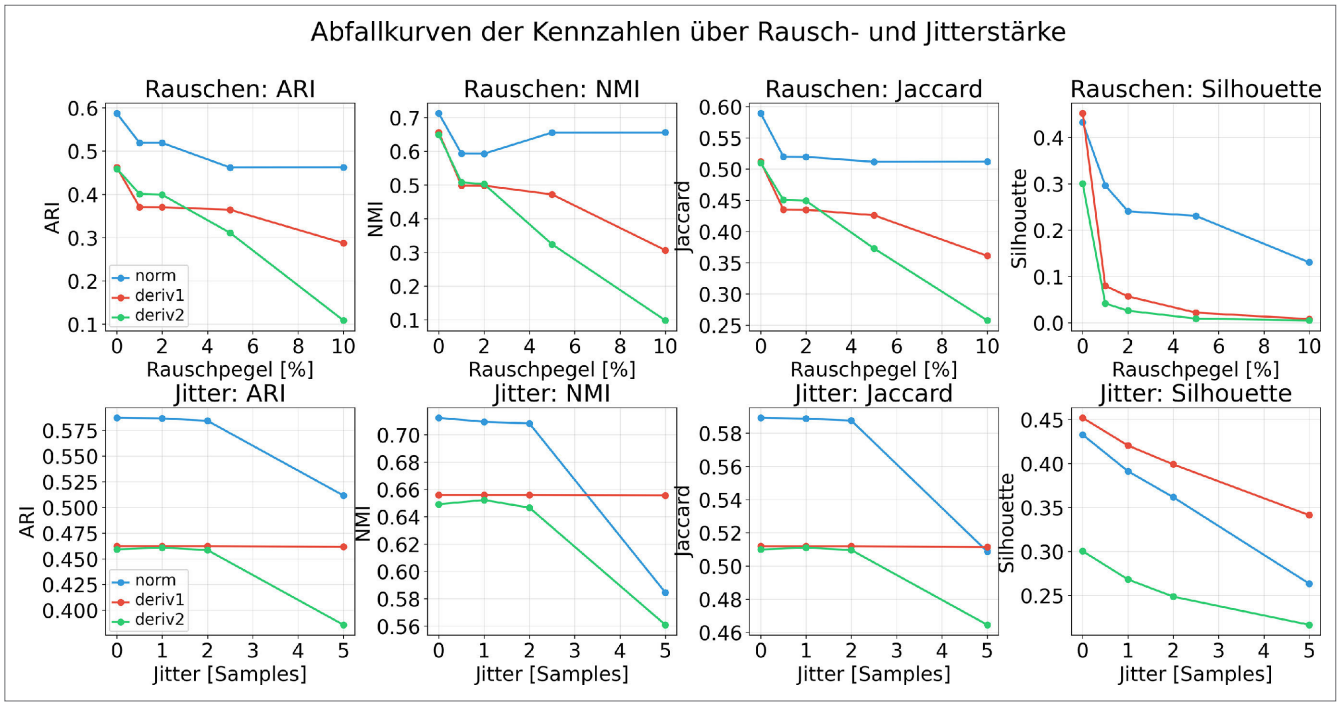

Bild 8 Abfallkurven der Kennzahlen über Rausch- und Jitterstärke. *Grafik: Eigene Darstellung*

Tabelle 6 Ausschlusskriterien fehlerhafter Messungen.

Kriterium	Anzahl Hübe
Geplant (3 Werkstoffe × 3 Zustände × 250)	2.250
In Rohdaten vorhanden	2.200
Nach Diskretisierung verbleibend	2.200
Ausgeschlossen gesamt	**0**
Davon: fehlende Rohdaten (vor Pipeline)	50
Davon: fehlender/mehrfacher Trigger	0
Davon: unplausible Fensterlänge	0
Davon: Sättigung/Clipping	0
Davon: fehlende Datenpunkte / NaN	0
Davon: abweichende Grundform	0

lässt sich quantifizieren, inwieweit die Ableitungsvarianten erwartungsgemäß empfindlicher reagieren als die normierten Kurven. Ergebnisse werden als Abfallkurven der Kennzahlen über Rauschbeziehungsweise Jitterstärke dargestellt (**Bild 8**).

Der Ausschluss fehlerhafter Messungen wird in **Tabelle 6** dokumentiert. Ausgeschlossen werden Hübe, die mindestens eines der folgenden Kriterien erfüllen: (1) fehlender oder mehrfacher Trigger/Peak-Index, (2) unplausible Fensterlänge (zum Beispiel <10 oder >300 Samples oberhalb der Schwelle), (3) Sättigung/Clipping der Kraftmesskette, (4) fehlende Datenpunkte beziehungsweise NaN-Werte, (5) offensichtlich abweichende Grundform (zum Beispiel doppelte Schneidphase).

4.4 Baselines und Einordnung des Mehrwerts

Um den Mehrwert der kurvenbasierten Repräsentation gegenüber einfachen Ansätzen einzuordnen, werden Baselines ohne Kurvenvektor betrachtet. Als Baseline-Features dienen

a. maximale Kraft F_{max},
b. Integral der Kraft über das Fenster (Energieproxy),
c. Fensterlänge beziehungsweise Dauer oberhalb der Schwelle,
d. Steigungskennwerte (zum Beispiel maximale erste Differenz).

Auf diesen Low-Dimensional-Features werden k-Means und Ward als Referenz angewendet. Die resultierenden Kennzahlen (ARI/NMI/Jaccard) werden den kurvenbasierten Varianten gegenübergestellt (**Tabelle 7**). Erwartungsgemäß zeigt sich, ob der Gewinn aus der Forminformation tatsächlich über triviale Kraftniveau- und Dauerinformationen hinausgeht.

Tabelle 7 Gegenüberstellung ARI, NMI und Jaccard bezogen auf den jeweiligen Ansatz.

Ansatz	ARI	NMI	Jaccard
Baseline k-Means	0,4	0,45	0,44
Baseline Ward	0,4	0,44	0,43
Kurvenbasiert k-Means	0,59	0,71	0,59

Tabelle 8 LOMO Ergebnisse

	Zustand			Werkstoff	
Hold-out	ARI	NMI	Jaccard	ARI	NMI
DP1000	0,6	0,75	0,62	0	0
DP600	0,57	0,72	0,6	0	0
DP800	0,57	0,73	0,6	0	0

Tabelle 9 Kontingenztabellen Cluster x Zustand und Cluster x Werkstoff.

	Cluster x Zustand				Cluster x Werkstoff			
Cluster	Kaputt	Mittel	Scharf	Summe	DP1000	DP600	DP800	Summe
0	0	500	700	1.200	450	250	500	1.200
1	750	0	0	750	250	250	250	750
2	0	250	0	250	0	250	0	250
Summe	***750***	***750***	***700***	***2.200***	***700***	***750***	***750***	***2.200***

4.5 Materialübergreifende Validierung (Leave-One-Material-Out)

Der zentrale Anspruch der Arbeit ist eine möglichst werkstoffunabhängige Erkennung von Werkzeugzuständen. Um diese Generalisierungsfähigkeit zu prüfen, wird ein Leave-One-Material-Out (LOMO)-Protokoll verwendet. Zwei Werkstoffe dienen zur Festlegung der Vorverarbeitung und Pipeline-Parameter, der dritte Werkstoff wird ausschließlich zur Validierung genutzt. Konkret werden drei Läufe durchgeführt:

1. Parameter auf DP600+DP800, Validierung auf DP1000
2. DP600+DP1000 → DP800
3. DP800+DP1000 → DP600

Wichtig ist, dass die Parameterwahl im LOMO-Protokoll nicht anhand der Zustandslabels optimiert wird. Die Parameter werden einmalig anhand plausibler, reproduzierbarer Heuristiken beziehungsweise anhand des Parametrisierungs-Sets festgelegt (wie etwa Kraftschwelle, Nachlauf, Bucketzahl, Alignment-Position p). Die Zustandslabels werden erst im Nachgang zur ex-post Bewertung der Generalisierungsfähigkeit herangezogen.

Für jeden LOMO-Lauf werden

a. ARI/NMI/Jaccard bezogen auf den Werkzeugzustand sowie
b. ARI/NMI bezogen auf den Werkstoff berichtet.

Letzteres dient als Material-Leakage-Check: Hohe Übereinstimmung „Cluster <-> Werkstoff“ bei gleichzeitig niedriger Übereinstimmung „Cluster <-> Zustand“ würde darauf hinweisen, dass die gefundenen Cluster primär Material- beziehungsweise Prozesseffekte trennen. Die Ergebnisse werden in **Tabelle 8** zusammengefasst.

4.6 Clusterinterpretation und fachliche Bedeutung

Die Metriken aus Abschnitt 4.2 belegen, dass im Merkmalsraum eine gewisse Struktur vorhanden ist. Für eine inhaltliche Interpretation ist aber eine Clustercharakterisierung erforderlich. Deshalb wird für jede Konfiguration zusätzlich ausgewertet:

1. die Zusammensetzung der Cluster nach Werkzeugzustand (scharf/mittel/kaputt)
2. die Zusammensetzung nach Werkstoff (DP600/DP800/DP1000).

Beides wird als Kontingenzmatrix „Cluster x Zustand“ beziehungsweise „Cluster x Werkstoff“ berichtet (**Tabelle 9**).

Zusätzlich werden pro Cluster repräsentative Verläufe visualisiert (Median und Interquartilsabstand der normierten Kurven beziehungsweise Ableitungen). Um Werkstoff- und Zustandsanteile innerhalb eines Clusters getrennt beurteilen zu können, werden diese Repräsentationsplots innerhalb des Clusters nach Werkstoff überlagert dargestellt (**Bild 9**). Somit lässt sich prüfen, ob ein Cluster eine konsistente Form über Materialien hinweg abbildet (indikativ für Zustandsinvarianz) oder ob innerhalb eines Clusters systematisch materialabhängige Substrukturen auftreten.

Da Clustering-Ergebnisse von Initialisierung und Stichprobenzusammensetzung abhängen können, wird die Clusterstabilität über Wiederholungen (20 Seeds beziehungsweise Bootstrap-Samples) quantifiziert. Als Stabilitätskennzahl wird die Übereinstimmung der Clusterzuordnungen zwischen Wiederholungen (zum Beispiel ARI über Clusterlabels) berichtet. Instabile Cluster würden die fachliche Interpretation einschränken und darauf hin-

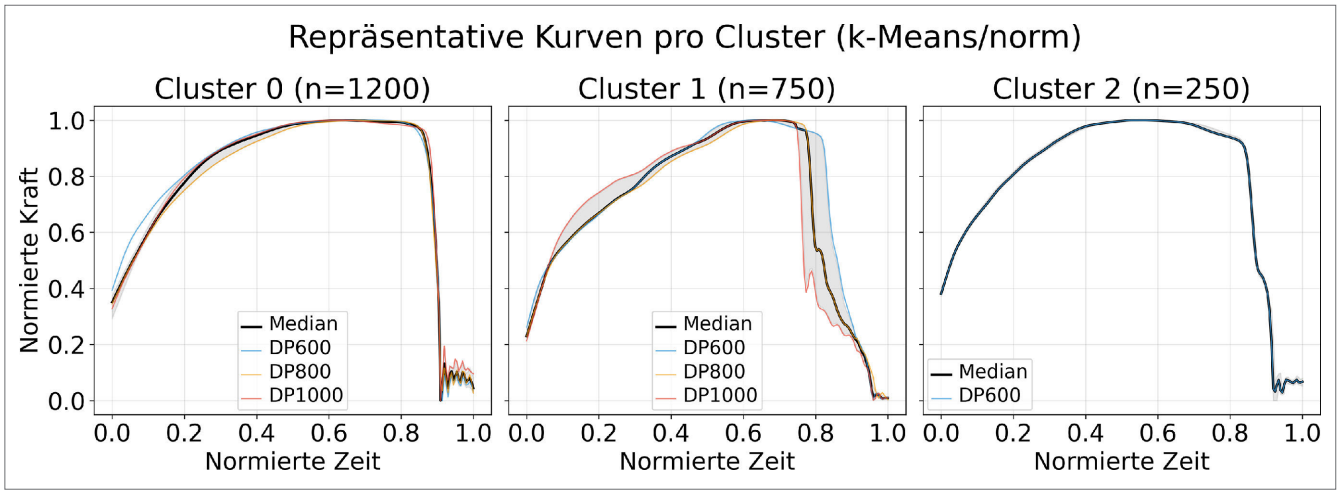

Bild 9 Repräsentative Kurven pro Cluster. *Grafik: Eigene Darstellung*

deuten, dass die beobachtete Struktur nicht robust genug für eine Zustandsdiagnose ist.

4.7 Drift- und Zeitachsenanalyse entlang des Verschleißverlaufs

Verschleiß ist typischerweise kein sprunghafter Zustandswechsel, sondern ein gradueller Prozess. Um zu prüfen, ob die Clusterzuordnung diese Drift abbildet oder lediglich verrauschte Schwankungen widerspiegelt, wird die Clusterzuordnung in Prozessreihenfolge ausgewertet. Dazu wird je Werkstoff und Versuchslauf die Folge der Hübe in zeitlicher Reihenfolge betrachtet und die Clusterlabels über dem Hubindex aufgetragen (**Bild 10**).

Als quantitative Drift-Indikatoren werden

1. die Transitionsrate (Anteil der Hübe, bei denen sich das Clusterlabel gegenüber dem vorherigen Hub ändert) sowie
2. die Entwicklung der Distanz zum zugewiesenen Clusterzentrum (als kontinuierlicher „Verschleißscore“) berechnet.

Eine niedrige Transitionsrate bei gleichzeitig systematischem Trend im Score wäre konsistent mit Drift. Ein häufiges Springen ohne Trend spricht eher für Sensitivität gegenüber Rauschen beziehungsweise Alignment-Fehlern.

Die Frage, ob ein Werkzeugzustand unabhängig vom Werkstoff erkennbar bleibt, wird in dieser Arbeit über das LOMO-Protokoll (Abschnitt 4.5) und den Material-Leakage-Check (Abschnitte 3.6 und 4.5) operationalisiert. Die Fallstudie umfasst drei DP-Stähle mit unterschiedlichen Festigkeitsniveaus. Kraftamplituden und Kurvenlängen hängen somit nicht nur vom Verschleiß, sondern stark vom Werkstoff ab. Entsprechend zeigen die Baseline-Ergebnisse (Abschnitt 4.4), inwieweit triviale Kraftniveau- und Dauerinformationen bereits eine Clusterstruktur erzeugen können, und die Robustheitsanalyse (Abschnitt 4.3) quantifiziert die Empfindlichkeit gegenüber Rauschen und Synchronisationsfehlern.

Aus Anwendersicht sind drei naheliegende, aber unzureichende Heuristiken zu unterscheiden:

1. Mehr Verschleiß → höhere Maximal-Kraft: plausibel, aber durch Zugfestigkeitsschwankungen überlagert.
2. Mehr Verschleiß → längere Schneidphase: ebenfalls plausibel, aber durch thermische Effekte und Einlaufzustände verzerrbar.
3. Mehr Verschleiß → markante Gradientenänderungen: möglich, aber Ableitungen sind besonders rauschempfindlich.

Eine praxistaugliche Diagnose muss daher mehrere Signalaspekte kombinieren und Störgrößen kontrollieren.

Der in dieser Fallstudie verfolgte Ansatz unterscheidet sich in Ziel und Informationspfad von feature-selection-basierten Arbeiten. Während diese typischerweise auf interpretierbare Kennwerte und deren Korrelation mit Qualitäts- oder Defektmerkmalen zielen, untersucht dieser Beitrag eine rein ähnlichkeitsbasierte Gruppierung der Signale. Dies reduziert den Engineering-Aufwand, verschiebt aber die Last auf die Robustheit der Repräsentation. Der Merkmalsraum muss so gestaltet sein, dass die gewünschten Zustandsunterschiede dominieren können.

Damit wird eine Grenze sichtbar. Wenn Werkstoff, Prozessführung und Temperatur den Merkmalsraum stärker strukturieren als der Werkzeugzustand, ist ein Clustering für Verschleißdiagnosen prinzipiell unterbestimmt. In diesem Fall sind zusätzliche Informationsquellen (weitere Sensorik, Prozessmodelle oder kontrollierte Validierungsdaten) erforderlich, um die Ursachenebenen zu entflechten.

5 Zusammenfassung zur Generalisierungsfähigkeit des Verfahrens

Die Generalisierung wird in dieser Arbeit als materialübergreifende Stabilität der Clusterzuordnung verstanden. Im LOMO-Protokoll (Abschnitt 4.5) werden die Pipeline-Parameter ausschließlich auf dem Parametrisierungs-Subset festgelegt. Anschließend wird auf dem Hold-out-Material evaluiert. Berichtet werden sowohl zustandsbezogene Kennzahlen (ARI/NMI/Jaccard gegen scharf/mittel/kaputt) als auch werkstoffbezogene Kennzahlen (ARI/NMI gegen DP600/800/1000), um die Trennrichtung der Cluster zu prüfen.

Für eine operative Nutzung wird eine Zuordnung „Cluster → Zustand“ auf Basis des Parametrisierungs-Sets festgelegt (zum Beispiel per Mehrheitszuordnung beziehungsweise Permutation, welche die Übereinstimmung maximiert). Diese Zuordnung wird danach unverändert auf das Hold-out-Material übertragen. Die dabei erzielten Kennzahlen (zum Beispiel ARI = 0,57 – 0,60, NMI = 0,72 – 0,75) sind ein direkter Indikator, ob die Cluster tatsächlich zustandsgetrieben sind oder ob sich die Struktur beim Materialwechsel grundlegend ändert.

Die LOMO-Ergebnisse zeigen, in welchen Konfigurationen die Zustandsstruktur materialübergreifend stabil bleibt und wo Werk-

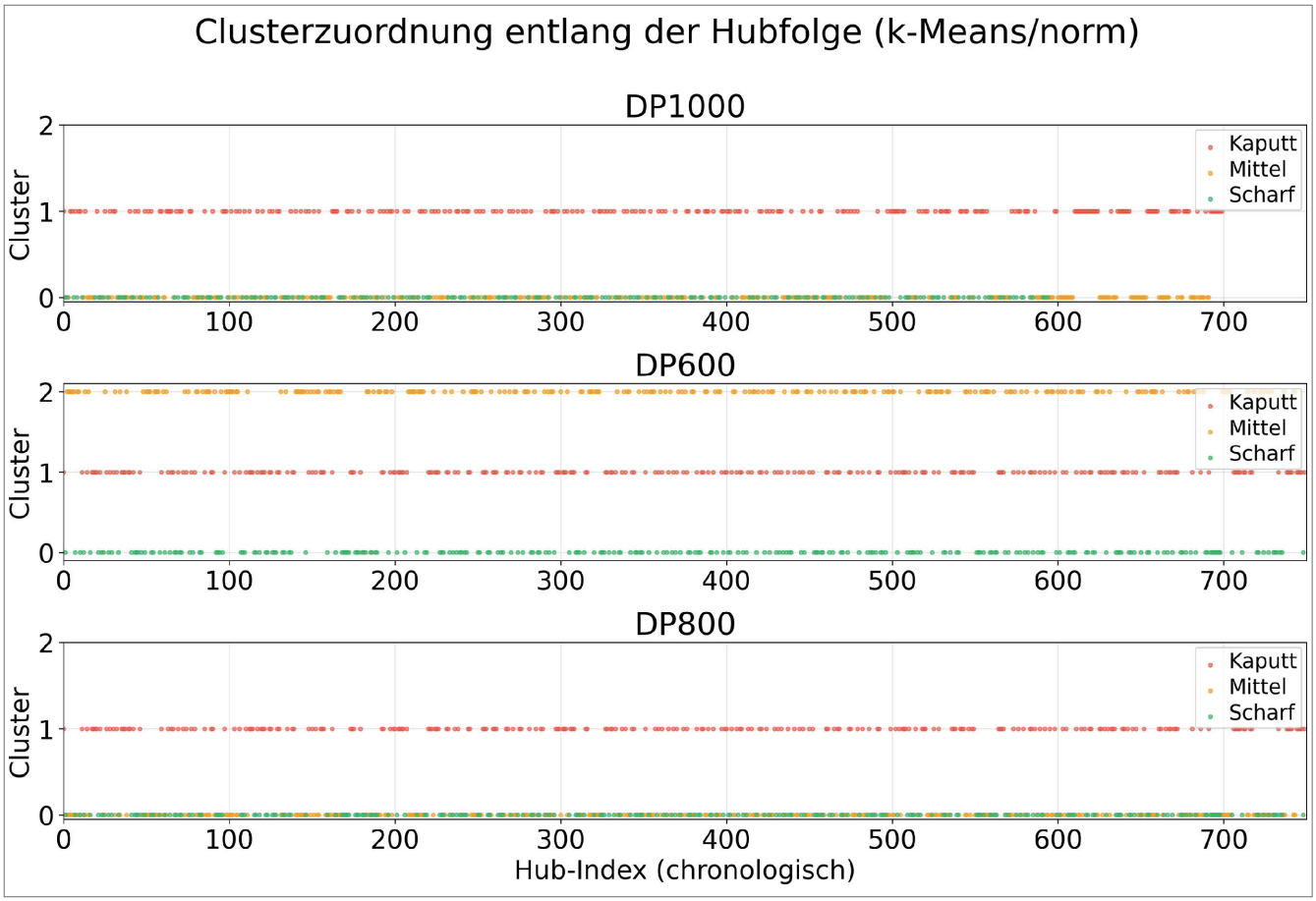

Bild 10 Clusterzuordnung entlang der Hubfolge (k-Means/norm). *Grafik: Eigene Darstellung*

stoffeffekte dominieren. In Fällen, in denen die werkstoffbezogenen Kennzahlen die zustandsbezogenen untertreffen (zum Beispiel $ARI_{Werkstoff} = 0{,}00 < ARI_{Zustand} = 0{,}58$), ist die Interpretation „Werkzeugzustand“ belastbar. Umgekehrt stützen stabile, zustandsbezogene Kennzahlen über alle drei LOMO-Läufe (zum Beispiel ARI in $[0{,}57,\ 0{,}00,\ 0{,}60]$) die Aussage, dass die Repräsentation Invarianten gegenüber dem Werkstoff gelernt beziehungsweise durch Normierung erzeugt hat.

Die Fallstudie zeigt, dass eine reine Clusterung diskretisierter Kraftkurven ohne domänenspezifische Features als alleinige, vollautomatische Diagnose nur dann tragfähig ist, wenn

1. die materialübergreifende Validierung (Abschnitt 4.3) stabil ausfällt,
2. die Cluster fachlich interpretierbar sind (Abschnitt 4.4) und
3. Robustheit gegenüber Rauschen und Synchronisationsfehlern gegeben ist.

Gleichzeitig wird belegt, dass in den Daten strukturierende Informationen vorhanden sind und mehrere Verfahren konsistent ähnliche Partitionen liefern. Daraus ergeben sich nächste Schritte entlang der Kette Signal → Invarianz → Regime → Entscheidung.

Erstens sollte Drift explizit adressiert werden, etwa durch Verfahren, die zeitliche Entwicklungen modellieren oder inkrementell arbeiten, sowie dichte- oder manifoldbasierte Ansätze mit stabiler Parametrisierung. Ergänzend ist ein Konzept zur Schwellendefinition über die Zeit erforderlich, da eine fixe Kraftschwelle zwar die Synchronisation erleichtert, bei geänderter Prozessführung oder Werkzeugtemperatur jedoch die Fensterung systematisch verschieben kann.

Zweitens ist eine robustere Invarianzbildung gegenüber Werkstoff und Prozesszustand notwendig. Eine Möglichkeit ist die Kombination aus normierter Kurvenlänge und formbasierten Deskriptoren, die nicht nur Ableitungen, sondern auch lokal integrierte Maßzahlen (wie Energieanteile in Phasen) berücksichtigen. Eine weitere Option ist die Nutzung von Distanzmaßen, die phasenverschobene Kurven tolerieren, ohne die Physik vollständig zu ignorieren.

Drittens muss die Validierungsstrategie an industrielle Anforderungen angepasst werden. Statt ausschließlich ex-post-Labels sollten online-fähige Gütefunktionen verwendet werden, ergänzt um seltene, gezielte Ground-Truth-Erhebungen (zum Beispiel mikroskopische Kantenvermessung in Stichproben). Zusätzlich ist eine Risikobetrachtung notwendig, die Fehlalarme und verpasste Verschleißereignisse in Kosten übersetzt, damit datengetriebene Verfahren in Wartungsentscheidungen integriert werden können.

Langfristig ist eine Übertragbarkeit auf weitere Werkstoffklassen (niederfeste Stähle, Aluminium, Kupfer) sowie auf andere Fertigungsprozesse mit vergleichbaren Signaturen (zum Beispiel Umform- oder Fügeverfahren) relevant. Der Beitrag für die Fertigungstechnik besteht darin, die Lücke zwischen datengetriebenen Methoden und prozessphysikalischer Interpretation systematisch zu schließen, sodass Assistenzsysteme nicht nur Abweichungen erkennen, sondern auch Ursachen plausibilisieren und Maßnahmen ableiten können. Für eine nächste Iteration ergeben sich konkret:

a. Generalisierungsprüfung über Werkstoffe (materialbasierte Parametrisierung und separate Validierung),
b. Erweiterung der Vorverarbeitungsvarianten, insbesondere Nutzung der diskretisierten Rohkurve ohne Min-Max-Normierung als Referenz sowie Ableitungen nach Längen- und Höhen-Normierung, um Form- und Gradienteninformation stärker zu isolieren, und
c. Einbettung in ein Online-Szenario, in dem Clusterstabilität über die Zeit bewertet und Drift systematisch erfasst wird.

Parallel dazu ist eine betriebswirtschaftliche Einbettung notwendig. Ein Wartungsalgorithmus wird nicht anhand einer Metrik wie ARI oder Silhouette eingeführt, sondern anhand der Reduktion von Stillstand, Ausschuss und ungeplanten Werkzeugereignissen.

Künftige Validierungen sollten daher nicht nur Signaltrennbarkeit prüfen, sondern auch eine Kostenfunktion definieren, die Fehlalarme (unnötige Wartung) und Verpasser (zu späte Wartung) gegeneinander abwägt. Erst wenn diese Übersetzung gelingt, kann datengetriebene Clusterung in ein belastbares Entscheidungsunterstützungssystem überführt werden.

Schließlich ist eine Standardisierung der Datenerfassung nötig. Ohne konsistente Abtastraten, reproduzierbare Triggerpunkte und dokumentierte Prozessrandbedingungen werden Unterschiede im Signal leicht zu Datenartefakten statt zu Zustandsindikatoren. Eine robuste Datenbasis ist damit keine Zusatzoption, sondern Voraussetzung für jede übertragbare Zustandsüberwachung.

Literatur

[1] Hoffmann, H.; Neugebauer, R.; Spur, G. (Hrsg.): Handbuch Umformen. München: Hanser 2012
[2] Kubik, C.; Hohmann, J.; Groche, P.: Exploitation of force displacement curves in blanking-feature engineering beyond defect detection. The International Journal of Advanced Manufacturing Technology 113 (2021) 1–2, pp. 261–278
[3] Makich, H.; Carpentier, L.; Monteil, G. et al.: Metrology of the burr amount – correlation with blanking operation parameters (blanked material – wear of the punch). International Journal of Material Forming 1 (2008) S1, pp. 1243–1246
[4] Hambli, R.: Prediction of burr height formation in blanking processes using neural network. International Journal of Mechanical Sciences 44 (2002) 10, pp. 2089–2102
[5] Görz, M.; Schenek, A.; Liewald, M. et al.: Einsatz von ML beim Scheren im offenen Schnitt / Use of ML in shearing with an open cutting line. wt Werkstattstechnik online 113 (2023) 10, S. 384–388. Düsseldorf: VDI Fachmedien, doi.org/10.37544/1436–4980–2023–10–6
[6] Schenek, A.; Görz, M.; Liewald, M. et al.: Data-Driven Derivation of Sheet Metal Properties Gained from Punching Forces Using an Artificial Neural Network. Key Engineering Materials 926 (2022), pp. 2174–2182
[7] Schenek, A.; Görz, M.; Riedmüller, K. R. et al.: Application of a neural network for predicting cutting surface quality of punching processes based on tooling parameters. IOP Conference Series: Materials Science and Engineering 1284 (2023) 1, #12014
[8] Kubik, C.; Knauer, S. M.; Groche, P.: 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. Journal of Intelligent Manufacturing 33 (2022) 1, pp. 259–282
[9] Breitling, J.; Pfeiffer, B.; Altan, T. et al.: Process control in blanking. Journal of Materials Processing Technology 71 (1997) 1, pp. 187–192
[10] Groche, P.; Hohmann, J.; Übelacker, D.: Overview and comparison of different sensor positions and measuring methods for the process force measurement in stamping operations. Measurement 135 (2019), pp. 122–130
[11] Görz, M.; Schenek, A.; Liewald, M. et al.: Evaluation of Feature Engineering Methods for the Prediction of Sheet Metal Properties from Punching Force Curves by an Artificial Neural Network. In: Zhang, M.; Peng, Z.; Li, B. et al. (Hrsg.): Characterization of Minerals, Metals, and Materials 2023. Cham: Springer Nature Switzerland 2023, pp. 75–85
[12] Hohmann, J. F.: Ermittlung und Bewertung von Prozesskenngrößen zur online Bauteil- und Prozessüberwachung bei Scherschneidprozessen. Dissertation, Technische Universität Darmstadt, 2020
[13] Schenek, A.; Görz, M.; Liewald, M. et al.: Prediction of Cutting Surface Parameters in Punching Processes Aided by Machine Learning. TMS 2023 152nd Annual Meeting & Exhibition Supplemental Proceedings. Cham: Springer Nature Switzerland 2023, pp. 607–619
[14] Hoppe, F.; Hohmann, J.; Knoll, M. et al.: Feature-based Supervision of Shear Cutting Processes on the Basis of Force Measurements: Evaluation of Feature Engineering and Feature Extraction. Procedia Manufacturing 34 (2019), pp. 847–856
[15] Kubik, C.; Molitor, D. A.; Rojahn, M. et al.: Towards a real-time tool state detection in sheet metal forming processes validated by wear classification during blanking. IOP Conference Series: Materials Science and Engineering 1238 (2022) 1, #12067
[16] Ramasubramanian, K.; Singh, A.: Machine Learning Using R. Berkeley, CA: Apress 2019
[17] Richard Bellman: Dynamic Programming. Science 153 (1957), pp. 34–37
[18] Venkat, N.: The Curse of Dimensionality: Inside Out. Carnegie Mellon University, 2018, doi.org/10.13140/RG.2.2.29631.36006
[19] Guyon, I.; Elisseeff, A.: An Introduction to Feature Extraction. In: Guyon, I.; Nikravesh, M.; Gunn, S. et al. (Hrsg.): Feature Extraction: Foundations and Applications. Heidelberg: Springer 2006, pp. 1–25
[20] Jia, W.; Sun, M.; Lian, J. et al.: Feature dimensionality reduction: a review. Complex & Intelligent Systems 8 (2022) 3, pp. 2663–2693
[21] Pudil, P.; Novovičová, J.; Kittler, J.: Floating search methods in feature selection. Pattern Recognition Letters 15 (1994) 11, pp. 1119–1125
[22] Dash, M.; Liu, H.; Yao, J.: Dimensionality reduction of unsupervised data. Proceedings Ninth IEEE International Conference on Tools with Artificial Intelligence, Newport Beach, CA, USA, 1997, pp. 532–539, doi.org/10.1109/TAI.1997.632300
[23] Emmert-Streib, F.; Moutari, S.; Dehmer, M.: Elements of Data Science, Machine Learning, and Artificial Intelligence Using R. Cham: Springer International Publishing 2023
[24] Mende, H.; Frye, M.; Vogel, P.-A. et al.: On the importance of domain expertise in feature engineering for predictive product quality in production. Procedia CIRP 118 (2023), pp. 1096–1101
[25] Nargesian, F.; Samulowitz, H.; Khurana, U. et al.: Learning Feature Engineering for Classification. Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, 2017, pp. 2529–2535, doi.org/10.24963/ijcai.2017/352
[26] Verdonck, T.; Baesens, B.; Óskarsdóttir, M. et al.: Special issue on feature engineering editorial. Machine Learning 113 (2021) 7, pp. 3917–3928
[27] Saarela, M.; Jauhiainen, S.: Comparison of feature importance measures as explanations for classification models. SN Applied Sciences 3 (2021) #272, doi.org/10.1007/s42452–021–04148–9
[28] Hooker, S.; Erhan, D.; Kindermans, P.-J. et al.: A benchmark for interpretability methods in deep neural networks. In: Proceedings of the 33rd International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc 2019, pp. 1–12
[29] Rengasamy, D.; Mase, J. M.; Kumar, A. et al.: Feature importance in machine learning models: A fuzzy information fusion approach. Neurocomputing 511 (2022), pp. 163–174
[30] Michelucci, U.: Fundamental Mathematical Concepts for Machine Learning in Science. Cham: Springer International Publishing 2024
[31] Suhaidi, M.; Abdul Kadir, R.; Tiun, S.: A Review of Feature Extraction Methods on Machine Learning. Journal of Information System and Technology Management 6 (2021) 22, pp. 51–59
[32] Chowdhary, K. R.: Fundamentals of Artificial Intelligence. New Delhi: Springer India 2020

Dipl.-Wirt.-Inf. Hendrik von Linde
hendrik.von-linde@iao.fraunhofer.de

Fraunhofer-Institut für Arbeitswirtschaft
und Organisation IAO
Nobelstr. 12, 70569 Stuttgart
www.iao.fraunhofer.de

Matthias Lück, M.Sc.

Universität Stuttgart, Institut für Arbeitswissenschaft
und Technologiemanagement (IAT)

Dr.-Ing. Adrian Schenek

Universität Stuttgart
Institut für Umformtechnik (IFU)
Holzgartenstr. 17, 70174 Stuttgart
www.ifu.uni-stuttgart.de
Nobelstr. 12, 70569 Stuttgart
www.iat.uni-stuttgart.de

Prof. Dr.-Ing. Oliver Riedel

Universität Stuttgart, Institut für Steuerungstechnik
der Werkzeugmaschinen und Fertigungseinrichtungen (ISW)
Seidenstr. 36, Stuttgart 70569
www.isw.uni-stuttgart.de