

## Steigerung der Vorhersagegüte modellbasierter Prozesskraftschätzungen

# Vorhersage von Restunsicherheiten mithilfe von KI

C. Brecher, M. Fey, M. Loba, J. Vos

**ZUSAMMENFASSUNG** Moderne Werkzeugmaschinen haben zahlreiche Sensoren, die sich zur Prozessüberwachung nutzen lassen. Die dahinterliegenden Modelle besitzen jedoch Restunsicherheiten, beispielsweise infolge von Reibung, die zu fehlerhaften Ergebnissen führen können. Datengetriebene Ansätze mithilfe neuronaler Netze bieten die Möglichkeit, diese zu verringern. In diesem Beitrag wird untersucht, inwiefern sich neuronale Netze zur Steigerung der Vorhersagegüte modellbasierter Prozesskraftschätzungen auf Basis steuerungsinnter Daten eignen. Untersucht wird die Eignung eines Transformer-Encoder-Modells. Es zeigt sich, dass sich hiermit Modellfehler um 66 % reduzieren lassen.

## STICHWÖRTER

Künstliche Intelligenz (KI), Werkzeugmaschinen, Automatisierung

## Predicting residual uncertainties using AI: Increasing the predictive quality of control data-based models for process force estimation

**ABSTRACT** Modern machine tools are equipped with numerous sensors that can be used for process monitoring. However, the underlying models exhibit residual uncertainties, for example due to friction, which can lead to inaccurate results. Data-driven approaches using neural networks offer the potential to reduce these uncertainties. This paper investigates the suitability of neural networks to improve the prediction accuracy of model-based process force estimations using internal machine data in milling. The suitability of a transformer-encoder model is investigated. It is shown that this can reduce model errors by 66 %.

## 1 Einleitung

Die zunehmende Verfügbarkeit von Daten aus Werkzeugmaschinen sowie steigende Rechenleistungen, ermöglichen es Produktionsanlagen effizienter zu betreiben. Basierend auf steuerungsinnter Daten lassen sich bereits heute Qualität [1, 2], Produktivität [3] und Verschleiß von Werkzeugen [4] sowie Maschinenkomponenten [5] erfassen. Vor allem im Hinblick auf die Qualität als auch den Verschleiß haben die entstehenden Lasten infolge von Prozesskräften einen entscheidenden Einfluss. Ein Vorteil der Verwendung steuerungsinnter Daten zur Vorhersage von Prozesskräften ist, dass sie helfen den aktuellen Zustand abzubilden. Die Vorhersage geschieht sowohl aufgrund einer physikalischen Beschreibung des Gesamtsystems (modellbasiert) als auch datengetrieben mithilfe von künstlicher Intelligenz (KI) in Form von Machine-Learning-Algorithmen (ML-Algorithmen) und künstlichen neuronalen Netzen (KNN). Vorteile einer modellbasierten Betrachtung liegen in einem Gesamtverständnis des Systems und einer hohen Recheneffizienz. Diese haben jedoch den Nachteil, dass sie eine gewisse Restunsicherheit besitzen, da sie, teils stochastische, Phänomene nicht abbilden können. In diesem Beitrag wird ein hybrider Ansatz vorgestellt, der zunächst modellbasiert die Prozesskraft vorhersagt (vgl. Kapitel 2) und mithilfe eines Black-Box-Anteils den Fehler zwischen realer und prognostizierter Kraft kompensiert (vgl. Kapitel 3). Der Ansatz wird anhand mehrerer Bauteile validiert.

## 2 Modellbasierte Prozesskraftvorhersage

Zur modellbasierten Prozesskraftvorhersage anhand von Steuerungsdaten ist ein Gesamtverständnis des Systems notwendig. Es existieren Ansätze, die sowohl anhand der resultierenden Achs- und Spindelströme [6–9] sowie der Differenz der Wegmesssysteme [10], die entstehenden Kräfte bestimmen. Die Modelleansätze bestimmen zunächst die Reib- und Trägheitsanteile sowie positionsabhängige Abweichungen (beispielsweise infolge des Spindelsteigungsfehlers) und subtrahieren diesen vom gemessenen Ist-Wert. Während die Vorhersage auf Basis der Achsströme sowie der Differenz der Wegmesssysteme die resultierende Gesamtkraft in der jeweiligen Achsrichtung ermittelt [6, 7, 10], lässt sich anhand des Spindelstroms nur der Schnittkraftanteil bestimmen [8, 9]. Park *et al.* nutzen in [9] anschließend einen konstanten Faktor um die Radialkräfte zu bestimmen. Ausgehend davon, dass Vorschubachsen nur Kräfte in Vorschubrichtung aufnehmen, ergibt sich vereinfacht für eine Achse der resultierende Strom zu:

$$I_x = \frac{\partial I}{\partial M_{M,x}} M_{M,x} = \frac{\partial I}{\partial M_{M,x}} \left( \underbrace{\frac{M_{F_x}}{F_x K_{F/M}}}_{\text{Prozesskraft-anteil}} + \underbrace{J_{\text{ges},M,x} \dot{\omega}_x}_{\text{Trägheitsanteil}} + \underbrace{M_{\text{fric},x}(\omega_x)}_{\text{Reibanteil}} + \underbrace{M_x(\varphi_x)}_{\text{pos.-abhängiger Fehler}} \right) \quad (1)$$

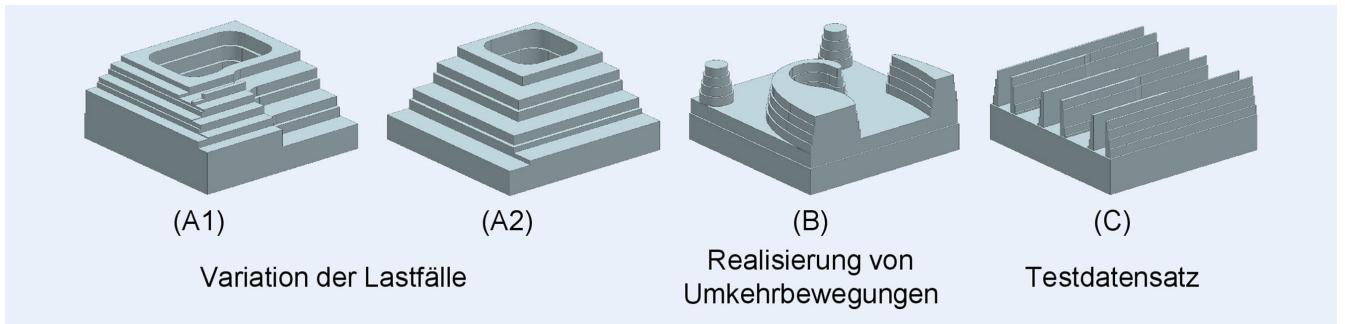


Bild 1. Vergleich modellbasierter Verfahren zur Bestimmung von Prozesskräften. Grafik: RWTH Aachen

sowie eine existierende Abweichung zwischen den Wegmesssystemen:

$$\Delta x_{2,1} = x_2 - \varphi_x \frac{r_2 h_{KGTx}}{r_1 2\pi} = \frac{\text{Achsstifigkeit}}{k_x(x_2)} \cdot \left( \underbrace{F_x^A}_{\text{Prozesskraftanteil}} + \underbrace{m_{ges} \ddot{x}_2}_{\text{Trägheitsanteil}} + \underbrace{F_{fric}(\dot{x}_2)}_{\text{Reibanteil}} \right) + \frac{\text{pos.-abhängiger Fehler}}{h_{err}(x_2)} \quad (2)$$

Hierbei gibt  $F_x$  den Anteil der Prozesskraft wieder,  $x_2$  den Wert des linearen Messsystems und  $x_1$  den des Drehgebers, der sich aus der Winkellage  $\varphi_x$ , der Getriebeübersetzung ( $r_2/r_1$ ) und der Spindelsteigung des Kugelgewindetriebes (KGT) ergibt. Mithilfe von Luftschnitten lassen sich die systematischen Anteile, welche nicht von der Prozesskraft abhängen, bestimmen. Stochastische Abweichungen, die vor allem bei Bewegungen mit niedrigen Vorschubgeschwindigkeiten auftreten, führen jedoch zu Abweichungen zwischen Soll- und Ist-Werten. Zusätzlich muss bei stillstehenden Achsen zunächst die Haftreibung überwunden werden, so dass eine Veränderung in den Strömen oder der Differenz der Wegmesssystem zu erkennen ist. Einen Vorteil bietet die Betrachtung der Hauptspindel, da diese während der Bearbeitung mit einer definierten Drehzahl rotiert. Nachteilig ist jedoch, dass sich nur das Schnittmoment berechnen lässt. Eine Abschätzung des Betrags und der Richtung der Schnittkraft ist mithilfe einer zusätzlichen Abtragsimulation möglich [8].

$$I_{SP} = \frac{\partial I}{\partial M_{SP}} M_{SP} - \frac{\partial I}{\partial M_{SP}} \left( \underbrace{\sum_i M_i}_{\sum_i F_{ci,r,i}^SP} + \underbrace{J_{ges,SP} \ddot{\theta}}_{\text{Trägheitsanteil}} + \underbrace{M_{fric,n}(n)}_{\text{Reibanteil}} + \underbrace{M_{SP}(\varphi_{SP})}_{\text{pos.-abhängiger Fehler}} \right) \quad (3)$$

Bild 1 zeigt einen Vergleich der Vorhersage der einzelnen Modelle für zwei Abschnitte eines Prozesses. Die Abschätzung der Prozesskraft erfolgt mit der in [1] vorgestellten dixel-basierten Abtragsimulation. Im oberen Teil des Bildes sind die Kräfte bei einer Linearbewegung dargestellt. Es ist erkennbar, dass das Modell auf Grundlage der Achsstrome  $F^1$  eine ähnliche Schwingungsamplitude wie die Kraftmessplattform (KMP) aufweist. Anhand einer Tiefpassfilterung (rechter Bildteil) wird sichtbar, dass das Modell auf Basis des Hauptspindelstroms  $F^{SP}$  den Trend am besten abbildet. Gleicher gilt für die Kreisbewegungen im unteren Bildteil. Hier sorgen im Bereich der Achsumkehrbewegungen Reibungseffekte zu einer Überbewertung der Kraft, die mit dem Modell berechnet wird, das auf der Differenz der Wegmesssysteme  $F^A$  beruht. Dieser Effekt ist beim Achsstrommodell weniger ausgeprägt. Eine Abbildung des Trends ist hier im Vergleich zu  $F^{SP}$  erkennbar schlechter. Um diese Vorhersage zu verbessern, wird nun ein Black-Box-Anteil beim Achsstrommodell hinzugefügt.

### 3 Neuronale Netze zur Restunsicherheitsabschätzung

Neuronale Netze wie rekurrente neuronale Netze (RNN) oder insbesondere Long Short-Term Memory-Netzwerke (LSTM) haben sich als leistungsfähige Ansätze für die Modellierung und Verarbeitung zeitabhängiger Daten etabliert

[11, 12]. Denkna et. al. [13] zeigen das Potenzial eines LSTM-basierten Ansatzes zur Modellierung der Zerspankräfte auf, wobei als Eingangsgrößen verschiedene Antriebsgrößen wie Position, Geschwindigkeit, Beschleunigung, Achsstrome, Spindeldrehzahl und Spindelstrom berücksichtigt werden. Im Vergleich zu einem modellbasierten Ansatz auf Basis des Achsstrommodells können mit dem LSTM insbesondere dynamische Prozessbereiche, etwa Richtungswechsel oder Änderungen in der Vorschubbewegung, präziser abgebildet werden [13]. Darüber hinaus zeigen Zhang et. al. [14], dass auch nicht-invasive Sensorsignale wie Spindelschwingungen in Verbindung mit LSTM eine präzise Vorhersage der Zerspankräfte ermöglichen. Durch die Zuordnung von Positionen zu aufeinanderfolgenden Rechenzeitschritten erzeugen rekurrente Architekturen eine Sequenz von verborgenen Zuständen  $h_{t-1}$ , wobei jeder Zustand als Funktion des vorherigen Zustands sowie der aktuellen Eingabe berechnet wird. Diese sequenzielle Struktur verhindert eine Parallelisierung der Verarbeitung innerhalb eines Trainingsbeispiels, was sich insbesondere bei langen Sequenzen nachteilig auf die Recheneffizienz auswirkt. Speicherbedingte Einschränkungen begrenzen zusätzlich die Möglichkeit zur gleichzeitigen Verarbeitung mehrerer Sequenzen (Batching). Transformer-Modelle, wie sie von Vaswani et. al. [15] eingeführt wurden, adressieren diese Limitation indem rekurrente Strukturen vollständig durch einen Attention-Mechanismus ersetzt werden. Dies ermöglicht eine parallele Verarbeitung von Sequenzen und erlaubt es bei langen Eingaben globale Abhängigkeiten effizient zu modellieren. [16]

Das in dieser Arbeit eingesetzte Modell basiert auf der Transformer-Architektur und verwendet ausschließlich den Encoder, um den Modellfehler des Achsstrommodells im Rahmen einer multivariaten Regressionsaufgabe zu bestimmen. Die Modellstruktur orientiert sich an dem Transformer-Ansatz von Zerveas et. al. [17]. Die Eingangsgrößen des Modells umfassen die Positionen ( $P_x, P_y, P_z$ ), Geschwindigkeiten ( $\dot{P}_x, \dot{P}_y, \dot{P}_z$ ) und Beschleunigungen ( $\ddot{P}_x, \ddot{P}_y, \ddot{P}_z$ ) der Maschinennachsen, den Spindelstrom ( $S_c$ ) sowie die physikalisch modellierten Kraftkomponenten ( $F_x^n, F_y^n, F_z^n$ ). Diese multivariate Zeitreihe

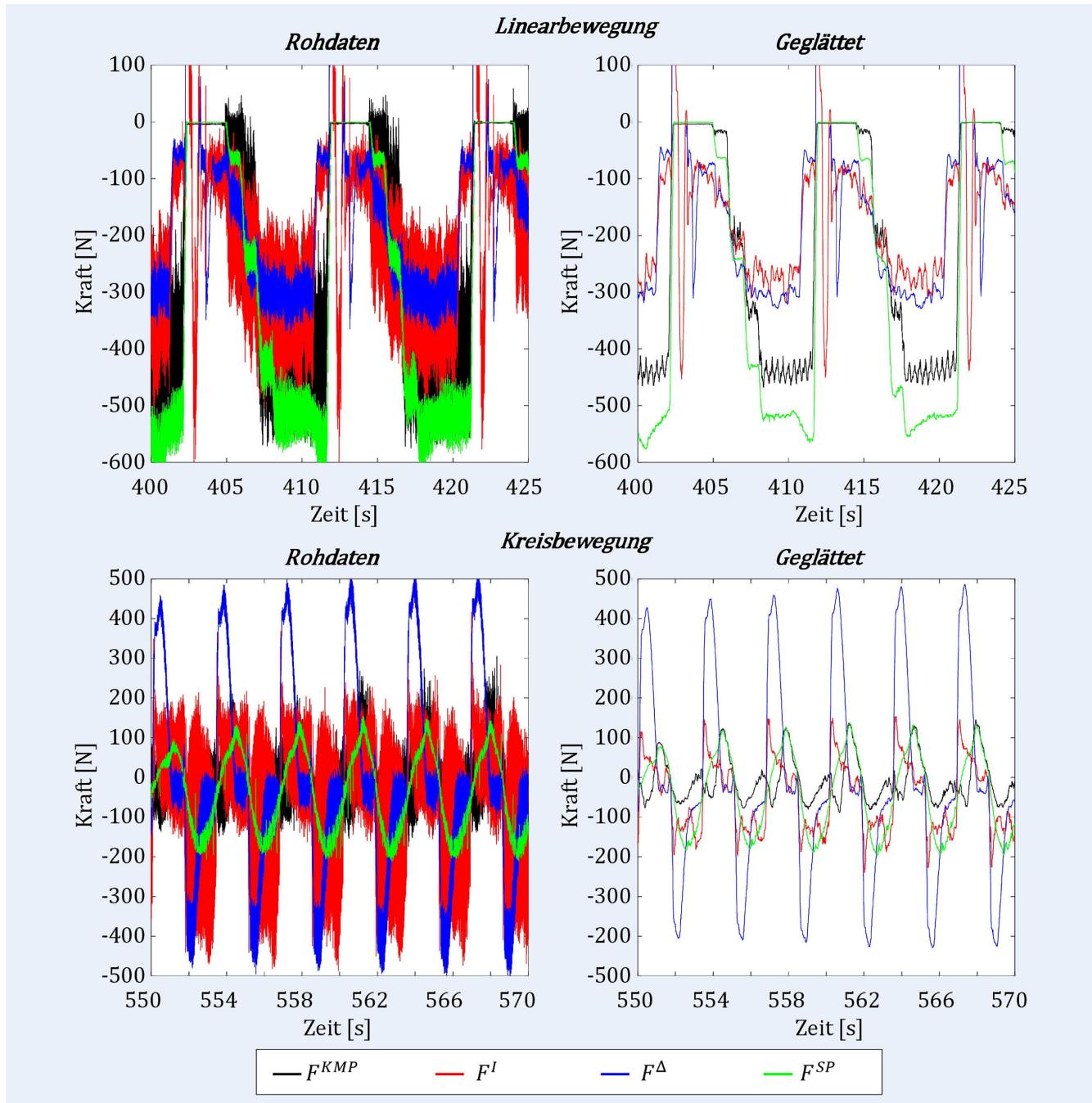


Bild 2. Transformer-Encoder zur Kompensation des Fehlers nach [17]. Grafik: RWTH Aachen

mit  $m$  Kanälen über  $T$  Zeitschritte wird mittels eines Sliding-Window-Verfahrens in überlappende Sequenzen der Länge  $w$  unterteilt. Jede Sequenz lässt sich als Matrix  $\Pi \in \mathbb{R}^{w \times m}$  darstellen und besteht aus einer Folge von Merkmalsvektoren  $\zeta_t \in \mathbb{R}^m$ . Die Merkmalsvektoren  $\zeta_t$  werden normalisiert, indem für jede Dimension der Mittelwert subtrahiert und durch die Standardabweichung im Trainingsdatensatz dividiert wird. Anschließend erfolgt eine lineare Projektion auf einen  $d$ -dimensionalen Modellraum.

$$u_i = W_u \zeta_t + b_b \quad (4)$$

Dabei sind  $Wp \in \mathbb{R}^{dxm}$  und  $b_b \in \mathbb{R}^d$  lernbare Parameter, und  $u_t \in \mathbb{R}^d$  für  $t = 1, \dots, w$  die Eingabevektoren des Modells, die analog

zu den Wortvektoren in NLP-Transformer interpretiert werden können. Da der Transformer keine Informationen über die zeitliche Reihenfolge der Eingabe erhält, wird zur Abbildung der zeitlichen Struktur ein Positional Encoding-Matrix  $W_{pos} \in \mathbb{R}^d$  zu dem Eingabevektor addiert:  $U = U + W_{pos}$ . Die sequenziellen Ausgaben des Transformer-Encoders werden abschließend durch eine vollständig verbundene lineare Schicht transformiert, und so für jeden Zeitschritt eine Regressionsvorhersage erzeugt:

$$\delta_t = W_{out} \cdot h_t + b_{out} \quad (5)$$

Hierbei entspricht der latente Repräsentation des Encoders,  $\delta_t \in \mathbb{R}^g$  der vorherzusagenden Zielgröße,  $W_{out} \in \mathbb{R}^{g \times d}$  der

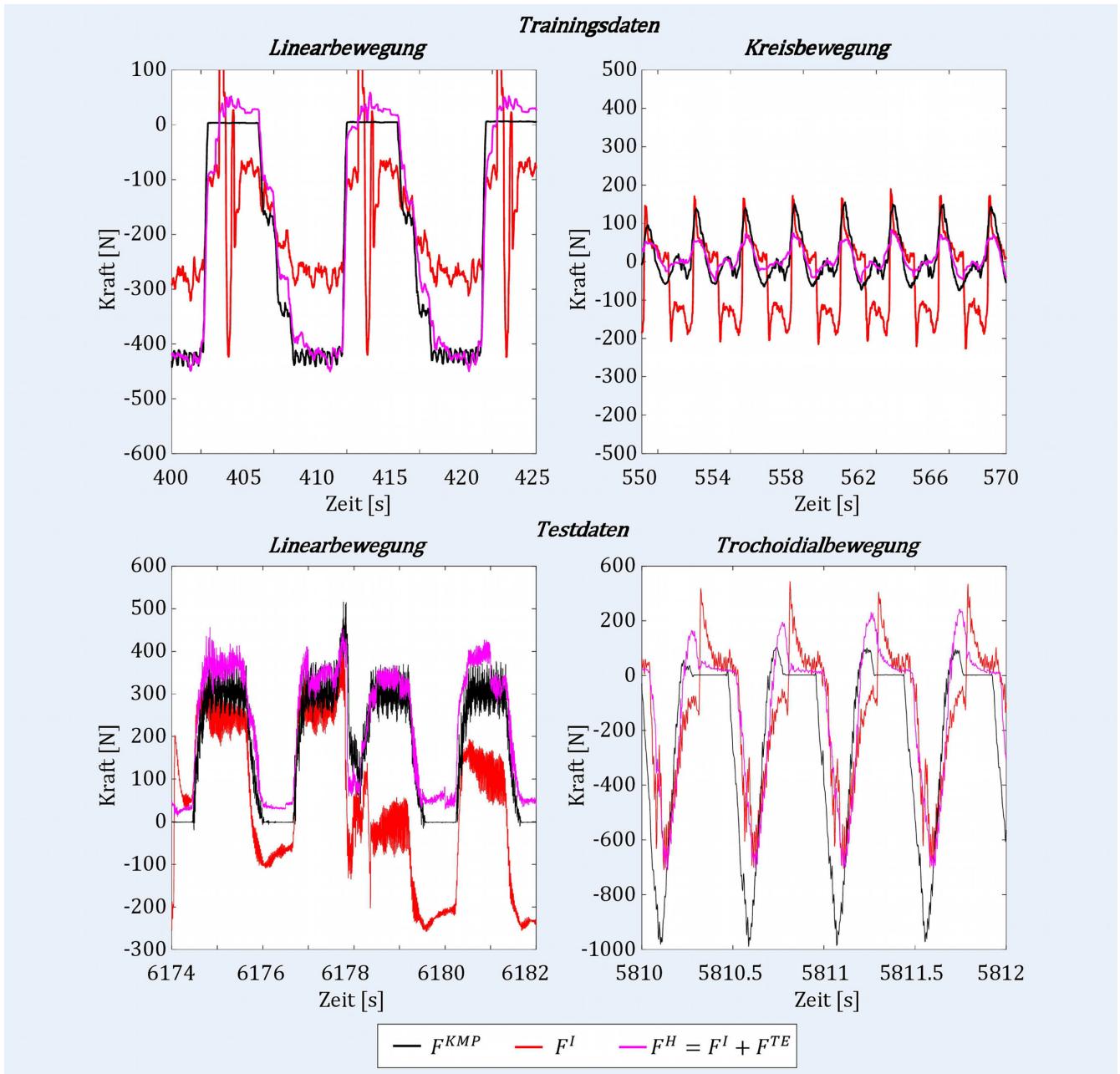


Bild 3. Versuchsbauenteile Grafik: RWTH Aachen

Gewichtsmatrix  $b_{out} \in \mathbb{R}^g$  und dem Bias der Ausgabeschicht. Die sequenzielle Zielgröße  $\Delta \in \mathbb{R}^{w \times g}$  beschreibt den Fehler des physikalischen Achstrommodells über das betrachtete Zeitfenster. Dieser ergibt sich aus der Differenz zwischen den gemessenen Kräften der Kraftmessplattform  $(F_{x_t}^{KMP}, F_{y_t}^{KMP}, F_{z_t}^{KMP})$  und der modellbasierten Kraftvorhersage  $(F_x^m, F_y^m, F_z^m)$ .

#### 4 Versuchsaufbau

Zur Validierung werden mehrere Musterbauteile (vgl. Bild 3) hergestellt. Die Bauteile A1 und A2 haben zum Ziel möglichst viele unterschiedliche Lastfälle abzudecken. Hierzu werden bei der Fertigung des Bauteils die Eingriffstiefe und -breite in Abhängigkeit zu der Zähnezahl, des Dralls und Durchmessers des jewei-

ligen Werkzeugs variiert. Die Fertigung von Bauteil B hingegen wird mit einer hohen Zahl an Kreisbewegungen realisiert. Dies soll dazu beitragen Effekte, die aus der Achsumkehr entstehen, mit zu berücksichtigen. Bauteil C dient zur Validierung und soll zugleich in weiteren Untersuchungen dazu dienen die Qualitätsvorhersage bei dünnwandigen Strukturen zu verbessern. Gefertigt werden die Bauteile aus C45 auf einem Bearbeitungszentrum „DMU65 monoblock“ von DMG und den in Tabelle 1 aufgeführten Werkzeugen. Die Erfassung der Prozesskräfte erfolgt mit einer Kraftmessplattform „9257A“ von Kistler. Äquivalent zu [1] werden die steuerungsinternen Daten mit 500 Hz aufgezeichnet und zu einer Prozesskraft verrechnet. Die Berechnung der einzelnen Korrekturwerte geschieht durch vorherige Messungen mithilfe von Last-Verformungskurven und Luftschnitten.

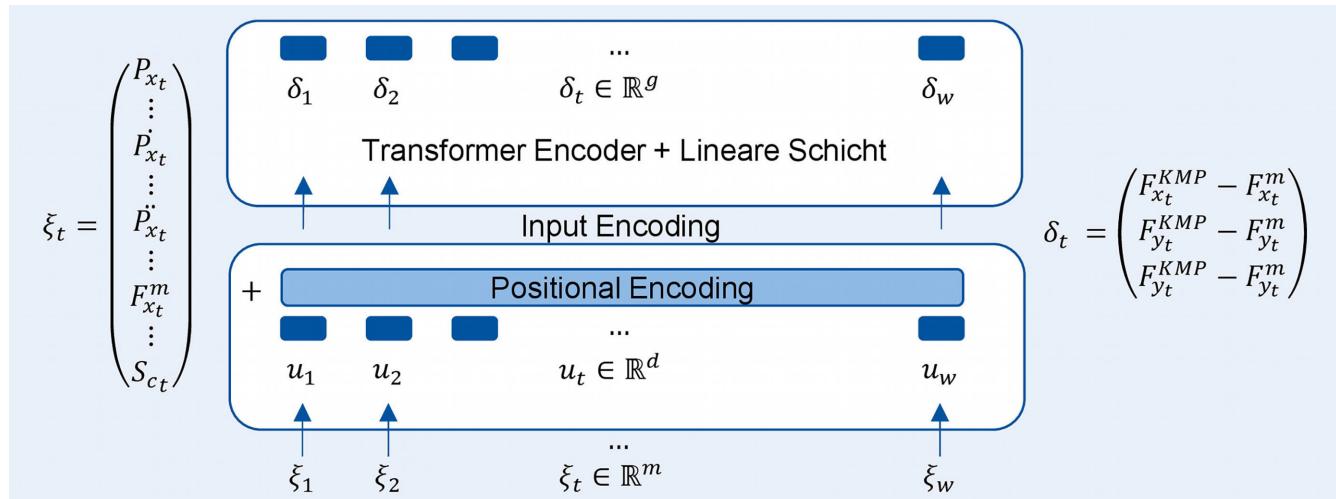


Bild 4. Ergebnisse des Transformer Encoder kompensierten Achsstrommodells. Grafik: RWTH Aachen

Tabelle 1. Versuchswerkzeuge.

Nr.	Wendeschneidplatten/Fräser	Halter*	Durchmesser	Zähnezahl
1	Sandvik 490R-08T308E-ML 1130	Sandvik 490-050HA06-08M	50 mm	5
2	Sandvik 490R-08T308E-ML 1130	Sandvik 490-040HA06-08H	40 mm	6
3	Sandvik 1K324-1600-XB 1730	Haimer A63.140.16	16 mm	4
4	Fraisa P45317610	Haimer A63.140.16.2	16 mm	4
5	Sandvik 1K232-1000-XB 1730	Haimer A63.140.10	10 mm	2
6	Fraisa P45322391	Haimer A63.144.08.3	8 mm	4

\* Zusätzlich wurde bei den Werkzeugen 2-4 neben den genannten Schrumpffuttern das Hydrodehnfutter Schunk „TENDO EC“ von HSK-A63 Ø20x80

Tabelle 2. RMSE der betrachteten Achsstrommodelle.

Vorhersage bas. Achsströmen $F^I$			Hybrides Modell $F^H = F^I + F^{TE}$	
RMSE	Trainingsdaten	Testdaten	Trainingsdaten	Testdaten
$F_x$	155,93 N	154,54 N	55,19 N	93,41 N
$F_y$	166,13 N	196,26 N	62,5 N	71,02 N
$F_z$	185,38 N	267,19 N	44,52 N	42,90 N
$F_{aes}$	293,73 N	365,77 N	94,52 N	124,93 N

## 5 Validierung

Die Validierung erfolgt anhand der gemessenen Kräfte der Kraftmessplattform (KMP). Zur Bewertung der Modellgüte dienen die Fehler des Achsstrommodells  $F^I$  sowie des hybriden Modells ( $F^H$ ), bei dem der verbleibende Modellfehler durch einen Transformer-Encoder (TE) kompensiert wird. Zunächst werden hierzu die Verläufe miteinander verglichen sowohl für den Trainings- als auch Testdatensatz. Anschließend erfolgt die Auswertung über den Root Mean Square Error (RMSE). Der Trainingsdatensatz umfasst die Bauteile A1, A2 und B, während der Testdatensatz aus Bauteil C sowie Luftschnitten der jeweiligen Prozesse besteht.

Der Vergleich der Verläufe zwischen dem Referenzdatensatz der Kraftmessplattform, dem Achsstrommodell und dem durch das TE kompensierten Modell wird erkennbar, dass beim Trainingsdatensatz eine signifikante Verbesserung sowohl bei der Linear- als auch der Kreisbewegung erzielt wird (vgl. oberer Teil Bild 4). Bei den Testdaten zeigt sich bei den Linearbewegungen ebenfalls eine Verbesserung des Achsstrommodells. Vor allem der Drift in den Achsströmen wird hier gut kompensiert. Bei den trochoidalen Bewegungen wird der Verlauf ebenfalls gut korrigiert, allerdings wird die Amplitude in der Kraftspitze unterschätzt.

Tabelle 2 bestätigt die Reduktion des Modellfehlers durch das hybride Modell ( $F^H$ ), bei dem der verbleibende Fehler des achsstrombasierten Modells ( $F^I$ ) durch den TE-Anteil  $F^{TE}$  kompensiert wird. Im Trainingsdatensatz sinkt der RMSE der Gesamt-

kraft  $F_{aes}$  von 293,73 N auf 94,52 N (−67,8 %), im Testdatensatz von 365,77 N auf 124,93 N (−65,8 %). Entsprechend wird der Fehler des Achsstrommodells  $F^1$  um 66,7 % reduziert.

## 6 Zusammenfassung & Ausblick

Die Erklärbarkeit und Verlässlichkeit von Modellausgaben sind im industriellen Kontext von zentraler Bedeutung. Eine Kenngröße für die Prozessüberwachung ist die Zerspankraft. Ihre präzise Modellierung stellt jedoch aufgrund komplexer Einflüsse, wie des Reibverhaltens, eine Herausforderung dar. Durch die Kombination physikalischer Modelle mit datengetriebenen Ansätzen zur Fehlerkompensation lassen sich diese komplexen Einflüsse ausgleichen. Da dabei ausschließlich der verbleibende Modellfehler korrigiert wird, bleibt das physikalisch fundierte Modellverständnis erhalten. Vor diesem Hintergrund verknüpft der vorliegende Ansatz ein physikalisches Modell mit einem datengetriebenen Modell auf Basis eines Transformer-Encoders (TE). Die Validierung zeigt, dass der TE den verbleibenden Fehler des physikalischen Modells signifikant reduzieren kann und über unterschiedliche Bauteilgeometrien und Verfahrbewegungen hinweg generalisiert. Insgesamt wurde eine Fehlerreduktion von über 66 % erreicht.

In weiteren Arbeiten soll der vorgestellte Ansatz mittels Transfer-Learning auf andere Maschinen übertragen werden, um die Übertragbarkeit und Generalisierbarkeit zu evaluieren. Neben hybriden Ansätzen, die physikalisches und datengetriebenes Wissen kombinieren, rücken zunehmend vollständig datengetriebene Verfahren wie das Self-Supervised Learning (SSL) in den Fokus. Diese bieten das Potenzial, beispielsweise über Masked Autoencoder, aussagekräftige Repräsentationen direkt aus unlabeled Prozessdaten zu extrahieren. [18, 19] Auf dieser Grundlage können KI-Modelle entwickelt werden, die robuster gegenüber veränderten Maschinenkonfigurationen und Prozessbedingungen sind. Perspektivisch ermöglicht dies eine verbesserte Generalisierbarkeit und Skalierbarkeit datenbasierter Modelle, wodurch sich der Aufwand für maschinenspezifische Anpassungen deutlich reduzieren lässt.

### FÖRDERHINWEIS

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) –  
Projektnummer 524834456.

### LITERATUR

- [1] Königs, M.; Brecher, C.: Process-parallel virtual quality evaluation for metal cutting in series production. *Procedia Manufacturing* (2018) 26:1087–1093. doi:10.1016/j.promfg.2018.07.145
- [2] Fertig, A.; Weigold, M.; Chen, Y.: Machine Learning based quality prediction for milling processes using internal machine tool data. *Advances in Industrial and Manufacturing Engineering* 4:100074. (2022) doi:10.1016/j.aime.2022.100074
- [3] Wellmann, F.: Datengetriebene, kontextadaptive Produktivitätssteigerung von NC-Zerspanprozessen, 1. Aufl. Aprimus Wissenschaftsverlag, Aachen (2019)
- [4] Bilgili, D.; Kecibas, G.; Besirova, C.; Chehrehzad, M.R.; Burun, G.; Pehlivan, T.; Uresin, U.; Emekli, E.; Lazoglu, I.: Tool flank wear prediction

using high-frequency machine data from industrial edge device. *Procedia CIRP* 118:483–488. (2023) doi:10.1016/j.procir.2023.06.083

- [5] Butler, Q.; Ziada, Y.; Stephenson, D.; Andrew Gadsden, S.: Condition Monitoring of Machine Tool Feed Drives: A Review. *J. Manuf. Sci. Eng* 144 (10) (2022). doi:10.1115/1.4054516
- [6] Yamada, Y.; Kakinuma, Y. (2016) Sensorless cutting force estimation for full-closed controlled ball-screw-driven stage. *Int J Adv Manuf Technol* 87(9-12):3337–3348. doi:10.1007/s00170-016-8710-5
- [7] Aslan, D.; Altintas, Y.: Prediction of Cutting Forces in Five-Axis Milling Using Feed Drive Current Measurements. *IEEE/ASME Trans. Mechatron.* 23 (2018) 2, S. 833–844. doi:10.1109/TMECH.2018.2804859
- [8] Saraie, H.; Sakahira, M.; Ibaraki, S.; Matsubara, A.; Kakino, Y.; Fujishima, M.: 326 Monitoring and Adaptive Control of Cutting Forces Based on Spindle Motor and Servo Motor Currents in Machining Centers. *LEM21* 2003 (0) S. 555–560. doi:10.1299/jsmelem.2003.555
- [9] Mostaghimi, H.; Park, C.; Kang, G.; Park, S.S.; Lee, D. Y.: Reconstruction of cutting forces through fusion of accelerometer and spindle current signals. *Journal of Manufacturing Processes* 68:990–1003, (2021) doi:10.1016/j.jmapro.2021.06.007
- [10] Fey, M.; Epple, A.; Kehne, S.; Brecher, C.: Verfahren zur Bestimmung der Achslast auf Linear- und Rundachsen G01L 1/04, (2016)
- [11] Hochreiter, S.; Schmidhuber, J. (1997) Long short-term memory. *Neural Computation* 9(8):1735–1780. doi:10.1162/neco.1997.9.8.1735
- [12] Pankaj Malhotra, Lovekesh Vig, Gautam Shroff, Puneet Agarwal: Long short term memory networks for anomaly detection in time series. *Proceedings 2015* (89), S. 89–94
- [13] Denkena, B.; Bergmann, B.; Stoppel, D.: (Hrsg) Reconstruction of Process Forces in a Five-Axis Milling Center with a LSTM Neural Network in Comparison to a Model-Based Approach. Springer, Cham, 2023
- [14] Li, B.; Liu, T.; Liao, J.; Feng, C.; Yao, L.; Zhang, J.: Non-invasive milling force monitoring through spindle vibration with LSTM and DNN in CNC machine tools. *Measurement* 210:112554. doi:10.1016/j.measurement.2023.112554, 2023)
- [15] Aswani, A.; Noam, S.; Parmar, N.; Uszkoreit, J.; Jones, L.; Aidan, N.; Kaiser, G. L.; Polosukhin, I.: Attention is all you need. *Advances in neural information processing systems*, 2017
- [16] Li, Shiyang, Xiaoyong Jin, Yao Xuan, Xiyou Zhou, Wenhui Chen, Yu-Xiang Wang, and Xifeng Yan: Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting. *Advances in neural information processing systems* 32:1–11 (2019)
- [17] Zerveas, G.; Jayaraman, S.; Patel, D.; Bhamidipaty, A.; Eickhoff, C.: A Transformer-based Framework for Multivariate Time Series Representation Learning (2020)
- [18] Li, Z.; Rao, Z.; Pan, L.; Wang, P.; Xu, Z.: Ti-MAE: Self-Supervised Masked Time Series Autoencoders (2023)
- [19] Cheng, M.; Liu, Q.; Liu, Z.; Zhang, H.; Zhang, R.; Chen, E.: TimeMAE: Self-Supervised Representations of Time Series with Decoupled Masked Autoencoders (2023)

Prof. Dr.-Ing. Christian Brecher

Dr.-Ing. Marcel Fey

Matthäus Loba, M.Sc.   
m.loba@wzl.rwth-aachen.de

Johannes Vos, M.Sc.

Werkzeugmaschinelabor WZL der RWTH Aachen  
Steinbachstr. 19, 52074 Aachen  
www.wzl.rwth-aachen.de

### LIZENZ



Dieser Fachaufsatz steht unter der Lizenz Creative Commons Namensnennung 4.0 International (CC BY 4.0)